Correlated Volatility Shocks*

Xiao Qiao[†,a] and Yongning Wang[‡,b]

[a,b]Booth School of Business, University of Chicago

August 25, 2016

Abstract

Commonality in idiosyncratic volatility cannot be completely explained by time-varying volatility. After removing the effects of time-varying volatility, idiosyncratic volatility innovations are still positively correlated. This result suggests correlated volatility shocks contribute to the comovement in idiosyncratic volatility. To capture this fact, we propose the Dynamic Factor Correlation (DFC) model, which fits the data well and captures the cross-sectional correlations in idiosyncratic volatility innovations. We decompose the common factor in idiosyncratic volatility (CIV) of Herskovic et al. (2014) into the volatility innovation factor (VIN) and time-varying volatility factor (TVV), and find VIN and TVV capture similar expected return variation and both contribute towards the asset pricing power of CIV.

Keywords: Volatility, GARCH models, Cross-section of stock returns, Idiosyncratic risk

*We thank Mark Huang for helpful discussions.

†Corresponding author. Email: xqiao0@chicagobooth.edu Website: https://sites.google.com/site/xiaoqiao10/

‡Email: ywang1@chicagobooth.edu

1 Introduction

The behavior of volatility has been a topic of interest for financial economists and econometricians because of its broad implications for volatility modeling and asset pricing. Recently a strand of literature has cropped up with a particular focus on idiosyncratic volatility. In a notable study, Herskovic et al. (2014) demonstrate that idiosyncratic volatilities of individual securities contain a common component which is priced in the cross section of stock returns.

An important question arises: What accounts for this commonality in idiosyncratic volatility? One possibility is that there is comovement in the time-varying idiosyncratic volatilities which drive the observed factor structure in idiosyncratic volatility. Alternatively, volatility innovations, or volatility shocks, could be correlated. Even with constant volatility, correlated volatility shocks can lead to a factor structure for idiosyncratic volatility. Herskovic et al. (2014) rationalize the asset pricing power of common idiosyncratic volatility through an incomplete markets heterogeneous-agent model, but does not attribute the comovement to time-varying volatility versus common volatlity shocks.

We show the commonality in idiosyncratic volatility cannot be fully explained by time-varying volatility alone. Correlated volatility shocks are necessary to capture this empirical regularity. To understand the importance of time-varying volatility, we fit univariate GARCH models to residual returns from factor models, and examine the GARCH volatility innovations. If time-varying volatility accounts for the commonality in idiosyncratic volatility, then the GARCH innovations should be cross-sectionally uncorrelated. We document that while the GARCH residuals are not autocorrelated, they are positively correlated in the cross section.

We propose a new multivariate GARCH model called the Dynamic Factor Correlation (DFC) model that captures the cross-sectional correlations of volatility innovations. Our model is most closely related to the Dynamic Conditional Correlation (DCC) model of Engle (2002), which focuses on the evolution of pairwise correlations among standardized residuals. Our DFC model extends the DCC model through introducing a factor structure in idiosyncratic volatility innovations which drives the comovement in those innovations. Both the DCC and DFC models examine the multivariate relationship after removing univariate volatility effects using GARCH. While Engle (2002) uses pairwise GARCH residuals to estimate a time-varying correlation model, we use the GARCH residuals from all individuals

to form a common factor that drives correlations in idiosyncratic volatility innovations. Our DFC model is also related to the Dynamic Equicorrelation (DECO) model of Engle and Kelly (2012). DECO assumes all pairwise correlations are equal on any given day, and this correlation changes over time. Our DFC model allows securities to have different correlations by carrying different exposures on the common factor. Under the restriction of equal exposures, the DFC model reduces to the DECO model.

The DFC model has a closed-form solution for its likelihood. As a result, estimation is straightforward even for a large number of assets. We use a two-stage quasi-maximum likelihood (QML) estimator for the DFC, which is consistent and asymptotically normal under regularity conditions, even if the model is misspecified. Empirically, the DFC model fits the data better than the DCC model on characteristic-sorted portfolios. We also propose an extension of the DFC model, the Group DFC model, to reduce the number of parameters and achieve better performance measured by the Bayesian Information Criterion (BIC).

Correlated volatility shocks cannot be diversified away in a portfolio, and may be a source of systematic risk. Our volatility innovations factor, VIN, appears to be related to the risk of cross-sectional return factors, but not macroeconomic risk. VIN is correlated with volatility and volatility innovations of the Fama and French. (1992, 2015) factors, but not highly correlated with the level, volatility, innovations, or volatility innovations of the Chicago Fed National Activities Index (CFNAI) or its subcomponents.

We construct our version of the common factor in idiosyncratic volatility factor (CIV) of Herskovic et al. (2014) using GARCH, and decompose it into VIN and a time-varying volatility factor, TVV. VIN captures more time-series variation of CIV. The correlation between VIN and CIV is 60%. In time-series regressions, VIN explains 36% of the variation of CIV, whereas TVV explains 1.6%. VIN forecasts next month's CIV with a 7% R^2, whereas TVV has a 2.4% forecasting R^2.

VIN and TVV are undiversifiable sources of risk, and are priced in the cross section of equity returns. We construct exposures of individual securities on VIN and TVV using 60-month windows. Quintile portfolios based on the exposure to VIN shows an average return difference of -1.92% per year between the highest and lowest portfolios. This difference becomes greater if we remove the market, size, and value factors as the CAPM and Fama and French. (1992) models are not able to explain the return spread. Stocks less exposed to volatility innovations earn higher average returns because their returns are relatively poor when volatility innovations are positive. Stocks with large exposure on VIN act as volatility

innovation hedges.

Quintiles based on the TVV exposure has an average return difference of -2.36% per year for the highest and lowest portfolios. The CAPM and the Fama and French. (1992) models are also not able to explain this average return spread. Stocks with high exposures to TVV effective hedges against spikes in time-varying volatility. Stocks with low exposures to TVV are more exposed to time-varying volatility risk, and earn higher returns on average.

Do VIN and TVV capture different aspects to cross-sectional average returns? We construct 25 portfolios based on exposures to both VIN and TVV to answer this question. We see return differences along both dimensions. A strategy taking a long position in the portfolio with the highest VIN and TVV loadings and a short position in the lowest VIN and TVV loadings portfolio generates an average annual return of -6.09%. Our results are consistent with those in Herskovic et al. (2014), who find a -5.4% spread in quintile portfolios formed on their CIV factor, a combination of VIN and TVV.

We also explore bivariate sorts of VIN and TVV loadings and market equity (ME), and find size does not diminish the average return spread across VIN and TVV portfolios. One exception is that in the smallest size quintile, the relationship between average returns and VIN or TVV loadings is reversed, although the results are not economically large.

Our paper fits into the literature on the properties of idiosyncratic volatility. Herskovic et al. (2014) document strong comovement in idiosyncratic volatility that is priced in the cross section, whereas we decompose this comovement into contributions from time-varying volatility and correlated volatility shocks. Campbell et al. (2001) investigate time-series variation in average idiosyncratic volatility, but does not address its cross-sectional properties. Duarte et al. (2014) look at the cross-sectional properties of the principal components of idiosyncratic volatility, but they do not try to distinguish common time-varying volatility with volatility innovation shocks. Kalnina and Tewou (2015) analyze cross-sectional dependence in idiosyncratic volatility with high-frequency data, but do not explore asset pricing implications. Ang et al. (2006) show that stocks with high idiosyncratic volatility earns low average returns, whereas Chen and Petkova (2012) find an average volatility factor can explain this factor. We show high sensitivity to idiosyncratic volatility innovations earn low average returns.

The paper is organized as follows. Section 2 provides motivating evidence that time-varying volatility is insufficient to describe the comovement in idiosyncratic volatility, and documents idiosyncratic volatility innovations are positively correlated. Section 3 proposes

the DFC model, derives its theoretical properties, and reports its empirical performance. Section 4 investigates the asset pricing implications of common volatility innovation shocks VIN and time-varying volatility TVV. Section 5 concludes.

2 Volatility Innovations are Correlated

Herskovic et al. (2014) show that idiosyncratic volatility contains a factor structure. They found that after they removed common factors from returns, idiosyncratic volatilities still tend to comove with one another. There are two possible sources for such comovement. First, there may be common drivers to time-varying idiosyncratic volatilities. The changes in these volatility drivers would lead to the rise and fall for all idiosyncratic volatilities. Second, idiosyncratic volatility innovations may be correlated. When the idiosyncratic volatility of one stock receives a positive innovation, other stocks are likely to experience it as well. Even in the absence of time-varying volatility, correlated volatility innovations can lead correlated idiosyncratic volatilities.

Our goal is to disentangle these two channels contributing to the factor structure in idiosyncratic volatility. As an exploratory exercise, we fit well-specified univariate volatility models to idiosyncratic volatilities to capture their time-variation. If time-varying volatility entirely accounts for the factor structure in idiosyncratic volatility, then the volatility model residuals should be uncorrelated. Instead, we find the volatility residuals are positively correlated.

2.1 Data

We use daily and monthly returns for Fama and French. (1992, 2015) factors and characteristic-sorted portfolios from Ken French's website[1]. We obtain daily and monthly returns on the Fama and French. (1992) 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.

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

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. All of the characteristic-sorted portfolio returns end on August 2015. We also obtain industry portfolio returns from July 1926 through August 2015 from French's website.

Individual stock returns, prices, and shares outstanding are from the Center for Research in Security Prices (CRSP) at the University of Chicago. For macroeconomic variables, we use the Chicago Fed National Activities Index (CFNAI) and its constituent components: production and income (PI); employment, unemployement, and hours (EUH); personal consumption and housing (CH); and sales, orders, and inventories (SOI). These series are from the Federal Reserve Bank of Chicago from March 1963 through August 2015. As an alternative measure of the macroeconomy, we use the Aruoba-Diebold-Scotti Business Conditions Index (ADS Index) from the Federal Reserve Bank of Philadelphia.

2.2 Correlated Volatility Innovations

We investigate the importance of common time-varying volatility versus correlated volatility shocks in producing comovements in idiosyncratic volatility. Our analysis starts with univariate and bivariate characteristic-sorted portfolios and industry portfolios. We construct idiosyncratic return series as the residuals after removing the Fama and French. (1992) factors. Similar to the results in Herskovic et al. (2014), we find the idiosyncratic volatilities of these portfolios are strongly correlated with one another.

For each of the idiosyncratic return series, we fit the best univariate GARCH model based on residual diagnostics. For a well-specified GARCH model, the GARCH residuals appear to be independently and identically distributed with unit variance. We search over different combinations of GARCH(p,q) and pick the model based on the GARCH residual behavior. Fitting GARCH models removes effects of time-varying volatility on the correlation of idiosyncratic volatility. If time-varying volatility is the main driver for common movements in idiosyncratic volatility, GARCH residuals should be cross-sectionally uncorrelated. We find that the GARCH residuals are positive correlated in the cross section, indicating that the factor structure in idiosyncratic volatility cannot be explained by time-varying volatility alone.

