Correlated Idiosyncratic Volatility Shocks

Xiao Qiao[*,a] and Yongning Wang[†,b]

[a]City University of Hong Kong
[b]Booth School of Business, University of Chicago

February 10, 2021

Abstract

The commonality in idiosyncratic volatility cannot be fully explained by time-varying volatility; correlated idiosyncratic volatility shocks are an important contributing factor. We empirically document this fact for various characteristic-sorted portfolios and industry portfolios. To capture the commonality in idiosyncratic volatility, we propose a novel multivariate GARCH model called Dynamic Factor Correlation. The DFC model has a closed-form likelihood function which allows computationally cheap estimation, even for a large number of assets. The DFC model improves statistical fit compared to existing multivariate GARCH models such as the Dynamic Conditional Correlation model and the Dynamic Equicorrelation model, achieving the lowest root mean square error and mean absolute error in simulations. Empirical tests also corroborate simulations in demonstrating the improved statistical fit of the DFC model. Mean-variance portfolio optimization using the DFC model outperforms alternative volatility models such as the historical covariance matrix and the DECO model. Out-of-sample mean-variance efficient portfolios using the DFC have the lowest volatility and highest Sharpe ratios, thereby improving the investor's opportunity set. Under parametric restrictions, the Dynamic Factor Correlation model reduces to the Constant Conditional Correlation model or the Dynamic Equicorrelation model.

[*]Corresponding author. Address: 83 Tat Chee Ave, Kowloon Tong, Hong Kong Email: xiaoqiao@cityu.edu.hk Website: https://sites.google.com/site/xiaoqiao10/

[†]Address: 250 W 100th St, New York, New York 10025. Email: ywang1@chicagobooth.edu

Keywords: volatility, GARCH, stock returns, idiosyncratic risk
Classification: C58, G12, G17

List of Tables and Figures

1 Introduction

The behavior of volatility is of central interest for financial economists and portfolio managers because of its importance for volatility modeling, asset pricing, and risk management. Much of volatility research has focused on the total volatility of individual securities or the aggregate stock market, relegating the role of idiosyncratic volatility as secondary. Theoretical justification for ignoring idiosyncratic volatility is often made through diversification. Indeed, Markowitz (1952) demonstrates that in well-diversified portfolios, idiosyncratic risk should provide no additional compensation for the investor because its contribution to overall portfolio risk is negligible.

If investors are not sufficiently diversified, the undiversified component of idiosyncratic volatility could impact the distribution of portfolio returns (Merton, 1987). In a notable recent study, Herskovic, Kelly, Lustig, and Van Nieuwerburgh (2016) show that idiosyncratic volatility contains a common factor structure. Rather than being reduced to insignificant by diversification, idiosyncratic volatilities comove together and contribute significantly towards portfolio risk. Herskovic et al. (2016) demonstrate that Merton's (1987) idea is not just theoretical fascination, but has empirical merit.

How do we model correlated idiosyncratic volatility? If we consider idiosyncratic volatility as the combination of a predictable component and an unexpected component, the comovement in idiosyncratic volatility could arise from either part. For the predictable component, there may be comovement in expected time-varying idiosyncratic volatility which drives the common factor in idiosyncratic volatility. Alternatively, if the predictable component does not exhibit comovement, unexpected idiosyncratic volatility shocks could be correlated in the cross section which would also result in a common factor structure in the overall idiosyncratic volatility. We want to distinguish between these two channels because they can help us understand how to model correlated idiosyncratic volatility.

Herskovic et al. (2016) establish a link between commonality in idiosyncratic volatility and shocks to firm-level idiosyncratic risk. They also show that the innovations to a common idiosyncratic volatility factor is priced in the cross section. These findings suggest that the unexpected idiosyncratic volatility shocks, rather than expected time-varying idiosyncratic volatility, drive the common factor in idiosyncratic volatility.

In this paper, we find that the commonality in idiosyncratic volatility cannot be fully explained by expected time-varying idiosyncratic volatility; correlated idiosyncratic volatility shocks are necessary to capture such commonality. We estimate factor models on characteristic-sorted portfolios to obtain residual returns. Then we fit univariate GARCH models to the residual returns. If expected time-varying idiosyncratic volatility completely explains the commonality in idiosyncratic volatility, we would expect the GARCH standardized residuals, free from time-varying volatility effects, to be cross-sectionally uncorrelated. Rather than having pairwise correlations of zero, the GARCH residuals are positively correlated in each set of portfolios, with average correlations ranging from 0.10 to 0.52. Correlated idiosyncratic volatility shocks are an important contributing factor to the commonality in idiosyncratic volatility.

Our empirical finding provides a challenge to existing multivariate GARCH models. The Dynamic Conditional Correlation (DCC) of R. Engle (2002) is designed to capture time-varying volatilities and correlations. The Dynamic Equicorrelation (DECO) of R. Engle and Kelly (2012) is another popular model for time-varying correlations. However, neither accounts for correlated idiosyncratic volatility shocks.

We propose a new multivariate GARCH model, Dynamic Factor Correlation (DFC), to capture the commonality in idiosyncratic volatility. Compared to DCC and DECO, the DFC directly model correlated idiosyncratic volatility innovations. In the DFC model, GARCH standardized residuals are driven by a common factor that induces cross-sectional comovement in standardized residuals, leading to comovement in idiosyncratic volatility.

The DFC model is closely related to the DCC model. Both the DCC and DFC seek to capture multivariate relationships after removing univariate volatility effects using GARCH models. Whereas R. Engle (2002) uses pairwise GARCH residuals to estimate time-varying correlations, we construct a common factor that drives the correlations of GARCH residuals. The DFC model is also related to the Constant Conditional Correlation (CCC) of Bollerslev (1990) and DECO of R. Engle and Kelly (2012). Under the assumption that the common factor in idiosyncratic volatility innovations is homoscedastic, the DFC model reduces to the CCC. DECO assumes all pairwise correlations are equal, whereas the DFC model allows for different pairwise correlations. Under the restriction of equal factor loadings on the common factor in idiosyncratic volatility innovations, the DFC model becomes the DECO model.

We use a two-stage quasi-maximum likelihood (QML) estimator for the DFC, which is consistent and asymptotically normal under regularity conditions. From the Sherman and Morrison (1950) formula, the likelihood function of the DFC model has a closed-form expression. As such, estimation is straightforward and computationally cheap even for a large number of assets. In contrast, the DCC requires numerical optimization and may be difficult to estimate for a large cross section.

We conduct simulations to compare the model performance of the DFC and DCC. The simulated returns assume correlated idiosyncratic volatility as part of the data-generating process. We compare the two models using root mean squared error (RMSE) and mean absolute error (MAE). The DFC model outperforms the DCC model under different parameter values for the size of the cross section, the number of time periods, and the strength of comovement in idiosyncratic volatility. A greater number of assets is associated with a greater reduction in RMSE and MAE for the DFC compared to the DCC. Stronger comovement in idiosyncratic volatility is associated with a larger improvement in RMSE and MAE for the DFC relative to the DCC.

Empirical tests corroborate simulations in demonstrating the improved statistical fit of the DFC compared to the DCC. We estimate the Fama and French (1993) model to seven characteristic-sorted portfolios, and we fit DFC, DCC, DECO models to the residual returns. Comparing the likelihood values, the DFC model has a better statistical fit, compared to the DCC and DECO, for all seven sets of portfolios. The DCC model compared favorably against the DECO model and outperforms in six of the seven sets of portfolios. We also propose an extension of the DFC, the Group DFC, to reduce the number of parameters for a more parsimonious model. The Group DFC model performs even better compared to the DFC on characteristic-sorted portfolios; the Bayesian Information Criterion (BIC) of the Group DFC is consistently lower compared to that of the DFC.

We apply the DFC model to construct optimal mean-variance portfolios. Mean-variance efficient portfolios are formed using 49 industry portfolios, and different volatility models are used to estimate the covariance matrix. We consider 12 volatility models in three classes. The first class includes two volatility models fit to raw returns, the historical covariance matrix and the DFC. The second class includes five models that fit the Capital Asset Pricing Model (CAPM) to returns, then DFC, Group DFC, or DECO to the residual returns of the CAPM. The third class includes five models that fit the Fama and French (1993) model to returns, then DFC, Group DFC, or DECO to the residual returns.

For all volatility models, we form out-of-sample mean-variance efficient portfolios by iteratively solving volatility-minimization problems given portfolio expected returns target of 5% or 10%. Each month, we form mean-variance portfolios based on each volatility model, and we track their performance for the following month. Compared to portfolios formed using the historical covariance matrix, portfolios formed using the DFC or Group DFC have lower volatility and higher Sharpe ratios. Similarly, portfolios formed using the DFC or Group DFC have lower volatility and higher Sharpe ratios than those using the DECO model, although the improvements are smaller than those comparisons to the historical covariance matrix. The DFC and Group DFC models can help expand the investor's opportunity set with out-of-sample portfolios that have more attractive risk-return trade-off.

Our paper fits into the literature studying the properties of idiosyncratic volatility. Herskovic et al. (2016) document strong comovement in idiosyncratic volatility, and show that a common factor is priced in the cross section. Duarte, Kamara, Siegel, and Sun (2014) examine the cross-sectional properties of the principal components of idiosyncratic volatility. Compared to these papers, we view the comovement in idiosyncratic volatility in terms of contributions from two channels, time-varying idiosyncratic volatility and correlated idiosyncratic volatility shocks, and we propose a volatility model to capture them.

Our paper is also related to the GARCH literature. R. F. Engle (1982) and Bollerslev (1986) provide the classic univariate models for time-varying volatility. Bollerslev (1990) extends GARCH into the multivariate setting with constant pairwise correlations. R. Engle (2002) puts forward a workhorse multivariate GARCH model, the Dynamic Conditional Correlation, to model time-varying correlations. R. Engle and Kelly (2012) simplify the DCC model by assuming all pairwise correlations are equal. These models are among the most popular used in volatility modeling, but they do not capture the common factor in idiosyncratic volatility. The DFC model is designed for this common factor and outperforms the DCC and DECO models in simulations and empirical tests.

The paper is organized as follows. Section 2 provides evidence that idiosyncratic volatility innovations are correlated in the cross section. Section 3 introduces the DFC model and examines its empirical performance. Section 4 applies the DFC model to portfolio construction. Section 5 concludes.

2 Correlated Idiosyncratic Volatility Shocks

2.1 Data

For our analysis, we use daily and monthly returns for Fama and French (1993); Fama and French. (2015) factors and characteristic-sorted portfolios from Ken French's website[1]. We obtain daily and monthly returns on the Fama and French (1993) factors from July 1926 through August 2015, and Fama and French. (2015) factors from July 1963 to August 2015. We use monthly returns on deciles formed on market equity (ME), the book-to-market ratio (BE/ME), long-term reversal (LT Rev), operating profitability (OP), investment (Inv), momentum (Mom), and short-term reversal (ST Rev), and double-sorted five-by-five portfolios based on these characteristics.

The data have different start dates: ME and BE/ME start in July 1926; OP and Inv start in July 1963; LT Rev starts in January 1927; ST Rev starts in February 1926; Mom starts in January 1931. The bivariate sorts have monthly returns starting at the later date of the two characteristics. We also obtain 49 industry portfolio returns from July 1926 through August 2015 from French's website.

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

2.2 Idiosyncratic Volatility Innovations are Correlated

Herskovic et al. (2016) demonstrate that idiosyncratic volatility contains a factor structure: after removing common factors from returns, idiosyncratic volatilities still tend to comove together. What accounts for this comovement? There are two possible channels. First, predictable comovement in time-varying idiosyncratic volatility could lead to the observed factor structure. Second, unpredictable idiosyncratic volatility innovations could be driven by a common factor. We seek to understand the relative importance of these two channels, as they have different implications for portfolio risk management.

Let $r_{i,t}$ be the return of security i at time t, which follows factor model,

$$r_{i,t} \quad = \quad \alpha_i + \mathbf{f}_t' \boldsymbol{\beta}_i + a_{i,t}, \tag{1}$$

where α_i is the intercept, $\mathbf{f}_t = (f_{t,1}, \cdots, f_{t,K})'$ is a vector of K factors, $\boldsymbol{\beta}_i$ is a vector of K factor loadings, and $a_{i,t}$ is the residual return. $h_{i,t}$ is the expectation of squared residual return conditional on the $t-1$ information set, $h_{i,t} \equiv \mathbb{E}_{t-1}[a_{i,t}^2]$. Herskovic et al. (2016) show that the volatility of $a_{i,t}$ is correlated across i's.

We investigate the relative importance of expected time-varying idiosyncratic volatility compared with correlated idiosyncratic volatility shocks. Our analysis starts with univariate and bivariate characteristic-sorted portfolios and industry portfolios. Portfolio returns are diversified relatively to individual stock returns, so the idiosyncratic component of returns is smaller. However, small idiosyncratic returns do not imply uncorrelated idiosyncratic volatility. Herskovic et al. (2016) find correlated idiosyncratic volatility for size and value-sorted Fama-French portfolios, indicating this phenomenon is not limited to individual stocks. Following Herskovic et al. (2016), we first fit factor models to capture return variation, using the Fama and French (1993) factors or principal components in place of $\mathbf{f}_t$ in Equation (1). We then explore the behavior of residual returns $a_{i,t}$. Idiosyncratic volatility is the volatility of these residual returns.

Similar to Herskovic et al. (2016), we find that the volatility of residual returns to be strongly correlated. It is not the common factors that drive the comovement in idiosyncratic volatility. Whether we use the Fama and French (1993) model or principal components to capture return variation, the volatility of residual returns are correlated. To isolate the effect of time-varying idiosyncratic volatility, we use GARCH models to capture the expected component of idiosyncratic volatility.

We fit univariate GARCH models for each residual return series. For a well-specified GARCH model, its standardized GARCH residuals should appear to be independently and identically distributed with unit variance. We search over different combinations of GARCH(p,q) to pick the model based on the best GARCH residual behavior. GARCH models capture expected time-varying idiosyncratic volatility, and remove its influence on the common factor in idiosyncratic volatility. If time-varying idiosyncratic volatility is the main driver for the comovement in idiosyncratic volatility, standardized residuals, free from time-varying volatility effects, should be cross-sectionally uncorrelated.

Table 1 presents summary statistics for the pairwise correlations between GARCH standardized residuals. The middle three columns report correlations among GARCH residuals of residual returns from the Fama and French (1993) model. Rather than being uncorrelated, GARCH residuals in idiosyncratic volatility are positively correlated. For example, the range in correlation for ME deciles is 0.05 to 0.52, with an average value of 0.19. The range for BE/ME is 0.02 to 0.22 with an average value of 0.12. The average pairwise correlation for univariate portfolios is positive for all seven characteristics.

It is possible that the Fama and French (1993) factors do not fully capture return variation in these portfolios, so that omitted factors could drive the commonality in residual return volatility. To address this concern, we use principal components to remove any common variation in returns. Principal components are selected such that residual returns are uncorrelated. Herskovic et al. (2016) show that even when residual returns are uncorrelated, the second moments of these returns remain positively correlated. Then, we fit GARCH models to these residual returns and examine the GARCH residual correlations. The right three columns in Table 1 show that even after removing principal components, the GARCH residuals are correlated in the cross section. Correlations for the ME deciles range between 0 and 0.75 with an average of 0.20, and the range for BE/ME portfolios range between 0.03 to 0.72 with an average of 0.20. Other univariate portfolios also show large range and positive averages. Evidently, correlated idiosyncratic volatility innovations are not due to model misspecification from using the Fama and French (1993) model.

The three left columns labeled "Raw Returns" contain results for GARCH models fitted on raw returns and illustrates the importance of first removing common factors. If we do not remove common factors driving returns, even after controlling for time-varying volatility, GARCH residuals are strongly correlated across portfolios for all univariate sorts. For example, correlations for ME deciles range from 0.69 to 0.98 and have an average value of 0.91, whereas correlations for BE/ME deciles range from 0.71 to 0.92 with an average of 0.86.

Bivariate portfolios show the same patterns as univariate portfolios. Before removing the Fama and French (1993) factors, GARCH residuals are strongly correlated. After removing the Fama and French (1993) factors, correlations are lower but still mostly positive. For example, the 25 ME and BE/ME portfolios have correlations ranging from 0.62 to 0.94 before removing common factors.

Table 1: **Correlations of Idiosyncratic Volatility Innovations**
We compute pairwise correlations of GARCH standardized residuals for characteristic-sorted portfolios and industry portfolios. The left three columns show results for fitting GARCH models to raw returns series. The middle three columns contain results for GARCH standardized residuals of idiosyncratic returns from the Fama and French (1993) model. In the right three columns, we first fit principal components to the portfolios such that idiosyncratic returns are uncorrelated, then we examine GARCH standardized residuals of the idiosyncratic returns.

	Raw Returns			FF3F Removed			PCs Removed		
	Univariate Sorts								
	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg
ME	0.98	0.69	0.91	0.52	0.05	0.19	0.75	-0.00	0.20
BE/ME	0.92	0.71	0.86	0.22	0.02	0.12	0.72	0.03	0.20
LT Rev	0.91	0.72	0.84	0.54	0.01	0.24	0.87	0.05	0.22
OP	0.92	0.80	0.87	0.28	-0.02	0.09	0.77	-0.00	0.20
Inv	0.91	0.80	0.87	0.26	-0.01	0.09	0.68	-0.00	0.19
Mom	0.92	0.66	0.82	0.57	0.08	0.28	0.82	-0.01	0.20
ST Rev	0.92	0.72	0.84	0.53	0.03	0.18	0.81	0.06	0.25
	Bivariate Sorts								
25 ME, BE/ME	0.94	0.62	0.83	0.75	-0.04	0.12	0.72	-0.03	0.15
25 ME, OP	0.97	0.66	0.85	0.49	-0.03	0.13	0.48	-0.06	0.08
25 ME, Inv	0.96	0.65	0.85	0.39	-0.05	0.10	0.49	-0.04	0.10
25 ME, Mom	0.96	0.65	0.85	0.71	0.02	0.22	0.48	-0.01	0.12
25 ME, ST Rev	0.95	0.63	0.84	0.63	0.01	0.19	0.57	-0.00	0.12
25 ME, LT Rev	0.91	0.61	0.82	0.50	0.01	0.14	0.48	-0.02	0.10
	Industry Portfolios								
5 Industries	0.86	0.71	0.79	0.35	0.11	0.16	0.95	0.18	0.52
10 Industries	0.88	0.48	0.67	0.36	0.02	0.13	0.86	0.05	0.25

After removing common factors, the range becomes -0.04 to 0.75, with an average value of 0.12. Importantly, average correlations are positive for all six bivariate portfolios, rather than zero. Even after removing principal components, GARCH residuals are still positively correlated across all six sets of portfolios.

Industry portfolios also have the same pattern as univariate and bivariate portfolios. Residual returns from the Fama and French (1993) show correlated idiosyncratic volatility innovations. For 5 industries, correlations of GARCH residuals range from 0.11 to 0.35 with an average value of 0.16. For 10 industries, GARCH residual correlations range from 0.02 to 0.36 with an average of 0.13. GARCH residuals are also positively correlated for idiosyncratic returns from principal components.

If expected time-varying volatility were the sole contributing channel towards the comovement in idiosyncratic volatility, we would expect to see largely uncorrelated GARCH residuals with an average correlation near zero. Positively correlated GARCH residuals shows that time-varying idiosyncratic volatility is not the only important force in driving idiosyncratic volatility comovement; correlated idiosyncratic volatility shocks is also important. Commonality in idiosyncratic volatility cannot be explained by time-varying idiosyncratic volatility alone. We must account for correlated idiosyncratic volatility shocks if we want to model commonality in idiosyncratic volatility.

3 A Model with Correlated Idiosyncratic Volatility

Various multivariate GARCH models have been proposed to study the comovement of asset returns and volatilities. They help us understand the correlation structure of portfolio constituents and improve risk management. We evaluate two models in light of our empirical findings, and propose a new multivariate GARCH model called the Dynamic Factor Correlation.

The Dynamic Conditional Correlation model introduced by R. Engle (2002) allows for time-varying correlations and volatilities. It has become a workhorse model to capture multivariate relationships among portfolio constituents[2]. While the DCC model proves to be a powerful tool, one potential disadvantage of the DCC model is the computing power required to estimate the model for large cross sections. R. Engle and Kelly (2012) propose the Dynamic Equicorrelation model to lower the computational cost by simplifying the correlation structure. They assume all pairwise correlations are equal at any point in time, so the number of parameters is greatly reduced. DCC and DECO models are widely used in academic research and portfolio management.

Commonality in idiosyncratic volatility is a robust empirical fact (Herskovic et al., 2016).

[2]Surveys by Tsay (2006) and Bauwens, Laurent, and Rombouts (2006) provide discussions on DCC and related multivariate GARCH models.