Table 1: **Volatility Innovation Correlations**
We fit GARCH models to return series and examine the correlation matrices of GARCH residuals. The left three columns display results for fitting GARCH models to raw returns series; the right three columns display results for fitting GARCH models on the idiosyncratic components after removing the Fama and French. (1992) factors. The top panel reports summary statistics for univariate portfolio sorts. the middle panel reports statistics for bivariate sorts. The bottom panel contains industry portfolios.

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

Table 1 presents summary statistics of the off-diagonal elements of correlation matrices of GARCH residuals, before and after removing common factors in returns[2]. The top panel contains univariate sorts, the middle panel contains bivariate sorts, and the bottom panel contains industry portfolios. For each panel, the left three columns present results for GARCH models fit to raw return series. The right three columns present results for GARCH models fit to the idiosyncratic return component for each portfolio after removing the Fama and French. (1992) factors.

Removing common factors reduces the correlations, but GARCH innovations in idiosyncratic volatility are still positively correlated. Before removing the Fama and French. (1992) factors, GARCH innovations are strongly positively correlated across portfolios for all univariate sorts. For example, the off-diagonal elements of the ME deciles range from 0.69 to 0.98 and have an average value of 0.91. Perhaps not surprisingly, after removing the common factors, the correlations decrease across the board. However, GARCH residuals are still positively correlated for all of the univariate portfolios. The range for ME deciles after removing common factors is 0.05 to 0.52, with an average value of 0.19. Other characteristic-sorted portfolios exhibit similar correlations.

Bivariate portfolios show similar patterns as univariate portfolios. Before removing common factors, GARCH residuals are strongly correlated. After removing common 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. After removing common factors, the range becomes -0.04 to 0.75, with an average value of 0.12. Industry portfolios also exhibit similar patterns.

To understand our observations in Table 1, we derive the covariance of volatility innovations under a factor model with GARCH volatilities in Appendix A. We assume raw returns follow a factor model in which both the factors and residual returns evolve with GARCH volatilities. Results in Appendix A suggest that comovement in volatility innovations of individual raw returns are driven by the pairwise correlations among volatility shocks of the factors and individual residual returns. The higher the correlation between idiosyncratic volatility innovations and the volatility innovations of the risk factors, the stronger the comovement observed in the GARCH residuals. This explains why we observe higher correlations in the left panel of Table 1 compared with those in the right panel.

If time-varying volatility were the sole contributor towards the idiosyncratic volatility

[2]Complete correlation matrices are omitted in the interest of space and are available upon request.

factor structure, we would expect to see largely uncorrelated GARCH residuals with an average value near zero. The fact that we see positively correlated GARCH residuals shows that time-varying volatility is not the only important force in driving idiosyncratic volatility comovements. The factor structure in idiosyncratic volatility cannot be explained by common time-varying volatility alone. With this fact in mind, we propose a GARCH-family model to better describe the data.

3 A Volatility Model with Correlated Shocks

Various multivariate GARCH models have been proposed to study the comovement of idiosyncratic volatilities motivated by the need for measuring correlations in financial risk management. The Dynamic Conditional Correlation (DCC) model introduced by Engle (2002) focuses on the evolution of correlations among standardized GARCH innovations and has become the benchmark multivariate volatility model for many purposes. Surveys by Tsay (2006) and Bauwens et al. (2006) provide discussions on DCC and related multivariate GARCH models. Engle and Kelly (2012) propose the Dynamic Equicorrelation (DECO) model to simplify the correlation structure by assuming all pairwise correlations are equal at any point in time. However, neither of these models imply the comovement in idiosyncratic volatility innovations we document in the previous section.

We propose the Dynamic Factor Correlation (DFC) model to capture the comovement in idiosyncratic volatility innovations. The DFC model extends Engle's DCC model by introducing a factor structure in standardized volatility innovations. Our model nests the DECO model of Engle and Kelly (2012), and is related to the constant conditional correlation (CCC) model of Bollerslev (1990). We estimate the DFC model using a two-stage quasi-maximum likelihood (QML) estimation, and demonstrate its superior performance relative to DCC in the data.

3.1 Dynamic Factor Correlation Model

Comovement in volatility innovations indicates the presence of common factors. These factors will impact the cross-section of residual returns. Let $r_{i,t}$ be the return of security i

at time t, $\mathbf{r}_t = (r_{1,t}, \cdots, r_{N,t})'$ be the vector of asset returns. $r_{i,t}$ follows a factor model,

$$r_{i,t} = \mathbf{f}_t' \beta_i + a_{i,t}, \tag{1}$$

where $\mathbf{f}_t = (f_{t,1}, \cdots, f_{t,K})'$ is a vector of K factors, β_i is a vector of factor loadings, and $a_{i,t}$ is the residual return. Suppose $h_{i,t}$ is the expectation of squared residual return conditional on the $t-1$ information set $\mathcal{F}_{t-1}$, $h_{i,t} \equiv \mathbb{E}[a_{i,t}^2 | \mathcal{F}_{t-1}] \equiv \mathbb{E}_{t-1}[a_{i,t}^2]$. $e_{i,t} \equiv a_{i,t}/\sqrt{h_{i,t}}$ is the standardized residual with $\mathbb{E}[e_{i,t}^2] = 1$. The econometrician may specify the volatility process of $h_{i,t}$. For example, the GARCH(1,1) model by Bollerslev (1986)

$$h_{i,t} = \gamma_{i,0} + \gamma_{i,1} a_{i,t-1}^2 + \gamma_{i,2} h_{i,t-1}, \tag{2}$$

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

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

Based on the comovement observed among idiosyncratic volatility innovations, we consider a factor structure embedded in the standardized residuals $e_{i,t}$ as follows

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

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 that affects $q_{i,t}$ for all i, and ξ_t is the factor loading of i on the 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 the 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 $\mathbb{E}[q_{i,t}^2] = 1$.

One simple way to empirically construct $\{h_{v,t}\}$ is to compute the average of a set of

standardized residuals from portfolios of interest:

$$h_{v,t} = \frac{1}{N}\sum_{i=1}^{N} e_{i,t-1}^2 \equiv \overline{\mathbf{e}_{t-1}^2} \tag{6}$$

Under Eq. (5), the conditional variance and covariance for $q_{i,t}$ follow

$$\begin{aligned} var_{t-1}[q_{i,t}] &= 1 + (h_{v,t} - 1)\xi_i^2 \\ cov_{t-1}(q_{i,t}, q_{j,t}) &= \xi_i \xi_j h_{v,t} \end{aligned} \tag{7}$$

Define $\mathbf{Q}_t = var_{t-1}(\mathbf{q}_t)$, we can represent the above relationship in matrix form.

$$\mathbf{Q}_t = \mathbf{\Lambda} + h_{v,t} \cdot \boldsymbol{\xi}\boldsymbol{\xi}', \tag{8}$$

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 $i \neq j$, then the correlation matrix of e_t is given by

$$\begin{aligned} \mathbf{R}_t &= cor_t(\mathbf{e}_t) \\ &= \tilde{\mathbf{Q}}_t^{-\frac{1}{2}} \mathbf{Q}_t \tilde{\mathbf{Q}}_t^{-\frac{1}{2}} \end{aligned} \tag{9}$$

Under the DFC model, the pairwise correlation between security i and j's standardized residuals, $\rho_t(i,j)$, can be expressed as the following

$$\begin{aligned} \rho_t(i,j) &= \frac{\sqrt{h_{v,t}}\xi_i\xi_j}{\sqrt{1 + (\sqrt{h_{v,t}} - 1)\xi_i^2}\sqrt{1 + (\sqrt{h_{v,t}} - 1)\xi_j^2}} \\ &= \frac{\xi_i\xi_j}{\sqrt{\xi_i^2 + \frac{(1-\xi_i^2)}{\sqrt{h_{v,t}}}}\sqrt{\xi_j^2 + \frac{(1-\xi_j^2)}{\sqrt{h_{v,t}}}}} \end{aligned} \tag{10}$$

Furthermore, we can obtain the correlation between squared standardized residuals as fol-

lows.

$$
\begin{aligned}
cor_t(e^2_{i,t-1}, e^2_{j,t-1}) &= \frac{\xi_i^2\xi_j^2\left(\mathbb{E}_{t-1}[v_t^4] - \sqrt{h_{v,t}}\right)}{2[1+(\sqrt{h_{v,t}}-1)\xi_i^2][1+(\sqrt{h_{v,t}}-1)\xi_j^2]} \\
&= \frac{\xi_i^2\xi_j^2\left(\frac{\mathbb{E}_{t-1}[v_t^4]}{\sqrt{h_{v,t}}} - 1\right)}{2[\xi_i^2+\frac{1-\xi_i^2}{\sqrt{h_{v,t}}}][\xi_j^2+\frac{1-\xi_j^2}{\sqrt{h_{v,t}}}]} \\
&= \frac{\xi_i^2\xi_j^2}{[\xi_i^2+\frac{1-\xi_i^2}{\sqrt{h_{v,t}}}][\xi_j^2+\frac{1-\xi_j^2}{\sqrt{h_{v,t}}}]}
\end{aligned} \tag{11}
$$

Among others, Longin and Solnik (2001), Ang and Chen (2002), and Cappiello et al. (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 empirical regularity. Eq. (10) and Eq. (11) imply that an increase in the common factor of volatility innovations, v_t, leads to stronger pairwise correlations in standardized residuals (given two securities with the same sign of factor loadings) and squared standardized residuals.

Based on Sherman-Morrison formula (see Sherman and Morrison (1950)), we have the following lemma.

Lemma 3.1. *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*

$$
\mathbf{Q}_t^{-1} = \boldsymbol{\Lambda}^{-1} - \frac{\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}\boldsymbol{\xi}'\boldsymbol{\Lambda}^{-1}}{1+h_{v,t}\boldsymbol{\xi}'\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}} \tag{12}
$$

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

$$
\begin{aligned}
det(\mathbf{Q}_t) &= det(\boldsymbol{\Lambda})[1+h_{v,t}\boldsymbol{\xi}'\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}] \\
&= [1+h_{v,t}\boldsymbol{\xi}'\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}]\prod_{i=1}^{N}(1-\xi_i^2)
\end{aligned} \tag{13}
$$

Assumption 3.1. $\forall i$, $\xi_i^2 < 1$.

Proposition 3.2. *Under Assumption 3.1, $\mathbf{R}_t$ is positive definite.*

Proof of Proposition 3.2 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.