A statistical model that aims to capture time-varying volatility cannot fully do so if it does not account for the common factor in idiosyncratic volatility. Moreover, risk management based on a model that does not capture correlated idiosyncratic volatility may lead the portfolio manager to think the portfolio is more diversified than it really is. DCC and DECO models were not designed to accommodate correlated idiosyncratic volatility.

We propose the Dynamic Factor Correlation model to account for the commonality in idiosyncratic volatility. The DFC model is an extension of Engle's DCC model that introduces a factor structure in the standardized idiosyncratic volatility innovations. Compared to DCC and DECO, DFC improves model fit and outperforms in portfolio construction.

3.1 Dynamic Factor Correlation Model

We start with a factor model as in Equation (1):

$$r_{i,t} = \alpha_i + \mathbf{f}_t' \boldsymbol{\beta}_i + a_{i,t}, \tag{2}$$

Recall $h_{i,t}$ is the conditional expectation of squared residual returns, $h_{i,t} = \mathbb{E}_{t-1}[a_{i,t}^2]$. The volatility process of $h_{i,t}$ can be specified by the econometrician, such as the GARCH(1,1) model of Bollerslev (1986). $e_{i,t} \equiv a_{i,t}/\sqrt{h_{i,t}}$ is the standardized residual with $\mathbb{E}[e_{i,t}^2] = 1$.

Define $\mathbf{a}_t = (a_{1,t}, \cdots, a_{N,t})'$, $\mathbf{e}_t = (e_{1,t}, \cdots, e_{N,t})'$, and $\mathbf{D}_t = diag\{\sqrt{h_{1,t}}, \cdots, \sqrt{h_{N,t}}\}$. The relationship between residuals and standardized innovations can be represented in matrix form

$$\mathbf{a}_t = \mathbf{D}_t \mathbf{e}_t \tag{3}$$

Motivated by our empirical finding that idiosyncratic volatility innovations are correlated, we impose a factor structure on the standardized residuals $e_{i,t}$

$$e_{i,t} = \frac{q_{i,t}}{s_{i,t}} \tag{4}$$

where

$$\begin{aligned} q_{i,t} &= v_t \xi_i + \sigma_i \epsilon_{i,t}, \\ s_{i,t}^2 &= \mathbb{E}_{t-1}[q_{i,t}^2] \end{aligned} \tag{5}$$

v_t is a common factor in idiosyncratic volatility innovations that affects $q_{i,t}$ for all i, and ξ_i is the factor loading of security i on this common factor. We assume $v_{t|t-1} \sim \mathcal{N}(0, h_{v,t})$ where

$h_{v,t} \equiv \mathbb{E}_{t-1}[v_t^2]$, $\mathbb{E}[v_t] = 0$, and $\mathbb{E}[v_t^2] = 1$. $\sigma_i \epsilon_{i,t}$ is the idiosyncratic component of $q_{i,t}$ not driven by a common factor. $\epsilon_{i,t} \sim i.i.d.\,\mathcal{N}(0,1)$. σ_i^2 satisfies $\sigma_i^2 = 1 - \xi_i^2$ with $|\xi_i| < 1\,\forall i$ so that $\mathbb{E}[q_{i,t}^2] = 1$.

Under Equation (5), the conditional variance and covariance for $q_{i,t}$ follow these processes:

$$\begin{aligned} var_{t-1}[q_{i,t}] &= 1 + (h_{v,t} - 1)\xi_i^2 \\ cov_{t-1}(q_{i,t}, q_{j,t}) &= \xi_i \xi_j h_{v,t} \end{aligned} \tag{6}$$

Let $\mathbf{Q}_t = var_{t-1}(\mathbf{q}_t)$, we can write Equation (6) in matrix form.

$$\mathbf{Q}_t = \mathbf{\Lambda} + h_{v,t} \cdot \boldsymbol{\xi}\boldsymbol{\xi}', \tag{7}$$

where $\mathbf{\Lambda}$ is a $N \times N$ diagonal matrix with $\mathbf{\Lambda}(i,i) = 1 - \xi_i^2$, and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_N)'$. Let $\tilde{\mathbf{Q}}_t$ be a diagonal matrix with the same diagonal elements as $\mathbf{Q}_t$, i.e. $\tilde{\mathbf{Q}}_t(i,i) = 1 + (h_{v,t} - 1)\xi_i^2$, and $\tilde{\mathbf{Q}}_t(i,j) = 0$ for $i \neq j$. The correlation matrix of standardized residuals e_t is given by

$$\begin{aligned} \mathbf{R}_t &= cor_t(\mathbf{e}_t) \\ &= \tilde{\mathbf{Q}}_t^{-\frac{1}{2}} \mathbf{Q}_t \tilde{\mathbf{Q}}_t^{-\frac{1}{2}} \end{aligned} \tag{8}$$

The correlation between standardized residuals i and j, $\rho_t(i,j)$, can be expressed as the following

$$\rho_t(i,j) = \frac{\xi_i \xi_j}{\sqrt{\xi_i^2 + \frac{(1-\xi_i^2)}{\sqrt{h_{v,t}}}}\sqrt{\xi_j^2 + \frac{(1-\xi_j^2)}{\sqrt{h_{v,t}}}}} \tag{9}$$

The correlations between squared standardized residuals i and j are:

$$cor_t(e_{i,t-1}^2, e_{j,t-1}^2) = \frac{\xi_i^2 \xi_j^2}{[\xi_i^2 + \frac{1-\xi_i^2}{\sqrt{h_{v,t}}}][\xi_j^2 + \frac{1-\xi_j^2}{\sqrt{h_{v,t}}}]} \tag{10}$$

Equations (9) and (10) imply that an increase $h_{v,t}$ leads to stronger pairwise correlations in standardized residuals[3] and squared standardized residuals. Among others, Longin and Solnik (2001), Ang and Chen (2002), and Cappiello, Engle, and Sheppard (2006) have documented that during volatile periods when volatility innovations tend to be larger, pairwise security-level correlations tend to be higher as well. Our model captures this feature for idiosyncratic volatility.

A key distinction between the DFC model and the DCC model of R. Engle (2002) lies in

[3]Provided the two factor loadings of the same sign

the dynamics of $\mathbf{Q}_t$. $\mathbf{Q}_t$ processes considered in DCC often include an exponential smoother or the MARCH model of Ding and Engle (2001), both with explicit dependence on $\mathbf{Q}_{t-1}$. The DFC model uses the factor structure in Equation (7). The evolution of $\mathbf{Q}_t$ is driven by the conditional variance of the common factor in idiosyncratic volatility innovations, $h_{v,t}$, and does not contain explicit dependence on $\mathbf{Q}_{t-1}$.

With additional restrictions, the DFC model reduces to existing multivariate GARCH models. If v_t were homoskedastic, i.e., $\forall t, h_{v,t} = 1$, $\mathbf{R}_t$ will be equal to $\mathbf{Q}_t$ where $\mathbf{R}_t(i,i) = 1$ and $\mathbf{R}_t(i,j) = \xi_i \xi_j$. The DFC model then becomes the Bollerslev (1990) Constant Conditional Correlation model with

$$var_{t-1}(\mathbf{a}_t) = \mathbf{D}_t \mathbf{R} \mathbf{D}_t \tag{11}$$

which has a static correlation matrix

$$\mathbf{R}(i,j) = \xi_i \xi_j \tag{12}$$

Under the restriction that all securities' factor loadings on ν_t were equal: $\xi_i = \xi_j \equiv \tilde{\xi}$, the DFC model reduces to the Dynamic Equicorrelation model in R. Engle and Kelly (2012). The correlation matrix has the form

$$\mathbf{R}_t = (1 - \rho_t)\mathbf{I}_N + \rho_t \mathbf{J}_N \tag{13}$$

with all pairwise correlation equal to

$$\rho_t = \frac{h_{v,t}\tilde{\xi}^2}{1 + (h_{v,t} - 1)\tilde{\xi}^2} \tag{14}$$

where $\mathbf{I}_N$ is the N-dimensional identity matrix and $\mathbf{J}_N$ denotes a $N \times N$ matrix of ones.

3.2 Estimation

The DFC model has a closed-form likelihood for estimation based on the Sherman and Morrison (1950) formula (see Appendix A for details), which makes estimation computationally inexpensive. We propose a two-stage quasi-maximum likelihood estimator for the DFC model. It is consistent and asymptotically normal under regularity conditions (see Appendix B), despite the possibility of a misspecified model.

The (scaled) log likelihood $\mathcal{L}$ for the estimator of DFC model can be expressed as

$$\mathcal{L} \quad = \quad -\frac{1}{T}\sum_t \Big(\log|\mathbf{D}_t\mathbf{R}_t\mathbf{D}_t| + \mathbf{a}_t'\mathbf{D}_t^{-1}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1}\mathbf{a}_t \Big) \tag{15}$$

where $\mathbf{D}_t$ and $\mathbf{a}_t$ are from Equation (3) and $\mathbf{R}_t$ is from Equation (8).

Let the parameters of univariate volatility (diagonal elements in $\mathbf{D}_t$) be denoted by $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and the correlation parameters (off-diagonal elements in $\mathbf{R}_t$) be denoted by $\boldsymbol{\phi} \in \boldsymbol{\Phi}$. Define the part of the likelihood which only includes $\boldsymbol{\theta}$ as the volatility part $\mathcal{L}_V(\boldsymbol{\theta})$, and the remainder of the likelihood as the correlation part $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$

$$\mathcal{L}_V(\boldsymbol{\theta}) \quad = \quad -\frac{1}{T}\sum_t \Big(2\log|\mathbf{D}_t| + \mathbf{a}_t'\mathbf{D}_t^{-2}\mathbf{a}_t - \mathbf{e}_t'\mathbf{e}_t \Big) \tag{16}$$

$$\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) \quad = \quad -\frac{1}{T}\sum_t \Big(\log|\mathbf{R}_t| + \mathbf{e}_t'\mathbf{R}_t^{-1}\mathbf{e}_t \Big) \tag{17}$$

The log-likelihood $\mathcal{L}$ can be written as the sum of the two components.

$$\mathcal{L} \quad = \quad \mathcal{L}_V(\boldsymbol{\theta}) + \mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) \tag{18}$$

We proceed with a two-step approach. In the first step of the QML estimation, we find the volatility parameter estimates that maximize $\mathcal{L}_V(\boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}} \quad = \quad \arg\max_{\boldsymbol{\theta}} \{\mathcal{L}_V(\boldsymbol{\theta})\} \tag{19}$$

Then we use the estimates in $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ for the second stage to obtain the maximizer $\hat{\boldsymbol{\phi}}$ (see Appendix A):

$$\hat{\boldsymbol{\phi}} \quad = \quad \arg\max_{\boldsymbol{\phi}} \{\mathcal{L}_C(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})\} \tag{20}$$

For a cross section of size N, the DFC estimates N factor loadings on the common factor ν_t. As the cross-sectional dimension increases, the number of parameters in the DFC also becomes large. To keep our model parsimonious, we consider imposing a group structure on factor loadings to reduce the number of parameters. Factor loadings are restricted to be equal within a group. We refer to this model variation as Group DFC. In the extreme case of one group, all of the factor loadings are equal and our model reduces to the Dynamic Equicorrelation model of R. Engle and Kelly (2012). To estimate Group DFC, we first estimate initial values of factor loadings. We then form groups based on the similarity in the factor loadings (see Appendix A for details).

3.3 A New Volatility Factor

In the DFC model, the conditional variance of the common factor in idiosyncratic volatility shocks, $h_{v,t}$, drives the correlation structure in idiosyncratic volatility. We examine this variable to better understand its property. It is natural to compare $h_{v,t}$ to the second moments of key variables such as asset pricing factors and macroeconomic quantities. Table 2 reports the correlations between the conditional variance $h_{v,t}$ and volatility measures of the Fama and French (1993) and Fama and French. (2015) factors. We construct monthly realized volatilities (RV) of the factors using daily data, as well as realized volatility innovations as residuals from autoregressive (AR) models chosen by the Akaike information criterion (AIC). Each realized volatility series may have its own AIC-selected AR order.

The top panel contains results for the long sample of the Fama and French (1993) factors, from 1926 to 2015. $h_{v,t}$ is positive correlated with the realized volatilities of the market, size, and value factors. The correlations are moderate, ranging from 0.35 to 0.39. $h_{v,t}$ is also positively correlated with the realized volatility innovations. Correlations with RV innovations are generally smaller compared to the RV levels.

The bottom panel contains the correlations between $h_{v,t}$ and the Fama and French. (2015) factors, which are the original three factors plus profitability (RMW) and investment (CMA). RMRF, SMB, and HML are constructed differently from the top panel, and are only available starting in 1963[4]. $h_{v,t}$ is positively correlated with the realized volatility of each of the five factors, with values ranging from 0.34 to 0.43. $h_{v,t}$ is positively but less strongly correlated with the RV innovations, with values from 0.22 to 0.34.

The undiversifiable nature of $h_{v,t}$ may be related to the prevailing macroeconomic conditions. We explore the relationship between $h_{v,t}$ and macroeconomic variables in Table 3. We use the Chicago Fed National Activities Index (CFNAI), as well as its constituent components PI, EUH, CH, and SOI[5] as barometers for macroeconomic conditions. We examine the correlations between $h_{v,t}$ and four transformations of these variables. First, we look at how $h_{v,t}$ comoves with the raw values. Second, we calculate the innovations to these variables from autoregressive models chosen by the AIC. Third, we fit GARCH(1,1) models to CFNAI and its components to calculate their volatilities. Fourth, we take the GARCH residuals to be macroeconomic volatility innovations.

$h_{v,t}$ is negatively correlated with the raw CFNAI and its components, indicating that it may underline some form of macroeconomic risk. The correlation magnitudes for the components are

[4]Please refer to Ken French's website for variable construction. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[5]PI: production and income, EUH: employment, unemployment, and hours, CH: personal consumption and housing, SOI: sales, orders, and inventories.

Table 2: **Correlations between Factor Volatility Measures and $h_{v,t}$**
We report correlations between the conditional variance $h_{v,t}$ and realized volatility and realized volatility innovations of the Fama and French (1993) and Fama and French. (2015) factors. We compute monthly realized volatility using daily data. RV innovations are calculated as the residuals from the best autoregressive model fitted to each volatility series. The AR models are chosen using the Akaike information criterion. The top panel displays values for the Fama and French (1993) three factors for 1926-2015. The bottom panel displays values for the Fama and French. (2015) five factors for 1963-2015.

Realized Vol		RV Innovations	
1926-2015			
RMRF	0.35	RMRF	0.21
SMB	0.39	SMB	0.30
HML	0.39	HML	0.38
1963-2015			
RMRF	0.37	RMRF	0.24
SMB	0.40	SMB	0.32
HML	0.43	HML	0.34
RMW	0.37	RMW	0.32
CMA	0.34	CMA	0.22

Table 3: **Correlations between Macroeconomic Measures and $h_{v,t}$**
We report correlations between the conditional variance $h_{v,t}$ and four macroeconomic measures. We include macro indexes from the Chicago Fed National Activities Index (CFNAI) and its constituent series: production and income (PI); employment, unemployment, and hours (EUH); personal consumption and housing (CH); and sales, orders, and inventories (SOI). Macro innovations are calculated as the residuals from the best autoregressive model fitted to each volatility series selected using the Akaike information criterion (AIC). Macro vol are the fitted values from a GARCH(1,1) model. Macro vol innovations are standardized residuals from the GARCH models.

	Raw Macro Variables	Macro Vol
CFNAI PI	-0.10	0.11
CFNAI EUH	-0.13	0.09
CFNAI CH	-0.02	0.01
CFNAI SOI	-0.12	0.15
CFNAI	-0.13	0.14
	Macro Innovations	Macro Vol Innovations
CFNAI PI	-0.03	-0.10
CFNAI EUH	-0.04	-0.13
CFNAI CH	-0.03	-0.02
CFNAI SOI	-0.08	-0.12
CFNAI	-0.05	-0.13

similar with the exception of CFNAI CH, which is only weakly related to $h_{v,t}$. $h_{v,t}$ and macroeconomic volatility are somewhat positively correlated. Again, the correlation with CFNAI CH is particularly weak. $h_{v,t}$ does not seem to be related to macroeconomic innovations, as shown in the bottom panel, and $h_{v,t}$ negatively comoves with macroeconomic volatility innovations.

$h_{v,t}$ appears to partially capture the variation in Fama and French (1993) Fama and French. (2015) factor volatilities. $h_{v,t}$ is negatively correlated with raw macroeconomic variables and their volatility innovations, and is positively correlated with macroeconomic volatility. However, the correlations are considerably weaker compared to those for return factors. Overall, $h_{v,t}$ appears to be somewhat related to the undiversifiable risk in second moments of the cross-sectional return factors. Its positive correlation to market volatility suggests that $h_{v,t}$ tends to be elevated in higher marginal utility states of the world which investors would want to avoid. However, the magnitude of the correlations indicate that a significant fraction of its variation remains unexplained.

3.4 Simulations

We can compare the DFC and DCC models using simulations. Since correlated idiosyncratic volatility innovations appears to be an robust feature of the data, we simulate returns under the hypothesis that the true data generating process includes correlated idiosyncratic volatility innovations. We first simulate the conditional variance of the common factor in idiosyncratic volatility shocks, $h_{v,t}$, with a GARCH-style process[6]. We choose values for the factor loadings on the common factor in idiosyncratic volatility shocks, ξ_i, and we use Equation (7) to generate $\{\mathbf{Q}_t\}$ and Equation (8) to generate conditional correlation matrices $\{\mathbf{R}_t\}$. We then form standardized residuals $\{\mathbf{e}_t\}$ from $\{\mathbf{Q}_t\}$ using Equations (4) and (5).

Next, we generate residual returns $\{\mathbf{a}_t\}$ and time-varying volatilities $\{\mathbf{D}_t\}$ iteratively using Equation (3) and the following process:

$$\mathbf{D}_t^2 = \mathbf{\Omega} + \mathbf{A} \circ \mathbf{a}_{t-1}\mathbf{a}_{t-1}' + \mathbf{B} \circ \mathbf{D}_{t-1}^2 \tag{21}$$

Where $\circ$ is the element-by-element (Hadamard) product. We choose the parameters as $\mathbf{\Omega} = 0.003\mathbf{I}_N, \mathbf{A} = 0.1\mathbf{I}_N$, and $\mathbf{B} = 0.85\mathbf{I}_N$, typical values for individual stock return volatility. For simplicity, we do not specify the factor model for returns $\mathbf{f}_t'\boldsymbol{\beta}_i$ in Equation (2), and we treat $\{\mathbf{a}_t\}$ as the simulated returns data set.

[6] $h_{v,t} = 0.05 + 0.05 * v_{t-1}^2 + 0.9 * h_{v,t-1}$ where the common factor in idiosyncratic volatility innovations $v_t \sim \mathcal{N}(0, h_{v,t})$. The initial value of $h_{v,t}$ is set to be 1, the unconditional variance of standardized residuals $e_{i,t}$.

For each simulated data set $\{\mathbf{a}_t\}$, we estimate the DFC and DCC models and compare them using two metrics: root mean squared error and mean absolute error. Both metrics are with respect to the true underlying process[7]. Table 4 presents RMSEs and MAEs of fitted correlation matrices under DFC and DCC. We consider $N = 3$ or $N = 10$ return series, and $T = 1000$ or $T = 5000$ time periods, as well as different choices of the factor loading vector $\boldsymbol{\xi}$. For each set of parameter values, we repeat the simulations $M = 1000$ times. In Panel A and Panel D, we set all ξ_i equal to 0.1. Panel C and Panel F set factor loadings all equal to 0.5. Panel B and Panel E set factor loadings ranging from 0.2 to 0.6.

Table 4: **Comparing DFC and DCC Models using Simulations**
We simulate return series assuming the data-generating process contains correlated idiosyncratic volatility innovations. We consider three or 10 assets in the cross section over 1000 or 5000 time periods, and we include different choices of factor loadings ξ on the common factor in idiosyncratic volatility shocks ν_t (see Equation (5)). For each set of parameter values, we repeat the simulation 1000 times. We estimate the DFC and DCC models on the simulated data sets, and we compare the root mean squared error and mean absolute error.