Definition 3.1. *A time series* $\mathbf{a}_t$ *follows the Dynamic Factor Correlation (DFC) model if* $\mathbf{a}_{t|t-1} \sim \mathcal{N}(0, \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t)$*, where for all* t*,* $\mathbf{D}_t$ *is the diagonal matrix of conditional volatility of* $\mathbf{a}_t$*,* $\mathbf{R}_t$ *is given by Eq. (8) and Eq. (9) with* $\boldsymbol{\Lambda}$ *satisfying Assumption 3.1.*

The distinction between our DFC model and the Dynamic Conditional Correlation (DCC) model of Engle (2002) lies in the dynamics of $\mathbf{Q}_t$. $\mathbf{Q}_t$ processes considered in DCC include exponential smoother, often expressed as the following

$$\mathbf{Q}_t = (1-\lambda)(\mathbf{e}_{t-1}\mathbf{e}'_{t-1}) + \lambda\mathbf{Q}_{t-1} \tag{14}$$

or the MARCH model of Ding and Engle (2001)

$$\mathbf{Q}_t = \mathbf{S} \circ (\mathbf{J}_N - \mathbf{A} - \mathbf{B}) + \mathbf{A} \circ \mathbf{e}_{t-1}\mathbf{e}'_{t-1} + \mathbf{B} \circ \mathbf{Q}_{t-1} \tag{15}$$

whereas the DFC model uses a factor structure in Eq. (8). The evolution of $\mathbf{Q}_t$ is driven by $h_{v,t}$ for the DFC and does not contain explicit dependence on $\mathbf{Q}_{t-1}$.

With additional conditions, the DFC model reduces to existing multivariate GARCH models. If $\{v_t\}$ is conditional 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 turns into a special case of the Bollerslev (1990) constant conditional correlation (CCC) model with

$$\begin{aligned} var_{t-1}(\mathbf{a}_t) &= \mathbf{D}_t\mathbf{R}\mathbf{D}_t \\ \mathbf{R}(i,j) &= \xi_i\xi_j \end{aligned} \tag{16}$$

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

$$\mathbf{R}_t = (1-\rho_t)\mathbf{I}_N + \rho_t\mathbf{J}_N \tag{17}$$

with a non-negative equicorrelation

$$\rho_t = \frac{h_{v,t}\tilde{\xi}^2}{1 + (h_{v,t} - 1)\tilde{\xi}^2} \tag{18}$$

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

3.1.1 Estimation

We propose a two-stage quasi-maximum likelihood (QML) estimator for the DFC model. It will be consistent and asymptotically normally distributed under certain regularity conditions (see Appendix B), despite possibility of a misspecified model.

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

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{T}\sum_t \Big(\log|\mathbf{D}_t\mathbf{R}_t\mathbf{D}_t| + \mathbf{a}_t'\mathbf{D}_t^{-1}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1}\mathbf{a}_t \Big) \\
&= -\frac{1}{T}\sum_t \Big(2\log|\mathbf{D}_t| + \log|\mathbf{R}_t| + \mathbf{e}_t'\mathbf{R}_t^{-1}\mathbf{e}_t \Big) \\
&= -\frac{1}{T}\sum_t \Big(2\log|\mathbf{D}_t| + \mathbf{a}_t'\mathbf{D}_t^{-2}\mathbf{a}_t - \mathbf{e}_t'\mathbf{e}_t \Big) \\
&\quad -\frac{1}{T}\sum_t \Big(\log|\mathbf{R}_t| + \mathbf{e}_t'\mathbf{R}_t^{-1}\mathbf{e}_t \Big)
\end{aligned}$$

Let the parameters of univariate volatility be denoted by $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and the correlation parameters be denoted by $\boldsymbol{\phi} \in \boldsymbol{\Phi}$. Further define the volatility part of the log-likelihood as $\mathcal{L}_V(\boldsymbol{\theta})$ and the correlation part as $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$

$$\mathcal{L}_V(\boldsymbol{\theta}) = -\frac{1}{T}\sum_t \Big(2\log|\mathbf{D}_t| + \mathbf{a}_t'\mathbf{D}_t^{-2}\mathbf{a}_t - \mathbf{e}_t'\mathbf{e}_t \Big) \tag{19}$$

$$\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\frac{1}{T}\sum_t \Big(\log|\mathbf{R}_t| + \mathbf{e}_t'\mathbf{R}_t^{-1}\mathbf{e}_t \Big) \tag{20}$$

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

$$\mathcal{L} = \mathcal{L}_V(\boldsymbol{\theta}) + \mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) \tag{21}$$

where $\mathcal{L}_V(\boldsymbol{\theta})$ is proportional to the sum of individual log volatilities.

$$\mathcal{L}_V(\boldsymbol{\theta}) = -\frac{2}{T}\sum_t \sum_{i=1}^{N} \Big(\log(\sqrt{h_{i,t}}) \Big) \tag{22}$$

Also by Lemma 3.1, we can rewrite $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ as

$$\begin{aligned}\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) &= -\frac{1}{T}\sum_t \Big(\log\Big[1 + h_{v,t}\sum_{i=1}^N \frac{\xi_i^2}{1-\xi_i^2}\Big] + \sum_{i=1}^N \log\Big[1-\xi_i^2\Big] - \sum_{i=1}^N \log\Big[1 + (h_{v,t}-1)\xi_i^2\Big]\Big) \\ &\quad -\frac{1}{T}\sum_t \Big(\sum_{i=1}^N \Big[\frac{1+(h_{v,t}-1)\xi_i^2}{1-\xi_i^2}\Big] e_{i,t}^2 - \frac{1}{1 + h_{v,t}\sum_{i=1}^N \frac{\xi_i^2}{1-\xi_i^2}} \Big[\sum_{i=1}^N \frac{\sqrt{1+(h_{v,t}-1)\xi_i^2}}{1-\xi_i^2} e_{i,t}\xi_i\Big]^2\Big)\end{aligned} \tag{23}$$

In the first step of the QML estimation, we find the volatility parameter estimates that maximize $\mathcal{L}_V(\boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}}\{\mathcal{L}_V(\boldsymbol{\theta})\} \tag{24}$$

Then we plug the estimates into $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ in the second stage to obtain the maximizer $\hat{\boldsymbol{\phi}}$:

$$\hat{\boldsymbol{\phi}} = \arg\max_{\boldsymbol{\phi}}\{\mathcal{L}_C(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})\} \tag{25}$$

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 Eq. (23). Under regularity conditions (see Appendix B, B1 - B8), consistency of $\hat{\boldsymbol{\theta}}$ would ensure consistency of $\hat{\boldsymbol{\phi}}$ based on Theorem 6.11 in White (1994). Moreover, we can evaluate Eq. (23) using $h_{v,t}$, the standardized innovations $\hat{e}_{i,t}$, and the factor loadings $\{\xi_i\}$. From Lemma 3.1, we do not require numerical matrix inversions or determinants calculations for Eq. (23) which reduces the computational complexity of the likelihood optimization problem, making applications of the DFC to high-dimensional datasets desirable.

To reduce the number of estimated parameters for the DFC, we also consider imposing a group structure on individual factor loadings. Securities in the same group share the same loading on the common factor. Suppose there are K groups among N securities. Define g_i as the group that security i belongs to and $\mathcal{G}_k = \{i : g_i = k\}$. Within each group k, stocks share the same factor loading, denoted by $\tilde{\xi}_k$. Under this group structure, $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$ can be

rewritten as

$$\begin{aligned}\mathcal{L}_C^B(\boldsymbol{\theta},\boldsymbol{\phi}) &= -\frac{1}{T}\sum_t\Big(\log\Big[1+h_{v,t}\sum_{k=1}^K\frac{|\mathcal{G}_k|\tilde{\xi}_k^2}{1-\tilde{\xi}_i^2}\Big]+\sum_{k=1}^K|\mathcal{G}_k|\Big\{\log\Big[1-\tilde{\xi}_k^2\Big]-\log\Big[1+(h_{v,t}-1)\tilde{\xi}_k^2\Big]\Big\}\Big)\\ &\quad -\frac{1}{T}\sum_t\Big(\sum_{i=1}^N\Big[\frac{1+(h_{v,t}-1)\tilde{\xi}_{g_i}^2}{1-\tilde{\xi}_{g_i}^2}\Big]e_{i,t}^2-\frac{1}{1+h_{v,t}\sum_{i=k}^K\frac{|\mathcal{G}_k|\tilde{\xi}_k^2}{1-\tilde{\xi}_k^2}}\Big[\sum_{i=1}^N\frac{\sqrt{1+(h_{v,t}-1)\tilde{\xi}_{g_i}^2}}{1-\tilde{\xi}_{g_i}^2}e_{i,t}\tilde{\xi}_{g_i}\Big]^2\Big)\end{aligned} \tag{26}$$

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 $\hat{\rho}$

$$\min\frac{1}{2}\sum_{i\neq j}(\log|\xi_i|+\log|\xi_j|-\log|\hat{\rho}_{i,j}|)^2 \tag{27}$$

First order conditions for each i are

$$(N-2)\log|\xi_i|+\sum_{j=1}^N\log|\xi_j|-\sum_{j\neq i}\log|\hat{\rho}_{i,j}| = 0 \tag{28}$$

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

$$|\hat{\xi}_i|=\exp\Big\{\frac{1}{N-2}\Big(\sum_{j\neq i}\log|\hat{\rho}_{i,j}|-\frac{1}{2(N-1)}\sum_{i=1}^N\sum_{j=1}^N\log|\hat{\rho}_{i,j}|\Big)\Big\} \tag{29}$$

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 $\{\hat{\xi}_i\}$:

$$\hat{\xi}_i = \arg\min_{\xi_i\in\{-|\hat{\xi}_i|,|\hat{\xi}_i|\}}\sum_{j<i}(\xi_i\hat{\xi}_j-\hat{\rho}_{i,j})^2 \tag{30}$$

3.2 Simulations

We conduct Monte Carlo experiments to evaluate the performance of the DFC model. We simulate returns data for N assets over T periods under the hypothesis that the DFC is the true data generating process according to Definition 3.1. In the simulation, we explore

$N = 3, 10$ and $T = 1000, 5000$. We consider different choices of the factor loading vector $\boldsymbol{\xi}$ and repeat the simulations for $M = 1000$ times. For each set of simulation, we first generate standardized residuals $\{\mathbf{e}_t\}$ from the conditional correlation matrices $\{\mathbf{R}_t\}$ based on Eq. (6), (8) and (9). The initial value of $h_{v,t}$ is set to be the unconditional variance of $e_{i,t}$, i.e., $h_{v,1} = 1$. Next, we form a_t from a DCC model

$$\mathbf{D}_t^2 = \boldsymbol{\Omega} + \mathbf{A} \circ \mathbf{a}_{t-1}\mathbf{a}_{t-1}' + \mathbf{B} \circ \mathbf{D}_{t-1}^2 \tag{31}$$

where $\circ$ is the Hadamard product, or the element-by-element multiplication of two same sized matrices. We choose the parameters as $\boldsymbol{\Omega} = 0.003\mathbf{I}_N, \mathbf{A} = 0.1\mathbf{I}_N$, and $\mathbf{B} = 0.85\mathbf{I}_N$. These values are typical DCC estimates for stock returns.