N = 3	T = 1000		T = 5000		N = 10	T = 1000		T = 5000	
	DFC	DCC	DFC	DCC		DFC	DCC	DFC	DCC
Panel A: $\xi = (0.1, 0.1, 0.1)'$					Panel D: $\xi = 0.1 \cdot \mathbf{1}_{10}$				
RMSE	0.030	0.032	0.011	0.014	RMSE	0.019	0.029	0.009	0.016
MAE	0.023	0.026	0.009	0.010	MAE	0.014	0.023	0.007	0.013
Panel B: $\xi = (0.2, 0.3, 0.5)'$					Panel E: $\xi = (0.2, 0.3, 0.4, 0.5, 0.6)' \otimes \mathbf{1}_2'$				
RMSE	0.029	0.035	0.017	0.022	RMSE	0.022	0.028	0.016	0.020
MAE	0.023	0.027	0.013	0.017	MAE	0.017	0.022	0.012	0.015
Panel C: $\xi = (0.5, 0.5, 0.5)'$					Panel F: $\xi = 0.5 \cdot \mathbf{1}_{10}$				
RMSE	0.039	0.053	0.029	0.045	RMSE	0.037	0.042	0.032	0.037
MAE	0.032	0.042	0.025	0.037	MAE	0.031	0.037	0.028	0.031

The DFC model achieves better performance than DCC in both RMSE and MAE for different

[7]For M simulations of N assets over T time periods, the RMSE and MAE are calculated as follows:

$$RMSE = \sqrt{\frac{2}{MTN(N-1)} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\mathbf{R}}_t^{(m)}(i,j) - \mathbf{R}_t^{(m)}(i,j) \right)^2}$$

$$MAE = \frac{2}{MTN(N-1)} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left| \hat{\mathbf{R}}_t^{(m)}(i,j) - \mathbf{R}_t^{(m)}(i,j) \right|$$

where $\hat{\mathbf{R}}_t^{(m)}(i,j)$ are the estimated returns from the models and $\mathbf{R}_t^{(m)}(i,j)$ are the simulated returns.

portfolio sizes, time periods, and factor loadings we consider. Several observations emerge from Table 4. First, RMSE and MAE tend to be smaller under larger samples for both DFC and DCC models, for greater sample size allows more precise estimation. For example, Panel B shows that as sample size increases from 1000 to 5000 under $\xi = (0.2, 0.3, 0.5)'$, RMSE of DFC and DCC decrease from 0.029 and 0.035 to 0.017 and 0.022.

Second, the improvement in model fit for the DFC model relative to the DCC model tends to be larger when the comovement among idiosyncratic volatility innovations is stronger. This is because the DFC model captures the comovement whereas the DCC model does not. In Panel A, in which the comovement among idiosyncratic volatility innovations is relatively low, RMSE and MAE are similar for the DFC and DCC. In contrast, in Panel C where the factor loadings are higher and comovement is larger, the differences in RMSE and MAE between the two models become more pronounced.

Third, while a greater number of assets is associated with lower RMSE and MAE for both models, the improvement in DFC is greater than that of the DCC model. For example, Panel A and Panel D have the same parameter values with the exception that Panel A contains three assets whereas Panel D contains 10 assets. For the DFC, the RMSE and MAE decrease by 27% and 39%. For the DCC, the decreases are 9% and 12%. Overall, Table 4 show that in the presence of correlated idiosyncratic volatility shocks, the DFC model can provide a better fit compared to the DCC model.

3.5 Empirical Tests

We test the empirical performance of DFC and DCC models on characteristic-sorted portfolios. We also include the DECO model of R. Engle and Kelly (2012), which can be viewed as as restricted case of the DFC model. The characteristic-sorted portfolios are based on market equity, the book-to-market ratio, long-term reversal, operating profitability, investments, momentum, and short-term reversal. We first fit the Fama and French (1993) model to portfolio returns, then we fit DFC, DCC, or DECO model to the residual returns.

Table 5 reports the DFC factor loadings for each set of portfolios. Characteristics-sorted decile portfolios have economically large and statistically significant factor loadings on the common factor in idiosyncratic volatility innovations. We report the QMLE asymptotic t-statistics in the parentheses. The factor exposures of all 10 portfolios are significantly different from zero for ME, LTRev, Inv, and Mom, and nine of the 10 portfolios have significant factor loadings for each of BE/ME, OP, and STRev. From earlier sections, we saw that idiosyncratic volatility innovations are correlated across characteristic-sorted portfolios. Significant factor loadings is another way to quantify this result, using a formal statistical model in the DFC.

Table 5: **DFC Estimation for Characteristic-Sorted Portfolios**
We regress characteristic-sorted portfolio returns on the Fama and French (1993) factors, and we estimate the DFC, DCC and DECO models on the residual returns. We report the 10 DFC factor loadings for each set of portfolios, along with their QMLE asymptotic t-statistics in parentheses. We also report the likelihood values of the three models. Each column presents a different set of decile portfolios.

	ME	BE/ME	LTRev	OP	Inv	Mom	STRev
ξ_1	0.229	0.448	0.348	0.358	0.356	0.672	0.485
	(-12.06)	(18.67)	(21.75)	(11.19)	(11.48)	(67.2)	(30.31)
ξ_2	0.139	0.089	0.624	0.544	0.486	0.764	0.547
	(-6.32)	(2.70)	(44.57)	(20.15)	(15.68)	(95.50)	(39.07)
ξ_3	-0.236	-0.252	0.648	0.414	0.475	0.747	0.489
	(11.24)	(-8.69)	(49.85)	(14.79)	(16.38)	(83.00)	(32.60)
ξ_4	-0.395	-0.444	0.522	0.254	0.437	0.632	0.373
	(21.94)	(-19.30)	(32.63)	(7.70)	(13.66)	(52.67)	(20.72)
ξ_5	-0.471	-0.527	0.401	0.302	0.230	0.425	0.145
	(31.40)	(-25.10)	(20.05)	(8.88)	(6.39)	(28.33)	(7.63)
ξ_6	-0.630	-0.514	0.245	0.191	0.099	0.098	0.003
	(48.46)	(-23.36)	(10.65)	(5.16)	(2.41)	(4.45)	(0.75)
ξ_7	-0.682	-0.299	-0.076	-0.027	-0.199	-0.285	-0.331
	(56.83)	(-11.96)	(-3.04)	(-0.73)	(-5.24)	(-15.83)	(-17.42)
ξ_8	-0.687	-0.166	-0.328	-0.222	-0.301	-0.533	-0.385
	(62.45)	(-5.93)	(-15.62)	(-6.53)	(-8.60)	(-41.00)	(-24.06)
ξ_9	-0.629	0.053	-0.424	-0.427	-0.487	-0.633	-0.335
	(48.38)	(1.83)	(-23.56)	(-15.25)	(-16.23)	(-57.55)	(-20.94)
ξ_{10}	0.462	0.327	-0.554	-0.411	-0.202	-0.710	-0.458
	(-28.88)	(12.58)	(-32.59)	(-14.17)	(-5.32)	(-71.00)	(-28.63)
$\mathcal{L}$(DFC)	31813.6	27972.5	26164.0	17596.8	17535.7	26146.3	26032.7
$\mathcal{L}$(DCC)	31535.3	27602.1	25918.7	17430.4	17474.6	25746.7	25771.5
$\mathcal{L}$(DECO)	31080.7	27432.0	25524.4	17465.2	17405.3	24164.4	25463.5

The bottom three rows of Table 5 compare the likelihood values of the DFC, DCC, and DECO models. The DFC model consistently provides a superior fit compared to both the DECO and DCC models. Although the DCC model allows for time-varying correlations in standardized residuals, it does not account for the comovement in idiosyncratic volatility innovations that we observe in the data. In comparison, the DFC explicitly models this comovement, which leads to an improved fit compared to the DCC model. The DCC model compares favorably against DECO and outperforms in six of the seven sets of portfolios (operating profitability is the exception): DFC has greater flexibility than the DECO model by allowing for factor loadings to differ across portfolios, which can lead to better model fit.

We evaluate the performance of Group DFC against DFC for characteristic-sorted portfolios. Based on the factor loading estimates in Table 5, we group portfolios with similar loadings into the same group and estimate the Group DFC model (see Appendix A for details on Group DFC). Table 6 provides the groups for each set of portfolios and reports the Bayesian Information Criterion obtained from the Group DFC model and the DFC without group structure. The advantage of the Group DFC lies in its ability to reduce the number of unknown parameters compared to the DFC model. As a more parsimonious model, the Group DFC achieves lower BICs for six of the seven characteristic-sorted deciles (long-term reversal is the exception). One disadvantage of the Group DFC is that we lose some economic interpretation. While we achieve a better statistical fit with Group DFC, it is not clear what each group represents, or why certain portfolios should belong in the same group.

Table 6: **Group DFC Estimation for Characteristic-Sorted Portfolios**
We regress characteristic-sorted portfolio returns on the Fama and French (1993) factors, and we estimate Group DFC models on the residual returns. Groups are based on the factor loading estimates in Table 5. Deciles are numbered 1 through 10, and each cell contains the deciles in that group. For each set of portfolios, we report the Bayesian Information Criterion of the DFC and Group DFC.

	ME	BE/ME	LTRev	OP	Inv	Mom	STRev
Group 1	1,2	1	1	1,2,3	1,2,3,4	1	1,3
Group 2	3	2	2,3,4	4,5,6	5	2,3	2
Group 3	4	3	5	7	6	4	4
Group 4	5	4,5,6	6	8	7	5	5
Group 5	6	7,8	7	9,10	8,9,10	6	6
Group 6	7,8	9	8			7	7,8,9
Group 7	9	10	9,10			8	10
Group 8	10					9,10	
BIC(Group DFC)	-25766.5	-27711.3	-25887.5	-17364.4	-17297.8	-25876.1	-25775.3
BIC(DFC)	-25756.3	-27696.1	-25890.1	-17339.7	-17278.7	-25869.9	-25756.3

4 Application of DFC to Portfolio Construction

The DFC model could be used to improve portfolio optimization. We construct mean-variance portfolios using different volatility models to estimate the covariance matrix. We include the historical covariance estimate, the DECO model, and variations of the DFC model. The volatility model that produces the most attractive risk-return trade-off would be the most useful model for portfolio optimization. We focus on this metric when comparing volatility models.

For all competing models, we evaluate the out-of-sample portfolio performance of mean-variance (MV) portfolios of Markowitz (1952). Suppose at time t, we have N securities with expected return vector μ_t and covariance matrix $\mathbf{\Sigma}_t$. We follow Markowitz (1952) to form portfolios that minimize portfolio variance given an expected-returns target μ_0:

$$\min_{\mathbf{w}_t} \mathbf{w}_t'\mathbf{\Sigma}_t\mathbf{w}_t \quad s.t. \quad \mathbf{w}_t'\mathbf{1}_N = 1 \tag{22}$$

$$\mathbf{w}_t'\mu_t \geq \mu_0 \tag{23}$$

Define $A_t = \mathbf{1}_N'\mathbf{\Sigma}_t^{-1}\mathbf{1}_N$, $B_t = \mathbf{1}_N'\mathbf{\Sigma}_t^{-1}\mu_t$, and $C_t = \mu_t'\mathbf{\Sigma}_t^{-1}\mu_t$. The solution to Problem (22) given (23) is the minimum-variance portfolio for a target level of expected returns μ_0:

$$\mathbf{w}_t^{MV}(\mu_0) = \frac{C_t - \mu_0 B_t}{A_t C_t - B_t^2}\mathbf{\Sigma}_t^{-1}\mathbf{1}_N + \frac{\mu_0 A_t - B_t}{A_t C_t - B_t^2}\mathbf{\Sigma}_t^{-1}\mu_t \tag{24}$$

For portfolio target returns, we consider $\mu_0 = 5\%$ and $\mu_0 = 10\%$.