For each simulated dataset, we estimate the DFC and DCC models, and compare their root mean squared error (RMSE) and mean absolute error (MAE) with respect to the true underlying process. In particular,

$$\begin{aligned} RMSE &= \sqrt{\frac{2}{MTN(N-1)} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\mathbf{R}}_t^{(m)}(i,j) - \mathbf{R}_t^{(m)}(i,j) \right)^2} \\ MAE &= \frac{2}{MTN(N-1)} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left| \hat{\mathbf{R}}_t^{(m)}(i,j) - \mathbf{R}_t^{(m)}(i,j) \right| \end{aligned} \tag{32}$$

Table 2 presents RMSEs and MAEs of fitted correlation matrices under DFC and DCC. On the left, we report the results of $N = 3$ in Panel A, B, and C, whereas the results of $N = 10$ are reported in the three panels on the right. In Panel A and Panel D, we set all ξ_i equal to 0.1, a weak exposure to volatility innovation factor. Panel C and Panl F use relatively stronger exposures, with all ξ_i's equal to 0.5. Panel B and Panel E contain the results of mixtures of factor loadings ranging from 0.2 to 0.6. The DFC model achieves better performance than DCC in both RMSE and MAE for all sets of the experiments with different combination of sizes, time periods, and factor loading parameters. Not surprisingly, errors tend to be smaller under larger samples. 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 respectively. Also, the advantage of the DFC model over the DCC model tends to be larger when the comovement among volatility innovations

is stronger, as the DFC model captures the comovement but the DCC model does not. RMSE and MAE appear to be similar for the DFC and DCC models in Panel A when the correlations are relatively low, whereas Panel C shows a larger difference when the exposures are stronger. Finally, a greater number of securities is associated with better fit, given the DFC model is the true data generating process.

Table 2: **Monte Carlo Simulations of the DFC Model**
We simulate return data for 3 or 10 assets over 1000 or 5000 periods under the assumption that the DFC is the true data generating process (Definition 3.1) with different choices of factor loadings. We use Eq. (6) for the $\{h_{v,t}\}$ process. For each set of parameter values, we repeat the simulation 1000 times. For each simulated dataset, we estimate the DFC and DCC models and compare their root mean squared error(RMSE) and mean absolute error (MAE) compared to the true process. In Panel A and Panel D, we generate volatility innovations with weak exposures to the common factor as all ξ_i equal to 0.1. Panel B and Panel E contain the results of mixtures of factor loadings ranging from 0.2 to 0.6. Panel C and Panel F report results for relatively stronger exposures with all ξ_i equal to 0.5.

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

3.3 Empirical Applications

3.3.1 Characteristic-Sorted Decile Portfolios

To evaluate its performance, we fit the DFC model on characteristic-sorted portfolios and compare the results with Engle's DCC model. We obtain residual returns by removing the Fama and French. (1992) factors from raw returns. For decile portfolios under each characteristic, we fit the DFC model in which idiosyncratic volatilities follow a GARCH(1,1)

model and $\{h_{v,t}\}$ follows Eq. (6). The DFC factor loading estimates and asymptotic standard errors are reported in Table 3.

Table 3: **DFC Estimation for Characteristic-Sorted Portfolios**
We estimate DFC model on residual returns of characteristic-sorted portfolios after removing Fama and French. (1992) three factors. Each column presents the 10 DFC factor loading estimates for the decile portfolios under one of the seven characteristics, with asymptotic t-statistics reported in parentheses. 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. For each characteristic, we estimate both the DFC and the DCC model, and compare their log-likelihood, shown on the bottom two rows.

	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

Characteristics-sorted decile portfolios appear to load significantly on the common factor of idiosyncratic volatility innovations. 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. For ME and BE/ME portfolios, middle portfolios tend to have negative loadings and extreme portfolio tend to have

positive loadings. This pattern suggests that standardized residuals of the middle portfolios tend to be negatively correlated with those in the extreme portfolios. For other characteristics the factor loadings are generally positive for low-characteristic portfolios and negative for high-characteristic portfolios. In terms of the magnitude of factor loadings, deciles of ME, LTRev, and Mom show larger exposure than other characteristics, which is consistent with their larger volatility innovation correlations in Table 1.

The bottom two rows of Table 3 compare the log-likelihood of the DFC model with that of the DCC model. For all seven of the characteristic-sorted deciles, the DFC model provides a superior fit compared to the DCC model. Although the DCC model allows for time-varying correlation in standardized residuals, it does not account for the comovement in the volatility innovations that we observe in the data. The DFC explicit models this comovement, which results in an improved fit compared to the DCC model.

Table 4: **Group DFC Estimation for Characteristic-Sorted Portfolios**
We estimate the Group DFC model on residual returns of characteristic-sorted portfolios after removing Fama and French. (1992) three factors. Groups are determined based on the closeness of factor loadings estimates in Table 3. Each cell reports the deciles that belong to the same group. For each characteristic, we estimate both DFC and Group DFC model, and compare their Bayesian Information Criterion (BIC).

	ME	BE/ME	LTRev	OP	Inv	Mom	STRev
$G1$	1,2	1	1	1,2,3	1,2,3,4	1	1,3
$G2$	3	2	2,3,4	4,5,6	5	2,3	2
$G3$	4	3	5	7	6	4	4
$G4$	5	4,5,6	6	8	7	5	5
$G5$	6	7,8	7	9,10	8,9,10	6	6
$G6$	7,8	9	8			7	7,8,9
$G7$	9	10	9,10			8	10
$G8$	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

Based on the closeness of factor loading estimates in Table 3, we group deciles with similar exposures into the same group and estimate the Group DFC model using Eq. (26). Table 4 provides the classification under each characteristic and reports the BICs 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). The disadvantage of the Group DFC is that we lose interpretation in the group setup. It is not clear what each group represents, and why certain portfolios should belong in the same group.

3.3.2 Industry Portfolios

We evaluate the usefulness of the DFC model in a risk management setting by comparing minimum variance portfolios formed using historical covariance estimates and variations of the DFC model. If the DFC model fits the data well, we should be closer to the true *ex ante* minimum variance portfolios using the DFC model for covariance estimates. We use monthly returns of 49 industry portfolios from January 1946 to June 2015 from Ken French's website[1]. For Group DFC models, we follow French's industry classification and consider 5-group, 10-group, and 30-group classifications based on CRSP SIC code. Group definitions are provided in Appendix C.

For all competing models, we evaluate the out-of-sample risk management performances of the mean-variance optimization portfolios of Markowitz (1952) based on the return time series of N securities, which has expected return $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma}_t$ at each time t. We calculate the portfolio weights that solve the problem

$$\min_{\mathbf{w}_t} \mathbf{w}_t' \boldsymbol{\Sigma}_t \mathbf{w}_t \quad s.t. \quad \mathbf{w}_t' \mathbf{1}_N = 1 \tag{33}$$

$$\mathbf{w}_t' \boldsymbol{\mu}_t \geq \mu_0 \tag{34}$$

Define $A_t = \mathbf{1}_N' \boldsymbol{\Sigma}_t^{-1} \mathbf{1}_N$, $B_t = \mathbf{1}_N' \boldsymbol{\Sigma}_t^{-1} \mu_t$, and $C_t = \mu_t' \boldsymbol{\Sigma}_t^{-1} \mu_t$, we obtain the global minimum variance (GMV) portfolio weights by solving Problem (33)

$$\mathbf{w}_t^{GMV} = \frac{1}{A_t} \boldsymbol{\Sigma}_t^{-1} \mathbf{1}_N \tag{35}$$

Solution to Problem (33) given (34) results in the minimum variance (MV) portfolio weights

$$\mathbf{w}_t^{MV}(\mu_0) = \frac{C_t - \mu_0 B_t}{A_t C_t - B_t^2} \boldsymbol{\Sigma}_t^{-1} \mathbf{1}_N + \frac{\mu_0 A_t - B_t}{A_t C_t - B_t^2} \boldsymbol{\Sigma}_t^{-1} \mu_t \tag{36}$$

In our analysis, Σ_t is evaluated based on different volatility or correlation models. For minimum variance portfolios, we entertain three values for required annualized returns μ_0:

5%, 7.5%, and 10%.

We consider the following candidate models for Σ_t. In a moving window of past five years, we have

- **Model 1** Unconditional covariance: Unconditional covariance on raw returns.
- **Model 2** DFC: Fit DFC on raw return time series and predict $\mathbf{D}_t$ 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: In the first stage, regress raw returns on market excess return r_{mt} to obtain residual returns. In the second stage, fit a univariate GARCH(1,1) on r_{mt} and estimate a DFC model on the residual returns. Then predict covariance matrix as $h_{mt}\boldsymbol{\beta}_t\boldsymbol{\beta}_t' + \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t'$ where h_{mt}, $\mathbf{D}_t$ and $\mathbf{R}_t$ come from GARCH and DFC forecasts, $\boldsymbol{\beta}_t$ is the factor loading vector.
- **Model 4** FF3 DFC: In the first stage, regress raw returns on Fama and French. (1992) three factors, namely market excess return, SMB, and HML, and obtain residual returns. In the second stage, fit a univariate GARCH(1,1) on r_{mt}, SMB, and HML, respectively, and estimate a DFC model for the residual returns. Then predict covariance matrix as $\mathbf{B}_t\text{diag}\{h_{mt}, h_{SMB,t}, h_{HML,t}\}\mathbf{B}_t' + \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t'$, where $\mathbf{B}_t$ is the factor loading matrix.
- **Model 5-7** 30-Group, 10-Group, and 5-Group FF3 DFCs: First stage is the same as in FF3 DFC. In the second stage, fit a univariate GARCH(1,1) on r_{mt}, SMB, and HML, respectively, and estimate a Group DFC model for the residual returns based on the 30-Group, 10-Group, and 5-Group classification columns in Table A1 respectively, then predict the covariance matrix.

Under each of the above seven combinations of factor and correlation models, we obtain out-of-sample optimal minimum variance portfolio weights as follows. At the beginning of each month, we use the data in a rolling window of previous five years and

1. For Model 3-7, run regression to estimate factor loadings and obtain residual returns.
2. For Model 3-7, estimate volatily processes for each factor.
3. Estmiate volatility models on raw returns for Model 2, and on residual returns for Model 3-7.

4. Estimate covariance for Model 1, and correlation models for Model 2-7.

5. Use factor loadings from Step 1, one-month ahead factor volatility forecasts from Step 2, individual volatility predictions from Step 3, and correlation forecasts from Step 4 to predict the covariance matrix for the next month.

6. Find optimal weights using the covariance forecasts and expected return, which is set as the historical mean in the moving window.

7. Record portfolio realized return at the end of the current month.

Table 5 presents the annualized standard deviation of realized returns in percentage for all seven models. Panel A reports the results for the whole sample period. In Panel B and Panel C, we show the performance of competing models in the last 50 (1965-2015) and 20 years (1995-2015), respectively. In each panel, we report the standard deviation of GMV portfolio returns as well as MV portfolios with $\mu_0 = 5\%$, 7.5%, and 10%. The model that achieves the lowest standard deviation in each column is shown in bold.

Without a factor structure in the mean equation, DFC model achieves the much lower standard deviation than unconditonal covariance on all four GMV and MV portfolios in each panel. Based on the results for the whole sample, standard deviation of GMV portfolio reduces from 0.694% to 0.489% if we consider DFC model instead of unconditional covariance on raw returns. DFC also outperforms unconditional covariance in MV portfolios. For MV portfolios with required annual returns equal 10%, 7.5%, and 5%, Non-Factor DFC achieves standard deviations of 0.462%, 0.462%, and 0.468% respectively, whereas unconditional covariance method generates 0.681%, 0.682%, 0.691% under corresponding μ_0 values. We find the same results across other sample periods reported in Panel B and C.

Using CAPM and Fama and French. (1992) three-factor model in the mean equation further improve the risk management performance of DFC. For example, standard deviation decreases to 0.431% when CAPM is used for GMV portfolio in Panel A. Under Fama and French. (1992) structure, standard deviation further drops to 0.407%. FF3 DFC also outperforms Non-Factor DFC and CAPM DFC for MV portfolios as well, with standard deviations of 0.405%, 0.402%, and 0.403% under $\mu_0 = 10\%$, 7.5%, and 5% respectively. The same results are observed for the last 50 and 20 years.

Considering Group DFC could further achieve lower out-of-sample standard deviation. In both Panel A and B, the lowest standard deviations are achieved by Group DFC models

Table 5: **Out-of-sample Standard Deviations of Minimum Variance Portfolio Return Time Series: Factor and Correlation Models on 49 Industry Portfolios** This table reports the annualized standard deviation of realized returns in percentage for seven competing models that are entertained to predict the one-month ahead forecast of covariance matrix of 49 industry portfolio returns from 1946 to 2015. In Panel A, we show the results for the whole sample. Panel B and Panel C compare the performance of candidate models in the last 50 (1965-2015) and 20 years (1995-2015), respectively. In each panel, we report the standard deviation of GMV portfolio returns as well as MV portfolios with $\mu_0 = 5\%$, 7.5%, and 10%. The model achieving the lowest standard deviation in each column is represented in bold.

Model	GMV	MV($\mu_0 = 10\%$)	MV($\mu_0 = 7.5\%$)	MV($\mu_0 = 5\%$)
		Panel A: 1946 - 2015		
Unconditional covariance	0.694	0.681	0.682	0.691
Non-Factor DFC	0.489	0.462	0.462	0.468
CAPM DFC	0.431	0.413	0.412	0.416
FF3 DFC	0.407	0.405	0.402	0.403
FF3 30-Group DFC	**0.405**	0.402	0.400	**0.402**
FF3 10-Group DFC	0.406	**0.401**	**0.399**	0.402
FF3 5-Group DFC	0.408	0.405	0.402	0.404
		Panel B: 1965 - 2015		
Unconditional covariance	0.763	0.747	0.750	0.759
Non-Factor DFC	0.530	0.500	0.501	0.507
CAPM DFC	0.471	0.449	0.449	0.455
FF3 DFC	0.439	0.439	0.437	0.439
FF3 30-Group DFC	**0.436**	0.433	**0.431**	**0.435**
FF3 10-Group DFC	0.439	**0.432**	0.432	0.436
FF3 5-Group DFC	0.442	0.437	0.436	0.439
		Panel C: 1995 - 2015		
Unconditional covariance	0.786	0.753	0.757	0.770
Non-Factor DFC	0.461	0.451	0.457	0.470
CAPM DFC	0.444	0.425	0.428	0.436
FF3 DFC	**0.392**	0.395	**0.395**	**0.402**
FF3 30-Group DFC	0.395	**0.394**	0.396	0.403
FF3 10-Group DFC	0.402	0.397	0.399	0.408
FF3 5-Group DFC	0.398	0.397	0.398	0.405

across all four minimum variance portfolios. For MV portfolio with $\mu_0 = 10\%$, FF3 30-Group DFC outperforms FF3 DFC in all three panels. FF3 30-Group DFC generates standard deviations of 0.402% in Panel A, 0.433% in Panel B, and 0.394% in Panel C, compared with 0.405%, 0.439%, and 0.395% in corresponding sample periods under FF3 DFC.

In summary, DFC model improves the risk management performance of the minimum variance portfolio returns on the 49 industry portfolios. Compared with unconditional covariance method, DFC achieves lower out-of-sample standard deviations of the optimal portfolios when there is no factor structure in the mean equation. Combining factor models such as CAPM and Fama and French. (1992) with DFC further reduces the standard deviation. Finally, our results suggest that using Group DFC could achieve similar or even better risk management performance with fewer unknown parameters.

4 Correlated Volatility Shocks and Asset Pricing

Asset pricing theory states that risk premiums should be associated with systematic, undiversifiable risks in the economy. Market beta is the classic example. We expect a risk premium to be associated with beta because it is a characteristic that does not diminish in importance when we form large, diversified portfolios. Commonality in volatility innovations is a plausible priced state variable because it also represents systematic risk that cannot be diversified away with more securities. Motivated by this idea, we examine the asset pricing properties of correlated volatility shocks.

We construct the volatility innovation factor (VIN) from factor models. We remove the Fama and French. (1992) factors from characteristic-sorted portfolios, then fit GARCH models to residual returns. We form our candidate state variable by taking an equal-weight average of the GARCH residuals. One natural way of extracting common components from a panel of correlated variables is principal component analysis (PCA). However, estimates from PCA may be unstable in different samples. An alternative is to take a simple cross-sectional average of the panel, which does not suffer from estimation errors. We compare the largest principal component from the GARCH residuals with their simple average, and find the two series are 90% correlated. We use the equal-weight average for our asset pricing tests; the results are unchanged using the largest principal component.

4.1 Time Series Relationships

We first examine the time-series properties of VIN. Because VIN is constructed from GARCH residuals, it is natural to look at the relationship between VIN and second moments. Table 6 reports the correlations between VIN and volatility measures of the Fama and French. (1992, 2015) factors. We construct month realized volatilities (RV) of the factors using daily data. We also construct realized volatility innovations as residuals from autoregressive models chosen by the Akaike information criterion (AIC). Each realized volatility series may have its own AIC-selected AR order.

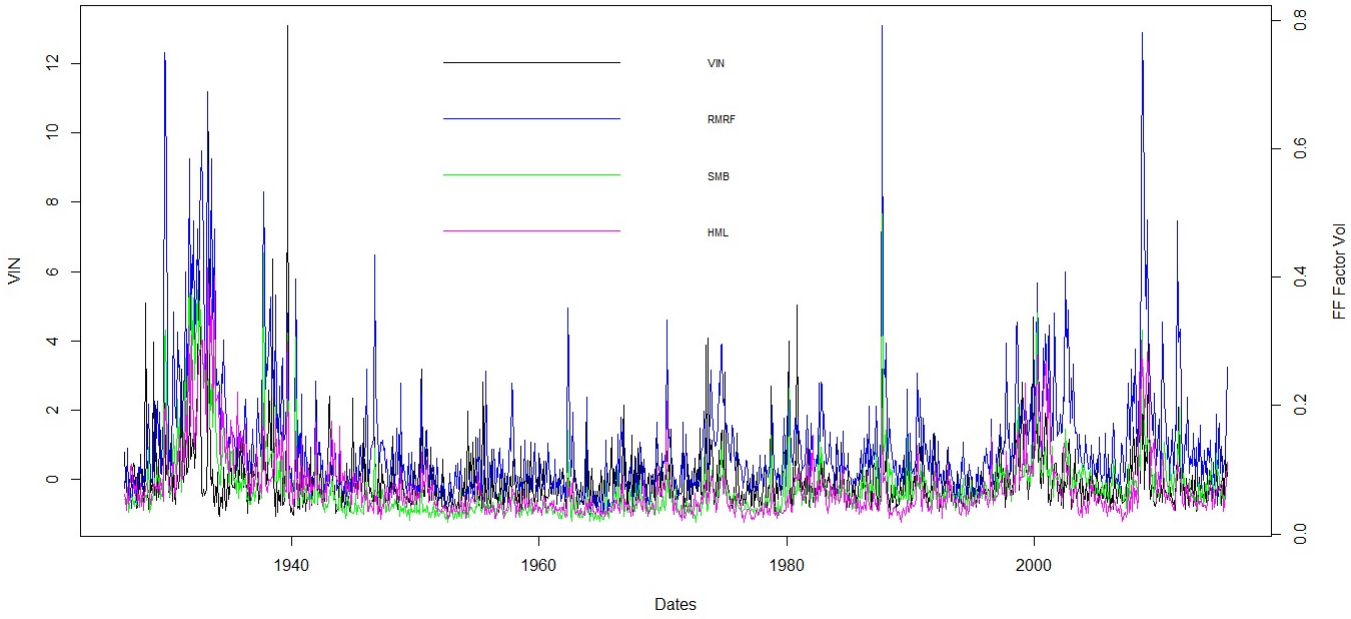
The top panel contains results for the long sample of the Fama and French. (1992) factors, from 1926 to 2015. VIN is positive correlated with the realized volatilities of the market, size, and value factors. The correlations are large, ranging from 0.35 to 0.39. VIN is also positively correlated with the realized volatility innovations. Correlations with RV innovations are generally smaller compared to the RV levels. The correlations with HML's RV or RV innovations are almost identical.

Table 6: **Correlations Between Factor Volatility Measures and VIN**
We report correlations between VIN and realized volatility (RV) and realized volatility innovations of the Fama and French. (1992) 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 (AIC). The top panel displays values for the Fama and French. (1992) 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

The bottom panel contains the correlations between VIN 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[3]. VIN is positively correlated with the realized volatility of each of the five factors, with values ranging from 0.34 to 0.43. VIN is positively but less strongly correlated with the RV innovations, with values from 0.22 to 0.34.