We compare different models for the covariance matrix Σ_t:

- **Model 1** Historical covariance matrix of raw industry portfolio returns.

- **Model 2** DFC: Each month, fit DFC on raw industry return series and forecast $\mathbf{D}_t$ and $\mathbf{R}_t$ for the next month. Form Σ_t as $\Sigma_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t$.

- **Model 3** CAPM DFC: We first regress raw returns on market excess returns r_{mt} to obtain residual returns. We fit a univariate GARCH(1,1) model on r_{mt} and estimate a DFC model on the residual returns. We then construct the covariance matrix as $h_{mt}\boldsymbol{\beta}_t\boldsymbol{\beta}_t' + \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t'$ where h_{mt} comes from the GARCH(1,1) model on the market, $\mathbf{D}_t$ and $\mathbf{R}_t$ are from the DFC forecasts, and $\boldsymbol{\beta}_t$ is a vector of market betas.

- **Model 4** 30-Group CAPM DFC: We regress raw returns on market excess returns r_{mt} to obtain residual returns. We then estimate a univariate GARCH(1,1) model on r_{mt} and estimate a Group DFC model for the residual returns based on the 30-Group industry classification. The covariance matrix is constructed similarly as Model 3 with h_{mt} coming from the GARCH(1,1) model on market returns, $\mathbf{D}_t$ and $\mathbf{R}_t$ from the 30-Group DFC, and a vector of market betas $\boldsymbol{\beta}_t$.

- **Model 5** 10-Group CAPM DFC: Similar set up as Model 4 with 10-Group industry classification in place of the 30-Group.

- **Model 6** 5-Group CAPM DFC: Similar set up as Model 4 with 5-Group industry classification in place of the 30-Group.

- **Model 7** CAPM DECO: Just like Model 3, we start by regression raw returns on market excess returns r_{mt} to obtain residual returns. We estimate a GARCH(1,1) model on r_{mt} and a DECO model for residual returns. The covariance matrix estimate is constructed similarly as Model 3 with DECO in place of DFC.

- **Model 8** FF3 DFC: We regress raw returns on Fama and French (1993) factors to obtain residual returns. We fit univariate GARCH(1,1) models on r_{mt}, SMB, and HML, and estimate a DFC model for the residual returns. The covariance matrix is constructed as $\mathbf{B}_t \mathrm{diag}\{h_{mt}, h_{SMB,t}, h_{HML,t}\}\mathbf{B}'_t + \mathbf{D}_t\mathbf{R}_t\mathbf{D}'_t$, where h_{mt}, $h_{SMB,t}$, and $h_{HML,t}$ are the GARCH variance estimates for the Fama and French (1993) factors, $\mathbf{D}_t$ and $\mathbf{R}_t$ are constructed from the DFC model, and $\mathbf{B}_t$ is the factor loading matrix.

- **Model 9** 30-Group FF3 DFC: We get residual returns from the Fama and French (1993) model and fit a Group DFC model based on the 30-Group industry classification. We also estimate univariate GARCH(1,1) models for the Fama and French (1993) factors. The covariance matrix is constructed as Model 8 with $\mathbf{D}_t$ and $\mathbf{R}_t$ coming from the 30-Group DFC model.

- **Model 10** 10-Group FF3 DFC: Similar set up as Model 9 with 10-Group industry classification instead of the 30-Group.

- **Model 11** 5-Group FF3 DFC: Similar set up as Model 9 with 5-Group industry classification instead of the 30-Group.

- **Model 12** FF3 DECO: Like Model 8, we regress raw returns on Fama and French (1993) factors to obtain residual returns. We then estimate GARCH(1,1) models on r_{mt}, SMB, and HML, and a DECO model for residual returns. The covariance matrix is formed in the same steps as Model 8 with DECO in place of DFC.

The volatility models can be categorized into three classes. Model 1 and Model 2 fit volatility models to raw returns without specifying a factor model for returns. These serve as the baseline models for comparison. Models 3 through 7 first fit a CAPM to returns, then five different volatility models to residuals returns. Models 8 through 12 first fit the Fama and French (1993) model to returns, then different volatility models for residual returns. Within the second and third classes, we could directly compare the performance of DFC, Group DFC, and DECO, since all the models are estimated on the same set of residual returns.

We form mean-variance portfolios using 49 industry portfolios[1]. For Group DFC models, we follow French's industry classification based on Standard Industry Classification (SIC) code. Group definitions are provided in Table 7. At the beginning of each month, we use 60-month rolling windows to estimate the models, using the 60-month average returns in place of μ_t. The portfolios are rebalanced each month.

Table 7: **Classifications of Industry Portfolios**
This table provides the 49 industry portfolios that we use for portfolio construction. Industry definitions are based on SIC codes from Ken French's website[1]. The 49 industries are further classified into 5, 10, and 30 groups based on coarser industry classifications, also from French's website.

Industry	30-Group	10-Group	5-Group	Industry	30-Group	10-Group	5-Group
Agric	Food	NoDur	Cnsmr	Guns	Other	Manuf	Manuf
Food	Food	NoDur	Cnsmr	Gold	Mines	Other	Other
Soda	Food	NoDur	Cnsmr	Mines	Mines	Other	Other
Beer	Food	NoDur	Cnsmr	Coal	Coal	Enrgy	Manuf
Smoke	Smoke	NoDur	Cnsmr	Oil	Oil	Enrgy	Manuf
Toys	Games	NoDur	Cnsmr	Util	Util	Utils	Manuf
Fun	Games	Other	Other	Telcm	Telcm	Telcm	HiTec
Books	Books	NoDur	Cnsmr	PerSv	PerSv	Shops	Cnsmr
Hshld	Hshld	Durbl	Cnsmr	BusSv	PerSv	Other	Other
Clths	Clths	NoDur	Cnsmr	Hardw	BusEq	HiTec	HiTec
Hlth	Hlth	Hlth	Hlth	Softw	BusEq	HiTec	HiTec
MedEq	Hlth	Hlth	Hlth	Chips	BusEq	HiTec	HiTec
Drugs	Hlth	Hlth	Hlth	LabEq	BusEq	HiTec	HiTec
Chems	Chems	Manuf	Manuf	Paper	Paper	Manuf	Manuf
Rubbr	Other	Manuf	Manuf	Boxes	Paper	Other	Other
Txtls	Txtls	NoDur	Cnsmr	Trans	Trans	Other	Other
BldMt	Cnstr	Other	Other	Whlsl	Whlsl	Shops	Cnsmr
Cnstr	Cnstr	Other	Other	Rtail	Rtail	Shops	Cnsmr
Steel	Steel	Manuf	Manuf	Meals	Meals	Shops	Cnsmr
FabPr	FabPr	Manuf	Manuf	Banks	Fin	Other	Other
Mach	FabPr	Manuf	Manuf	Insur	Fin	Other	Other
ElcEq	ElcEq	Manuf	Manuf	RlEst	Fin	Other	Other
Autos	Autos	Durbl	Manuf	Fin	Fin	Other	Other
Aero	Carry	Manuf	Manuf	Other	Other	Other	Other
Ships	Carry	Manuf	Manuf				

Table 8 presents the annualized portfolio volatility for all the models. Panel A reports the results for the full sample period, whereas Panel B includes results from 1995 to 2015. In each panel, we report the volatility of MV portfolios with target expected returns of 5% and 10%. For models with the same factor model for returns (CAPM or the Fama and French (1993) three-factor model), the one achieving the lowest volatility is represented in bold.

In the first class of volatility models, we compare the historical covariance matrix with DFC. The DFC model achieves the significantly lower volatility for each MV portfolio. For the full sample with a target return of 5%, the DFC model has a volatility of 13.5%, a 32% reduction compared to the 19.9% volatility from the historical covariance matrix. The reduction in portfolio volatility is similar for portfolios with 10% target return: 13.3% compared to 19.7%. The difference in volatility is more pronounced, up to 40%, in the recent sub-sample. For a target return of 5%, the portfolio volatility is 13.6% using the DFC model and 22.2% using the historical covariance matrix. Targeting a portfolio return of 10% leads to volatility of 13.0% for the DFC and 21.7% for the historical estimate.

The DFC model performs better when combined with a factor model for returns. Estimating first the CAPM or the Fama and French (1993) model to returns improves the stability of the mean-variance efficient portfolios, thereby reducing their out-of-sample standard deviations. Portfolio volatility is lower for CAPM DFC compared to DFC, and the portfolio volatility is lower still for FF3 DFC compared to CAPM DFC. In fact, all models that include the CAPM or the Fama and French (1993) model result in lower out-of-sample portfolio volatility compared to DFC. The reduction in volatility ranges between 4% to 15%.

Among models with the same factor model for returns (the CAPM or the Fama and French (1993) model), we compare the performance of DFC models against the DECO model using the Diebold and Mariano (1995) test. For example, we compare the portfolio volatility of CAPM DECO with those of CAPM DFC and the CAPM Group DFCs. If two volatility models have large differences in performance, we reject the null that the models have equal predictive accuracy for the covariance matrix. The DECO model assumes factor loadings on the common factor in idiosyncratic volatility shocks are all equal, so our test is essentially testing whether allowing for less restrictive factor loadings results in more attractive portfolios.