Figure **??** plots VIN alongside the Fama and French. (1992) factor volatilities. For ease of viewing, VIN is standardized and follows the scale on the left, whereas factor volatilities are not standardized and follow the scale on the right. VIN appears to comove with the factor volatilities, but only some of the time. For example, around Black Monday on October 19, 1987, VIN spikes up to its maximum value, but none of the factor volatilities spiked to their max values. Periods with particularly high VIN includes the Great Depression, Black Monday of 1987, busting of the tech bubble in the early 2000s, and the recent Great Recession.



The undiversifiable nature of VIN could have macroeconomic origins – it is possible that VIN may be a proxy for macroeconomic risk. We investigate this possibility in Table 7. Our macroeconomic variables are the CFNAI and its constituent components PI, EUH,

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

CH, and SOI. We examine the correlations between VIN and four transformations of these variables. First, we look at how VIN 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.

Table 7: **Correlations Between Macroeconomic Measures and VIN**
We report correlations between VIN 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

VIN is negatively correlated with the raw CFNAI and its components, indicating that it may underline some type of macroeconomic risk. The correlation magnitudes for the components are similar with the exception of CFNAI CH, which is only weakly related to VIN. VIN and macroeconomic volatility are somewhat positively correlated. Again, the correlation with CFNAI CH is particularly weak. VIN does not seem to be related to macroeconomic innovations, as shown in the bottom panel, and VIN negatively comoves with macroeconomic volatility innovations. As an alternative to the CFNAI, we use the Aruoba-Diebold-Scotti (ADS) Business Conditions Index, and find similar results. We do not find strong results that VIN reflects macroeconomic risk.

VIN partially captures the variation in Fama and French. (1992, 2015) factor volatilities. The large positive correlations with return factor volatilities suggest the possibility of a premium associated with VIN. VIN 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, VIN appears to capture the undiversifiable risk in the cross-sectional return factors, not macroeconomic shocks.

4.2 Portfolios Sorts on Volatility Innovations and Time-Varying Volatility

Herskovic et al. (2014) document that firms' exposure to idiosyncratic volatility shocks helps explain cross-sectional differences in average stock returns. Leaning on their results, we further investigate if the asset pricing power comes from correlated volatility shocks or time-varying volatility. By decomposing a stock's idiosyncratic volatility shock into (standardized) volatility innovation and time-varying volatility, we disentangle the contributions from these two parts to expected returns.

LHS of Eq.(37) measures firm i's idiosyncratic volatility shock. It is equal to the difference between realized idiosyncratic volatility and its expectation. Herskovic et al. (2014) construct their common factor in idiosyncratic volatility (CIV), which they show is priced in the cross section, from this difference. We rewrite the volatility shock as the product of expected time-varying volatility (TVV) and standardized volatility innovation (VIN).

$$\begin{aligned} \underbrace{\sigma^2_{i,t} - \mathbb{E}_{t-1}[\sigma^2_{i,t}]}_{\text{Idiosyncratic Volatility Shock}} &= \frac{\sigma^2_{i,t} - \mathbb{E}_{t-1}[\sigma^2_{i,t}]}{\mathbb{E}_{t-1}[\sigma^2_{i,t}]} \cdot \mathbb{E}_{t-1}[\sigma^2_{i,t}] \\ &\equiv \underbrace{\tilde{\nu}_{i,t}}_{\text{Volatility Innovation}} \cdot \underbrace{\mathbb{E}_{t-1}[\sigma^2_{i,t}]}_{\text{Time-Varying Volatility}} \end{aligned} \tag{37}$$

Let $\bar{\tilde{\nu}}$ and $\bar{\sigma}^2$ be the expected values for $\tilde{\nu}_{i,t}$ and $\sigma^2_{i,t}$ across all stocks. A first-order Taylor expansion around $(\bar{\tilde{\nu}}, \bar{\sigma}^2)$ gives a linear approximation of Eq.(37).

$$\tilde{\nu}_{i,t} \cdot \mathbb{E}_{t-1}[\sigma^2_{i,t}] \approx -\bar{\tilde{\nu}} \cdot \bar{\sigma}^2 + \bar{\tilde{\nu}} \cdot \mathbb{E}_{t-1}[\sigma^2_{i,t}] + \tilde{\nu}_{i,t} \cdot \bar{\sigma}^2 \tag{38}$$

Suppose at time t, the total number of stocks equals N_t and the weight of security i in forming the CIV factor is $w_{i,t}$. Then based on Eq.(37) and Eq.(38), the CIV factor can be written as follows

$$\begin{aligned} CIV_t &= \sum_{i=1}^{N_t} w_{i,t}\Big(\sigma_{i,t}^2 - \mathbb{E}_{t-1}[\sigma_{i,t}^2]\Big) \\ &\approx \sum_{i=1}^{N_t} w_{i,t}\Big(-\bar{\tilde{\nu}}\cdot\bar{\sigma}^2 + \bar{\tilde{\nu}}\cdot\mathbb{E}_{t-1}[\sigma_{i,t}^2] + \tilde{\nu}_{i,t}\cdot\bar{\sigma}^2\Big) \\ &= -\bar{\tilde{\nu}}\cdot\bar{\sigma}^2 + \bar{\sigma}^2\underbrace{\sum_{i=1}^{N_t} w_{i,t}\tilde{\nu}_{i,t}}_{VIN_t} + \underbrace{\bar{\tilde{\nu}}\sum_{i=1}^{N_t} w_{i,t}\mathbb{E}_{t-1}[\sigma_{i,t}^2]}_{TVV_t} \end{aligned} \tag{39}$$

Therefore, shocks to idiosyncratic volatility are driven by time-varying volatility and volatility innovations.

For our empirical results, the VIN and TVV factors are constructed as follows. We first fit the Fama and French. (1992) model to each CRSP stock with 60-month rolling windows. We then fit GARCH models to each residual series to obtain security-level VIN and TVV. Whereas citethkln2014 form their volatility innovations using monthly realized volatility, we use GARCH. Common factors in VIN and TVV are formed each month as the equally-weighted average in their security-level values.

To distinguish our GARCH-based CIV with the one in citethkln2014, we will refer to ours as the GARCH-CIV. To understand the relative importance of VIN and TVV as components of the GARCH-CIV, we run explanatory and forecasting regressions on the GARCH-CIV using VIN and TVV. Table 8 reports the regression results as well as correlations among these three factors. In Panel A, Columns (1) to (5) reports contemporaneous OLS regressions of GARCH-CIV from Eq.(39) against VIN, TVV factors and their interactions. Newey-West t-statistics with five lags are shown.

Columns (1) and (2) present estimates of the relative importance of VIN and TVV for CIV shocks. Column (1) shows that VIN captures 37% of the contemporaneous movements in GARCH-CIV, while column (2) shows that TVV captures under 2%. It is clear that VIN dominates in explaining the variation of GARCH-CIV. The negative coefficient on TVV suggests that a higher time-varying volatility level tend to be associated with a lower GARCH-CIV. When both VIN and TVV are included in the regression in Column (3), they

Table 8: **GARCH-CIV, VIN and, TVV: Time-Series Regressions and Correlations** Panel A reports contemporaneous (Columns (1) - (5)) and predictive (Columns (6) - (10)) time-series regressions of GARCH-CIV against VIN and TVV factors. The explanatory variables are VIN, TVV, and VIN · TVV. All coefficients are multiplied by 100. Newey-West t-statistics with five lags are reported. The estimation period is from January 1946 to December 2014. Panel B reports the sample correlation among GARCH-CIV, VIN, and TVV factors.

Panel A: Time-Series Regressions										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	-0.05	0.08	0.13	-0.06	0.19	-0.04	0.11	0.13	-0.05	0.14
	(-2.33)	(1.62)	(3.03)	(-3.24)	(4.13)	(-1.87)	(2.27)	(3.14)	(-2.03)	(3.28)
VIN_t	1.68		1.73		0.23					
	(10.66)		(10.02)		(0.95)					
TVV_t		-7.27	-10.59		-14.87					
		(-1.76)	(-3.63)		(-4.61)					
$\text{VIN}_t \cdot \text{TVV}_t$				79.67	77.23					
				(7.66)	(4.81)					
VIN_{t-1}						(0.73)		(0.78)		(0.55)
						(6.27)		(6.73)		(2.43)
TVV_{t-1}							-8.98	-10.48		-11.14
							(-2.33)	(-3.20)		(-3.33)
$\text{VIN}_{t-1} \cdot \text{TVV}_{t-1}$									27.51	11.85
									(3.01)	(0.83)
R^2	36.54	1.57	39.85	44.18	51.01	6.96	2.41	10.21	5.27	10.47

Panel B: Factor Correlations		
Factor	VIN	TVV
CIV Shocks	0.60	-0.13
VIN		0.09

account for 40% of the variation in the GARCH-CIV, and both coefficients are significant. Column (4) reports a contemporaneous regression of GARCH-CIV on the product of VIN and TVV. The R^2 of the regression is 44%, which indicates that the product of VIN and TVV captures more variation in GARCH-CIV than a linear combination of them, as Eq.(37) suggests. In Column (5), we regress the GARCH-CIV against VIN, TVV, and their product, resulting in a R^2 of 51%. The coefficient on VIN is no longer significant when we include the produce of VIN and TVV.

Columns (6) to (10) show the results of predictive regressions for the GARCH-CIV. In Columns (6) and (7), we report the simple regressions of CIV shocks on lagged VIN and TVV factors respectively. The results suggest both VIN and TVV have significant forecasting power in predicting the GARCH-CIV. In particular, VIN shows a higher forecasting power (6.95% R^2) than TVV (2.41% R^2). A higher VIN factor value is associated with a higher GARCH-CIV next period, whereas an increase in TVV is associated with lower value of next period's GARCH-CIV. Including both factors in column (8) improves the predictive regression R^2 to 10%. Column (9) reports the predictive regression of GARCH-CIV against the product of VIN and TVV, and the explanatory power of the model is 5%. Evidently, the produce of VIN and TVV is a good explanatory variable in the contemporaneous setting, but does not forecast the GARCH-CIV as well as including VIN and TVV separately. Regressing GARCH-CIV against lagged VIN, TVV, and their product, we obtain a R^2 of 10%.

Panel B of Table 8 reports the correlation among GARCH-CIV, VIN, and TVV. The results suggest a strong relationship between the GARCH-CIV and VIN, with a correlation of 0.60. A higher value of aggregate standardized volatility innovation is associated with a higher value of the idiosyncratic volatility shock. Consistent with the regression results, TVV is negatively correlated to the GARCH-CIV, with a correlation of -0.13. As time-varying volatility increases, the unexpected change in volatility tends to be smaller. VIN and TVV are only slightly positively correlated. Overall, although it appears VIN more closely captures the GARCH-CIV and forecasts better, both VIN and TVV are important components of the GARCH-CIV.

Next we explore the asset pricing power of VIN and TVV. For each month from January 1946 to December 2014, for each security, we regress its monthly excess returns on VIN and TVV separately using a 60-month window to obtain the VIN and TVV exposures. A stock's exposure to VIN and TVV factors are referred as its VIN-beta and TVV-beta respectively.

We examine univariate and bivariate sorted portfolios on VIN-beta and TVV-beta. We

start with univariate portfolios based on VIN and TVV exposures. The results are reported in Table 9. We sort stocks based on their VIN-beta or TVV-beta and form five value-weighted portfolios. The portfolios are rebalanced monthly. We report the average and abnormal returns of each portfolio, as well as the average returns on a strategy that invests in the highest quantile and shorts the lowest quantile.

Panel A reports the results for portfolios formed on VIN-beta. Average returns tend to decrease in stock's exposure to VIN. Stocks with a high VIN-beta pays off when volatility innovation is high, effectively serving as insurance for particularly high volatility innovations. All else equal, investors prefer holding these securities. As a result, their prices are bid up and expected returns are bid down. By similar logic, the low VIN-beta stocks are more exposed to volatility innovations risk, and earn higher average returns. The long-short portfolio has a CAPM alpha of -4.25% with a t-statistic of -2.35 and Fama and French. (1992) model alpha of -3.32% with a t-statistic of -1.84, suggesting the differences in average returns is economically and statistically large.

Table 9: **Single-Sorted Portfolios Formed on VIN Beta and TVV Beta**
We report average excess returns, CAPM, and Fama and French. (1992) alphas in annual percentages for portfolios on the basis of one-way sorts on monthly VIN beta and TVV beta separately using all CRSP stocks for 1946-2014 sample. Panel A reports value-weighted average excess returns and alphas based on VIN-beta sorts. Panel B shows results of TVV-beta sorted portfolios.

VIN beta	1 (Low)	2	3	4	5 (High)	5-1	t(5-1)
Panel A: One-way sorts on VIN-beta							
$\mathbb{E}[R]-r_f$	11.52%	10.18%	9.60%	9.52%	9.60%	-1.92%	-1.02
α_{CAPM}	4.66%	3.65%	2.64%	2.04%	0.42%	-4.25%	-2.35
α_{FF}	3.84%	2.63%	1.75%	1.25%	0.52%	-3.32%	-1.84
TVV beta	1 (Low)	2	3	4	5 (High)	5-1	t(5-1)
Panel B: One-way sorts on TVV-beta							
$\mathbb{E}[R]-r_f$	11.64%	9.36%	9.74%	9.41%	9.28%	-2.36%	-1.40
α_{CAPM}	3.82%	2.67%	2.93%	2.00%	0.39%	-3.43%	-2.04
α_{FF}	3.82%	2.02%	2.17%	1.08%	0.24%	-3.58%	-2.16

Panel B shows results of quintile portfolios sorted by TVV-beta and the spread portfolio. Similar to the results for VIN-beta, average returns are decreasing in TVV exposure. Similar logic follows. Stocks with high TVV-betas have high returns when time-varying volatility is

high, and are relatively safe by this measure. Stocks with low TVV-betas are more exposed to time-varying volatility risk, and earn higher returns as a result. The long-short portfolio has an average return of -2.36%, although it is not statistically significant. Its CAPM alpha is -3.43% with a t-statistic of -2.04 and Fama and French. (1992) alpha of -3.58% with a t-statistic of -2.16. Similar to the case for VIN-beta, portfolios forms on TVV-beta exhibit economically large differences in average returns.

Table 10 reports the results for VIN-beta and TVV-beta double-sorted portfolio. Stocks are first sorted into quintiles based on their VIN-beta, then within each VIN-beta quintile, we sort stocks into five TVV-beta portfolios. From the 25 portfolios, we form a strategy that takes long positions in stocks in the group with both the highest VIN-beta and TVV-beta, and short those in the portfolio with the lowest VIN-beta and TVV-beta.

Table 10: **Portfolios Formed on VIN Beta and TVV Beta**
Performance of the spread portfolio double-sorted on the basis of monthly VIN beta and TVV beta using all CRSP stocks: alphas in annual percentages, slope coefficients (t-statistics) and R^2 from month-by-month regression of spread portfolio returns on risk factors for 1946-2014 sample. Panel A, B, C, and D report value-weighted excess returns, CAPM, Fama and French. (1992), and Carhart (1997) model results respectively.

Model	Intercept	Mkt.RF	SMB	HML	WML	R^2
Panel A: Raw Return	-6.09% (-2.31)					
Panel B: CAPM	-8.52% (-3.28)	0.33 (6.56)				5.0%
Panel C: FF 3-factor	-8.03% (-3.10)	0.24 (4.56)	0.41 (5.26)	-0.12 (-1.46)		8.6%
Panel D: Carhart 4-factor	-11.42% (-4.37)	0.27 (5.32)	0.43 (5.56)	-0.03 (-0.43)	0.31 5.66	12.0%

In Table 10, the spread portfolio formed by taking long-short positions in the extreme portfolios has an average annual return of -6.09% with a t-statistic of -2.31. It also generates a CAPM alpha of -8.52% with a t-statistic of -3.28, a Fama and French. (1992) alpha of -8.03% with a t-statistic of -3.10, and a Carhart (1997) alpha of -11.42% with a t-statistic of -4.37. Taking a long position in the high VIN-beta and high TVV-beta portfolio and a short position in the low VIN-beta and low TVV-beta portfolio is similar to taking a long position in the high-CIV portfolio and a short position in the low-CIV portfolio, the strategy

taken by Herskovic et al. (2014). Our results are consistent with Herskovic et al. (2014) who show that stocks with higher exposure to idiosyncratic volatility change (CIV) tend to have lower average returns than those with lower exposure. Our findings further suggest that both volatility innovation and time-varying volatility exposures help explain cross-sectional differences in average stock returns. In particular, stocks with higher exposures to VIN and TVV tend to generate lower average and abnormal returns than those with lower exposures to both factors.

Table 11 reports the average and abnormal returns on bivariate portfolio sorts based on market equity (ME) and VIN-beta. We first sort stocks into five portfolios based on ME, then within each portfolio we form five portfolios based on VIN-beta. For each panel, columns are VIN-beta portfolios and rows are ME portfolios. Average excess returns are generally decreasing in VIN-beta and decreasing in market equity (the size effect). In the smallest size quintile, average excess returns are greatest for the two extreme VIN-beta portfolios. In the other four size quintiles, the high-minus-low portfolio exhibit economically large differences in average returns with moderately large t-statistics. Panel B reports the CAPM alphas, which are even larger compared to Panel A. In fact, in the smallest size quintile we now see a small spread in average returns in the same direction as the other size quintiles. Statistical significance is also improved in Panel B compared to Panel A. Panel C presents the results for the Fama and French. (1992) three factor alphas. Similar to Panel B, the spread in average returns are larger than those in Panel A. As a check, we consider an independent sort based on ME and VIN-beta by forming five portfolios on ME and five portfolios on VIN-beta, and then crossing the two sets of portfolios for 25 portfolios. The results are unchanged for the independently sort.

In Table 12, we report the results for portfolios sequentially sorted on ME and TVV-beta. Panel A reports the average excess returns. The long-short portfolios have negative returns for all the quintiles except for the smallest size quintile, with moderately significant t-statistics. Panel B provides the results for the CAPM alpha. Except for the smallest quintile, the high-minus low TVV-beta strategy generates alphas ranging from -2.65% to -3.36% per year, and the t-statistics for all four size quintiles are negative. Panel C reports the Fama and French. (1992) three factor alphas. The abnormal return again is decreasing in TVV-beta except for the smallest ME quintile portfolio. For all other ME quintiles, the high-beta-minus-low-beta strategy has annual FF alphas ranging from -1.62% to -3.49%.

Univariate and bivariate portfolios sorts show that for each of VIN-beta and TVV-

Table 11: **Portfolios Formed on VIN Beta and ME**
We report average excess returns, as well as CAPM and Fama and French. (1992) alphas for portfolios sorted on VIN-beta and market equity (ME) using all CRSP stocks for the 1946-2014 sample. Panel A shows value-weighted average excess returns in sequential two-way sorts on ME and VIN-beta. Panel B shows CAPM alphas. Panel C reports Fama and French. (1992) three factor model alphas. All reported numbers are annualized monthly measures.

VIN beta	1 (Low)	2	3	4	5 (High)	5-1	t(5-1)
Panel A: Average excess returns of Two-way sorts on ME and VIN-beta							
1 (Small)	25.70%	21.13%	21.46%	22.97%	26.08%	0.38%	0.23
2	16.45%	14.81%	13.21%	12.97%	13.49%	-2.96%	-1.92
3	13.31%	12.20%	12.09%	12.14%	10.96%	-2.35%	-1.56
4	12.48%	11.28%	10.82%	10.56%	9.61%	-2.87%	-1.81
5 (Big)	10.99%	9.46%	9.44%	8.45%	9.05%	-1.93%	-1.10
5-1	-14.72%	-11.67%	-12.03%	-14.52%	-17.03%	-	-
t(5-1)	-7.11	-6.66	-6.72	-7.29	-6.97	-	-
Panel B: CAPM α of Two-way sorts on ME and VIN-beta							
1 (Small)	18.11%	14.14%	14.53%	15.55%	16.88%	-1.23%	-0.76
2	8.64%	7.80%	6.33%	5.08%	3.80%	-4.83%	-3.25
3	5.41%	5.07%	5.19%	4.25%	1.17%	-4.25%	-2.92
4	4.82%	4.34%	3.61%	2.72%	0.18%	-4.64%	-3.00
5 (Big)	4.55%	3.21%	2.66%	1.25%	0.51%	-4.04%	-2.37
5-1	-13.56%	-10.93%	-11.87%	-14.29%	-16.37%	-	-
t(5-1)	-6.54	-6.20	-6.56	-7.11	-6.64	-	-
Panel C: FF α of Two-way sorts on ME and VIN-beta							
1 (Small)	15.89%	11.50%	11.95%	12.63%	14.18%	-1.71%	-1.06
2	6.58%	5.22%	3.67%	2.52%	1.74%	-4.83%	-3.24
3	3.23%	2.85%	3.11%	1.95%	-0.10%	-3.32%	-2.32
4	3.04%	2.47%	1.84%	0.81%	-0.45%	-3.50%	-2.30
5 (Big)	3.87%	2.37%	2.20%	1.05%	1.31%	-2.56%	-1.52
5-1	-12.02%	-9.12%	-9.75%	-11.58%	-12.87%	-	-
t(5-1)	-7.50	-7.27	-7.67	-8.26	-6.64	-	-

Table 12: **Portfolios Formed on TVV Beta and ME**
We report average excess returns, as well as CAPM and Fama and French. (1992) alphas for portfolios sorted on TVV-beta and market equity (ME) using all CRSP stocks for the 1946-2014 sample. Panel A shows value-weighted average excess returns in sequential two-way sorts on ME and TVV-beta. Panel B shows CAPM alphas. Panel C reports Fama and French. (1992) three factor model alphas. All reported numbers are annualized monthly measures.