We focus our attention to the case of CAPM as the factor model, and compare DFC, Group DFC, and DECO. For the full sample, DFC and Group DFCs achieve lower portfolio volatilities compared to that of DECO. Whereas the CAPM DECO model has an out-of-sample volatility of 12.3% for a target return of 5%, DFC has a volatility of 12.0%, and Group DFCs have the lowest volatilities ranging from 11.7% to 11.9%. The CAPM 10-Group DFC has the lowest portfolio volatility, 11.7%, which is statistically lower compared to the CAPM DECO model at the 1% level. Portfolios with expected return targets of 10% display the same pattern. CAPM DFC and CAPM Group DFC models have lower volatility compared to CAPM DECO, and the CAPM 10-Group

Table 8: **Out-of-Sample Portfolio Volatility**
We report annualized portfolio volatility for minimum-variance portfolios given target expected returns of 5% or 10% (see Equations (24)). We compare twelve volatility models used to construct the covariance matrix (see Section 4). The top panel is made up of two models fit to raw returns. The middle panel consists of all models that include the CAPM as the factor model for returns. The lower panel contains all models that include the Fama and French (1993) model. Among specifications with the same factor model for returns (the CAPM or the Fama and French (1993) model), the one with lowest volatility is shown in bold. Each bolded model is marked with *, **, or ***, if it achieves significant lower volatility than the DECO model with the same factor model specification, based on Diebold-Mariano tests at the 5.0%, 1.0%, or 0.1% (one-sided) levels. Panel A contains results for 1946 to 2015 and Panel B contains results for 1995 to 2015.

Model	Target Returns			
μ_0	5%	10%	5%	10%
	Panel A: 1946 - 2015		Panel B: 1995 - 2015	
Historical covariance	19.9%	19.7%	22.2%	21.7%
DFC	13.5%	13.3%	13.6%	13.0%
CAPM DFC	12.0%	11.9%	12.6%	12.3%
CAPM 30-Group DFC	11.9%	11.8%	12.4%	12.1%
CAPM 10-Group DFC	**11.7%****	**11.7%****	12.1%	11.7%
CAPM 5-Group DFC	11.9%	11.9%	**11.9%****	**11.6%****
CAPM DECO	12.3%	12.2%	12.9%	12.5%
FF3 DFC	11.6%	11.7%	**11.6%***	11.4%
FF3 30-Group DFC	**11.6%***	11.6%	11.6%	**11.4%***
FF3 10-Group DFC	11.6%	**11.6%****	11.8%	11.5%
FF3 5-Group DFC	11.7%	11.7%	11.7%	11.5%
FF3 DECO	11.7%	11.7%	11.7%	11.5%

DFC performs the best.

Turning our attention to models with the Fama and French (1993) model, the FF3 DFC and FF3 Group DFC models have somewhat lower volatilities compared to the FF3 DECO model. The portfolio volatility for FF3 DECO, when the target expected return is 5%, is 11.7%, similar to those observed for FF3 DFC and FF3 Group DFC models. FF3 30-Group DFC has the lowest volatility, which is statistically different from the FF3 DECO model at the 5% level. For portfolios with target expected returns of 10%, FF3 10-Group DFC has the lowest volatility, also lower than the FF3 DECO model at the 5% level.

We find the similar results across in the sub-sample in Panel B. DFC outperforms the historical covariance estimate. Among the CAPM models, CAPM DFC and CAPM Group DFCs have lower volatility compared to the CAPM DECO model. CAPM 5-Group DFC has significantly lower volatility compared to the CAPM DECO. The FF3 models show similar levels of volatility with smaller improvements of FF3 Group DFC models over the FF3 DECO. Nevertheless, FF3 DFC and FF3 30-Group DFC have significantly lower portfolio volatilities compared to the FF3 DECO model.

Comparing across all models, a combination of the Fama and French (1993) and the Group DFC models give the least volatile out-of-sample portfolios. Since we use the same historical average return inputs for each model and the portfolio target returns are identical in each column, lower portfolio volatility implies an improved risk-return trade-off. The model with the lowest volatility is also the one with the highest Sharpe ratio. By estimating the covariance matrix more precisely, the DFC models provide the investor with a superior investment opportunity set.

This section illustrates the practical value of the DFC model in portfolio construction. We provide evidence that the DFC model improves portfolio performance of the mean-variance portfolios based on the 49 industry portfolios. Compared with historical covariance matrix or DECO, DFC and Group-DFC achieve lower out-of-sample volatility of the optimized portfolios. Combining factor models such as the CAPM and the Fama and French (1993) model with DFC further reduces the portfolio volatility and improves out-of-sample Sharpe ratios.

5 Conclusion

This paper provides evidence that correlated idiosyncratic volatility shocks are an important force that contributes to the comovement in idiosyncratic volatility. We propose a new multivariate GARCH model, Dynamic Factor Correlation, to capture this empirical fact. Our model outperforms the DCC model of R. Engle (2002) and the DECO model of R. Engle and Kelly (2012) in simulations and empirical tests. Aside from DCC, the DFC model is also closely related to other multivariate GARCH models. Under certain parametric restrictions, the DFC reduces to the CCC model of Bollerslev (1990) or the DECO model of R. Engle and Kelly (2012).

We apply the DFC model to portfolio construction. We construct mean-variance portfolios using 49 industry portfolios as basis assets. Using DFC and Group DFC models to estimate the covariance matrix, we can achieve lower out-of-sample portfolio volatility compared to using the historical covariance matrix or DECO. Portfolio volatility estimates are significantly lower than those using DECO at the 5% and 1% levels. Lower portfolio volatility is associated with a more attractive risk-return trade-off for the investor. Therefore, DFC and Group DFC models can help expand the investment opportunity set in mean-variance optimization.

Like other multivariate GARCH models, the DFC requires a balanced panel in its estimation. An interesting future direction is to extend the estimation procedure to be able to handle unbalanced panels. For example, a generalization of the Group DFC model allowing for time-varying groups similar to the approach in Wang and Tsay (2019) will permit the estimation to handle unbalanced panels. Another potentially fruitful direction is to apply the DFC in other settings such as international equities or commodities.

Our application of the DFC model to portfolio construction demonstrates the usefulness of better covariance estimates. There are different ways to improve estimates of the covariance matrix. In this paper, we capture more realistic behavior of portfolio volatility by modeling correlated idiosyncratic volatility shocks. The portfolio manager oblivious to this empirical result may think her portfolio is more diversified than it really is - because there is a common component in idiosyncratic volatility, it cannot be entirely diversified away; when idiosyncratic volatility of one portfolio constituent increases, idiosyncratic volatilities of other portfolio constituents also tend to increase. The DFC model allows the portfolio manager to better monitor correlated idiosyncratic volatility, thereby improving risk management.

Acknowledgments

We thank Vitali Alexeev (Auckland Finance Meeting discussant), Redouane Elkamhi, Jing Wang (29th Australasian Finance and Banking Conference discussant), Mark Huang, and Claudio Michelacci for helpful discussions. We also thank the conference participants at the 2016 Econometric Society European Winter Meeting, 29th Australasian Finance and Banking Conference, and the Auckland Finance Meeting for their comments.

Declarations of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, *63*(3), 443-494. doi: http://dx.doi.org/10.1016/S0304-405X(02)00068-5

Bauwens, L., Laurent, S., & Rombouts, V. K. (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics*, *21*, 79-109. doi: http://dx.doi.org/10.1002/jae.842

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327. doi: http://dx.doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, *72*, 498-505.

Cappiello, L., Engle, R., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, *4*(4), 537-572. doi: http://dx.doi.org/10.1093/jjfinec/nbl005

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 253–263. doi: http://dx.doi.org/10.1080/07350015.1995.10524599

Ding, Z., & Engle, R. (2001). Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers*, *29*, 157-184.

Duarte, J., Kamara, A., Siegel, S., & Sun, C. (2014). The systematic risk of idiosyncratic volatility. *Available at SSRN 1905731*.

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. doi: http://dx.doi.org/10.1198/073500102288618487

Engle, R., & Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, *30*(2), 212-228. doi: http://dx.doi.org/10.1080/07350015.2011.652048

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 987–1007.

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

Fama, E. F., & French., K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1-22. doi: http://dx.doi.org/10.1016/j.jfineco.2014.10.010

Herskovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial*

Economics, *119*(2), 249–283. doi: http://dx.doi.org/10.1016/j.jfineco.2015.09.010

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance*, *56*(2), 649-676. doi: http://dx.doi.org/10.1111/0022-1082.00340

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, *7*(1), 77-91. doi: http://dx.doi.org/10.2307/2975974

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, *42*(3), 483–510.

Sherman, J., & Morrison, W. J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *Annals of Mathematical Statistics*, *21*(1), 124-127. doi: http://dx.doi.org/10.1214/aoms/1177729893

Tsay, R. S. (2006). Multivariate volatility models. *Time Series and Related Topics, IMS Lecture Notes-Monograph Series*, *52*, 210-222. doi: http://dx.doi.org/10.1214/074921706000001058

Wang, Y., & Tsay, R. S. (2019). Clustering multiple time series with structural breaks. *Journal of Time Series Analysis*, *40*, 182–202. doi: http://dx.doi.org/10.1111/jtsa.12434

White, H. (1994). Estimation, inference and specification analysis. *New York, NY: Cambridge University Press*. doi: http://dx.doi.org/10.1017/CCOL0521252806

A Sherman-Morrison Formula and the DFC Likelihood

We use the Sherman-Morrison formula (Sherman & Morrison, 1950) to write out a closed-form solution to the likelihood function of the Dynamic Factor Correlation model. By Sherman and Morrison (1950), we have the following lemma:

Lemma A.1. *Given $h_{v,t}$, suppose that $\forall i$, $\xi_i^2 \neq 1$, and $1 + \sqrt{h_{v,t}}^2 \boldsymbol{\xi}' \boldsymbol{\Lambda}_t^{-1} \boldsymbol{\xi} \neq 0$, then the inverse of $\mathbf{Q}_t$ is equal to*

$$\mathbf{Q}_t^{-1} = \boldsymbol{\Lambda}^{-1} - \frac{\boldsymbol{\Lambda}^{-1} \boldsymbol{\xi} \boldsymbol{\xi}' \boldsymbol{\Lambda}^{-1}}{1 + h_{v,t} \boldsymbol{\xi}' \boldsymbol{\Lambda}^{-1} \boldsymbol{\xi}} \tag{25}$$

and the determinant of $\mathbf{Q}_t$ is

$$\begin{aligned} det(\mathbf{Q}_t) &= det(\boldsymbol{\Lambda})[1 + h_{v,t} \boldsymbol{\xi}' \boldsymbol{\Lambda}^{-1} \boldsymbol{\xi}] \\ &= [1 + h_{v,t} \boldsymbol{\xi}' \boldsymbol{\Lambda}^{-1} \boldsymbol{\xi}] \prod_{i=1}^{N} (1 - \xi_i^2) \end{aligned} \tag{26}$$

A closed form expression of the determinant of $\mathbf{Q}_t$ will be useful to construct the likelihood of the DFC.

Assumption A.1. $\forall i$, $\xi_i^2 < 1$.

Proposition A.2. *Under Assumption A.1, $\mathbf{R}_t$ is positive definite.*

Proof of Proposition A.2 is straightforward. Under the condition that $\forall i$, $\xi_i^2 < 1$, $\boldsymbol{\Lambda}$ is positive definite. Since $h_{v,t} \cdot \boldsymbol{\xi}\boldsymbol{\xi}'$ is positive semidefinite, the sum of $\boldsymbol{\Lambda}$ and $h_{v,t} \cdot \boldsymbol{\xi}\boldsymbol{\xi}'$ is positive definite.

The log likelihood $\mathcal{L}$ for the estimator of DFC model is

$$\mathcal{L} = -\frac{1}{T} \sum_t \left(\log |\mathbf{D}_t \mathbf{R}_t \mathbf{D}_t| + \mathbf{a}_t' \mathbf{D}_t^{-1} \mathbf{R}_t^{-1} \mathbf{D}_t^{-1} \mathbf{a}_t \right) \tag{27}$$

which can be written as the sum of two parts:

$$\mathcal{L} = \mathcal{L}_V(\boldsymbol{\theta}) + \mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) \tag{28}$$

$\mathcal{L}_V(\boldsymbol{\theta})$ is a function of only univariate volatility parameters $\boldsymbol{\theta}$; it is proportional to the sum of individual log volatilities:

$$\mathcal{L}_V(\boldsymbol{\theta}) = -\frac{2}{T} \sum_t \sum_{i=1}^{N} \left(\log(\sqrt{h_{i,t}}) \right) \tag{29}$$

$\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ is a function of both univariate volatility parameters $\boldsymbol{\theta}$ and correlation parameters $\boldsymbol{\phi}$. By Lemma A.1, we can rewrite $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ as

$$\begin{aligned} \mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\tfrac{1}{T}\textstyle\sum_t \Big(\log\Big[1 + h_{v,t}\sum_{i=1}^N \tfrac{\xi_i^2}{1-\xi_i^2}\Big] + \sum_{i=1}^N \log\Big[1-\xi_i^2\Big] - \sum_{i=1}^N \log\Big[1+(h_{v,t}-1)\xi_i^2\Big]\Big) \\ -\tfrac{1}{T}\textstyle\sum_t \left(\sum_{i=1}^N \Big[\tfrac{1+(h_{v,t}-1)\xi_i^2}{1-\xi_i^2}\Big] e_{i,t}^2 - \tfrac{1}{1+h_{v,t}\sum_{i=1}^N \frac{\xi_i^2}{1-\xi_i^2}}\Big[\sum_{i=1}^N \tfrac{\sqrt{1+(h_{v,t}-1)\xi_i^2}}{1-\xi_i^2} e_{i,t}\xi_i\Big]^2\right) \end{aligned} \tag{30}$$

We employ a two-step estimation strategy. We first find the volatility parameter estimates that maximize $\mathcal{L}_V(\boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}}\{\mathcal{L}_V(\boldsymbol{\theta})\} \tag{31}$$

Then we use $\hat{\boldsymbol{\theta}}$ for the second stage to obtain the maximizer $\hat{\boldsymbol{\phi}}$:

$$\hat{\boldsymbol{\phi}} = \arg\max_{\boldsymbol{\phi}}\{\mathcal{L}_C(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})\} \tag{32}$$

In particular, we estimate $\hat{\boldsymbol{\phi}}$ by replacing the estimated standardized innovation $\hat{e}_{i,t}$ from the first stage estimation with $e_{i,t}$ in Equation (30). Under regularity conditions (see Appendix B), consistency of $\hat{\boldsymbol{\theta}}$ would ensure consistency of $\hat{\boldsymbol{\phi}}$ based on Theorem 6.11 in White (1994). Moreover, we can evaluate Equation (30) using $h_{v,t}$, the standardized innovations $\hat{e}_{i,t}$, and the factor loadings $\{\xi_i\}$. From Lemma A.1, we do not require numerical matrix inversions or determinants calculations for Equation (30) which reduces the computational complexity of the likelihood optimization problem, making the DFC model easily applicable to high-dimensional settings.

As the cross-sectional dimension increases, we could keep our model parsimonious by grouping factor loadings ξ_i. Suppose there are K groups among N factor loadings. Define g_i as the group that loading i belongs to and $\mathcal{G}_k = \{i : g_i = k\}$. Within each group k, all factor loadings are equal to $\tilde{\xi}_k$. Under this group structure, $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ can be rewritten as

$$\begin{aligned} \mathcal{L}_C^G(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\tfrac{1}{T}\textstyle\sum_t \Big(\log\Big[1 + h_{v,t}\sum_{k=1}^K \tfrac{|\mathcal{G}_k|\tilde{\xi}_k^2}{1-\tilde{\xi}_i^2}\Big] + \sum_{k=1}^K |\mathcal{G}_k|\Big\{\log\Big[1-\tilde{\xi}_k^2\Big] - \log\Big[1+(h_{v,t}-1)\tilde{\xi}_k^2\Big]\Big\}\Big) \\ -\tfrac{1}{T}\textstyle\sum_t \left(\sum_{i=1}^N \Big[\tfrac{1+(h_{v,t}-1)\tilde{\xi}_{g_i}^2}{1-\tilde{\xi}_{g_i}^2}\Big] e_{i,t}^2 - \tfrac{1}{1+h_{v,t}\sum_{i=k}^K \frac{|\mathcal{G}_k|\tilde{\xi}_k^2}{1-\tilde{\xi}_k^2}}\Big[\sum_{i=1}^N \tfrac{\sqrt{1+(h_{v,t}-1)\tilde{\xi}_{g_i}^2}}{1-\tilde{\xi}_{g_i}^2} e_{i,t}\tilde{\xi}_{g_i}\Big]^2\right) \end{aligned} \tag{33}$$

We determine the initial values of ξ for the likelihood optimization problem using the following procedure. First, we solve the optimization problem given the sample correlation matrix $\hat{\rho}$

$$\min \frac{1}{2} \sum_{i \neq j} (\log|\xi_i| + \log|\xi_j| - \log|\hat{\rho}_{i,j}|)^2 \tag{34}$$

First order conditions for each i are

$$(N-2)\log|\xi_i| + \sum_{j=1}^{N} \log|\xi_j| - \sum_{j \neq i} \log|\hat{\rho}_{i,j}| \quad = \quad 0 \tag{35}$$

where $\sum_{j=1}^{N} \log|\xi_j|$ is estimated by $\frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \log|\hat{\rho}_{i,j}|$, thus

$$|\hat{\xi}_i| = \exp\left\{ \frac{1}{N-2} \Big(\sum_{j \neq i} \log|\hat{\rho}_{i,j}| - \frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \log|\hat{\rho}_{i,j}| \Big) \right\} \tag{36}$$

The signs of ξ's are determined in the next step. We assume ξ_1 to be positive, then starting from $i = 2$, we solve the following problem iteratively to obtain $\{\hat{\xi}_i\}$:

$$\hat{\xi}_i \quad = \quad \arg \min_{\xi_i \in \{-|\hat{\xi}_i|, |\hat{\xi}_i|\}} \sum_{j<i} (\xi_i \hat{\xi}_j - \hat{\rho}_{i,j})^2 \tag{37}$$

B Assumptions for Consistency and Asymptotic Normality of QML Estimator

B1. The observed data $\{a_t\}$ are a realization of a stochastic process on a complete probability space.
B2. $\mathcal{L}_V$ and $\mathcal{L}_C$ are measurable for each $(\boldsymbol{\theta}, \boldsymbol{\phi})$ in $\{\boldsymbol{\Theta}, \boldsymbol{\Phi}\}$ and continuous on $\{\boldsymbol{\Theta}, \boldsymbol{\Phi}\}$ for all t.
B3(1). For each $(\boldsymbol{\theta}, \boldsymbol{\phi})$ in $\{\boldsymbol{\Theta}, \boldsymbol{\Phi}\}$, $\mathbb{E}[\log \mathcal{L}_V(a_t, \boldsymbol{\theta})]$ and $\mathbb{E}[\log \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})]$ exist and are finite for all t.
(2). $\mathbb{E}[\log \mathcal{L}_V(a_t, \boldsymbol{\theta})]$ and $\mathbb{E}[\log \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})]$ are continuous on $\{\boldsymbol{\Theta}, \boldsymbol{\Phi}\}$ for all t.
(3). $\{\log \mathcal{L}_V(a_t, \boldsymbol{\theta})\}$ and $\{\log \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})\}$ obeys the strong uniform law of large numbers (ULLN).
B4. $\mathcal{L}_V(a_t, \boldsymbol{\theta})$ and $\mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})$ are continuously differentiable of order 2 on $\{\boldsymbol{\Theta}, \boldsymbol{\Phi}\}$ for all t.
B5. $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\mathbb{E}[\nabla \mathcal{L}_V(a_t, \boldsymbol{\theta})] < \infty$ and $\forall \boldsymbol{\phi} \in \boldsymbol{\Phi}$, $\mathbb{E}[\nabla \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})] < \infty$, for all t.
B6(a). $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\mathbb{E}[\nabla^2 \mathcal{L}_V(a_t, \boldsymbol{\theta})] < \infty$ and $\forall \boldsymbol{\phi} \in \boldsymbol{\Phi}$, $\mathbb{E}[\nabla^2 \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})] < \infty$, for all t.
(b). $\mathbb{E}[\nabla^2 \mathcal{L}_V(a_t, \cdot)]$ and $\mathbb{E}[\nabla^2 \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \cdot)]$ are continuous on $\{\boldsymbol{\Theta}, \boldsymbol{\Phi}\}$ uniformly for all t. (c). $\{\nabla^2 \log \mathcal{L}_V(a_t, \boldsymbol{\theta})\}$ and $\{\nabla^2 \log \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})\}$ obeys the strong ULLN.
B7. $\mathbb{E}[\log \mathcal{L}_V(a_t, \boldsymbol{\theta})]$ is uniquely maximized by $\boldsymbol{\theta}^*$ interior to $\boldsymbol{\Theta}$ and $\mathbb{E}[\log \mathcal{L}_C(a_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\phi})]$ is uniquely maximized by $\boldsymbol{\phi}^*$ interior to $\boldsymbol{\Phi}$.
B8. The double array $\{T^{-\frac{1}{2}} \nabla'_\theta \log \mathcal{L}_V(a_t, \boldsymbol{\theta}^*), T^{-\frac{1}{2}} \nabla'_\phi \log \mathcal{L}_C(a_t, \boldsymbol{\theta}^*, \boldsymbol{\phi}^*)\}$ obeys the central limit theorem.