VIN beta	1 (Low)	2	3	4	5 (High)	5-1	t(5-1)
Panel A: Average excess returns of Two-way sorts on ME and TVV-beta							
1 (Small)	24.39%	20.73%	20.89%	24.17%	28.20%	3.81%	2.07
2	16.07%	14.25%	13.39%	13.46%	13.83%	-2.24%	-1.50
3	13.06%	12.58%	11.89%	12.02%	11.19%	-1.88%	-1.27
4	11.72%	11.08%	11.29%	10.87%	9.78%	-1.95%	-1.39
5 (Big)	11.14%	8.96%	9.47%	8.41%	8.20%	-2.93%	-1.96
5-1	-13.25%	-11.78%	-11.42%	-15.76%	-20.00%	-	-
t(5-1)	-6.08	-6.42	-6.15	-7.92	-8.39	-	-
Panel B: CAPM α of Two-way sorts on ME and TVV-beta							
1 (Small)	16.40%	13.92%	13.99%	16.45%	19.44%	3.04%	1.64
2	7.63%	7.30%	6.59%	5.80%	4.45%	-3.18%	-2.14
3	4.48%	5.49%	5.04%	4.35%	1.75%	-2.73%	-1.85
4	3.33%	4.03%	4.22%	3.34%	0.68%	-2.65%	-1.88
5 (Big)	3.67%	2.46%	2.89%	1.41%	0.31%	-3.36%	-2.23
5-1	-12.73%	-11.46%	-11.10%	-15.04%	-19.12%	-	-
t(5-1)	-5.79	-6.19	-5.92	-7.51	-7.97	-	-
Panel C: FF α of Two-way sorts on ME and TVV-beta							
1 (Small)	13.46%	11.28%	11.41%	13.97%	17.00%	3.54%	1.90
2	5.00%	4.81%	4.08%	3.40%	2.59%	-2.41%	-1.61
3	2.30%	3.16%	2.67%	2.24%	0.68%	-1.62%	-1.11
4	1.75%	2.04%	2.45%	1.50%	-0.06%	-1.81%	-1.29
5 (Big)	3.93%	2.13%	2.55%	1.00%	0.45%	-3.49%	-2.33
5-1	-9.53%	-9.15%	-8.86%	-12.96%	-16.55%	-	-
t(5-1)	-5.89	-7.09	-6.42	-8.57	-8.30	-	-

beta, stocks with low exposures on these factors earn economically higher average returns compared to stocks with high exposures. Accounting for the market, size, and value factors of Fama and French. (1992) does not diminish the differences in average returns. We adjust for size exposure through bivariate sorts based on ME and VIN-beta or TVV-beta, and observe average returns differences in along both dimensions. Evidently, size does not subsume the volatility factors. Intuitively, stocks with lower exposures to VIN or TVV are more exposed to volatility innovation risk and time-varying volatility risk, so they must earn higher average returns in order for investors to hold them. Our results support the hypothesis that cross-sectional differences in average stocks returns explained by firm's exposure to idiosyncratic volatility shocks of Herskovic et al. (2014) can be attributed to both exposures to volatility innovations risk and time-varying volatility risk.

5 Conclusion

This paper demonstrates that correlated volatility shocks is an important force that contributes to the factor structure in idiosyncratic volatility. After removing time-volatility effects from idiosyncratic volatilities, the residuals are still positively correlated. Common components in time-varying volatility are not sufficient to explain the idiosyncratic volatility factor structure. We propose a novel volatility model, the DFC model, to capture this empirical fact. Our model fits the data well and outperforms the DCC model of Engle (2002).

Correlated volatility shocks are undiversifiable and are priced in the cross section of equity returns. Univariate and bivariate portfolio sorts exhibit 2-4% difference in annualized average returns between the extreme quintile portfolios. The volatility innovations factor, VIN, appears to capture variation in Fama and French. (1992, 2015) factor volatility, but is not highly correlated with macroeconomic measures. We decompose the CIV factor of Herskovic et al. (2014) into VIN and a time-varying volatility component, TVV, and show both are associated with similar degrees of expected return variation.

While we provide a statistical model to describe the idiosyncratic volatility factor structure, an economic model is required to truly understand this phenomenon. Any model that seeks to delineate economic channels must take into account our finding of correlated volatility shocks. Investigation into such models may prove fruitful for future research. Extending our work to international equities or other asset classes is another interesting direction.

Suitable explanations for U.S. equities may not hold for other asset classes or geographical regions.

References

Ang, A. and Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61(1):259–299.

Bauwens, L., Laurent, S., and Rombouts, V. K. (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics*, 21:79–109.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

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.

Campbell, J. Y., Lettau, M., Malkiel, B. G., and Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance*, 56:1–43.

Cappiello, L., Engle, R., and Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4(4):537–572.

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

Chen, Z. and Petkova, R. (2012). Does idiosyncratic volatility proxy for risk exposure? *Review of Financial Studies*, 25:2745–2787.

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

Duarte, J., Kamara, A., Siegel, S., and 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.

Engle, R. and Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, 30(2):212–228.

Fama, E. F. and French., K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.

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

Herskovic, B., Kelly, B. T., Lustig, H., and Nieuwerburgh., S. V. (2014). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *No. w20076. National Bureau of Economic Research.*

Kalnina, I. and Tewou, K. (2015). Cross-sectional dependence in idiosyncratic volatility.

Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance*, 56(2):649–676.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Sherman, J. and 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.

Tsay, R. S. (2006). Multivariate volatility models. *Time Series and Related Topics, IMS Lecture Notes-Monograph Series*, 52:210–222.

White, H. (1994). Estimation, inference and specification analysis. *New York, NY: Cambridge University Press.*

Appendix A Volatility Innovations of Raw Returns

In this appendix, we derive the covariance of volatility innovations among decile portfolios, assuming raw returns follow a factor model in which both the factors and residual returns have GARCH volatilities. For simplicity, suppose raw returns of security i follow a single-factor model

$$r_{i,t} = f_t\beta_i + e_{i,t}, \tag{40}$$

where f_t is the factor and $e_{i,t}$ is the residual return. Let $\sigma^2_{i,t} = \mathbb{E}[r^2_{i,t}|\mathcal{F}_{i,t-1}]$, $\sigma^2_{f,t} = \mathbb{E}[f^2_t|\mathcal{F}_{i,t-1}]$, $\sigma^2_{e,it} = \mathbb{E}[e^2_{i,t}|\mathcal{F}_{i,t-1}]$, then the squared returns satisfy

$$\begin{aligned} r^2_{i,t} &= f^2_t\beta^2_i + e^2_{i,t} + 2f_t\beta_i e_{i,t}, \\ \sigma^2_{i,t} &= {\beta_i}^2\sigma^2_{f,t} + \sigma^2_{ei,t}, \end{aligned} \tag{41}$$

where $\mathcal{F}_{i,t-1}$ is the previous information set of historical returns and factor values $\mathcal{F}_{i,t-1} = \{r_{i,1}, \cdots, r_{i,t-1}, f_1, \cdots, f_{t-1}\}$ and we use the orthogonality that $\mathbb{E}[f_t e_{i,t}|\mathcal{F}_{i,t-1}] = \mathbb{E}[f_t|\mathcal{F}_{i,t-1}] \cdot \mathbb{E}[e_{i,t}|\mathcal{F}_{i,t-1}] = 0$. Define volatility innovations as follows.

$$\begin{aligned} \text{raw return volatility innovation: } \nu_{i,t} &= r^2_{i,t} - \sigma^2_{i,t}, \\ \text{factor return volatility innovation: } \nu_{ft} &= f^2_t - \sigma^2_{ft}, \\ \text{residual return volatility innovation: } \nu_{e,it} &= e^2_{i,t} - \sigma^2_{ei,t}, \end{aligned} \tag{42}$$

Hence we have the relationship among the volatility innovations of raw, factor, and residual returns.

$$\begin{aligned} \nu_{i,t} &= \beta^2_i\nu_{ft} + \nu_{e,it} + 2\beta_i f_t\epsilon_{i,t}, \\ \text{or } \sigma^2_{i,t}\tilde{\nu}_{i,t} &= \beta^2_i\sigma^2_{f,t}\tilde{\nu}_{ft} + \sigma^2_{ei,t}\tilde{\nu}_{ei,t} + 2\beta_i\sigma_{f,t}\sigma_{e,t}\tilde{f}_t\epsilon_{i,t}, \end{aligned} \tag{43}$$

where $\epsilon_i, t = e_{i,t}/\sigma_{i,t}$, $\tilde{\nu}_{i,t} = \nu_{i,t}/\sigma^2_{i,t}$, $\tilde{\nu}_{ft} = \nu_{ft}/\sigma^2_{ft}$, $\tilde{\nu}_{e,it} = \nu_{ei,t}/\sigma^2_{ei,t}$.
For two securities i and j, covariance betweem their volatility innovations has the form

$$\begin{aligned} \mathbb{E}[\tilde{\nu}_{i,t}\tilde{\nu}_{j,t}] &= 2\beta^2_i\beta^2_j\frac{\mathbb{E}[\sigma^4_{ft}]}{\mathbb{E}[\sigma^2_{i,t}\sigma^2_{j,t}]} + \beta^2_i\frac{\mathbb{E}[\sigma^2_{ft}\sigma^2_{ej,t}]}{\mathbb{E}[\sigma^2_{i,t}\sigma^2_{j,t}]}\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ej,t}] \\ &+ \beta^2_j\frac{\mathbb{E}[\sigma^2_{ft}\sigma^2_{ei,t}]}{\mathbb{E}[\sigma^2_{i,t}\sigma^2_{j,t}]}\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ei,t}] + \frac{\mathbb{E}[\sigma^2_{ei,t}\sigma^2_{ej,t}]}{\mathbb{E}[\sigma^2_{i,t}\sigma^2_{j,t}]}\mathbb{E}[\tilde{\nu}_{ei,t}\tilde{\nu}_{ej,t}] \end{aligned} \tag{44}$$

Assume factor and residuals follow GARCH(1,1) models respectively, particularly

$$\begin{aligned} \sigma_{ft}^2 &= \gamma_{f,0} + \gamma_{f,1} f_{t-1}^2 + \gamma_{f,2}\sigma_{f,t-1}^2, \\ \sigma_{ei,t}^2 &= \gamma_{ei,0} + \gamma_{ei,1} e_{i,t-1}^2 + \gamma_{ei,2}\sigma_{ei,t-1}^2, \end{aligned} \tag{45}$$

It could be derived that

$$\mathbb{E}[\sigma_{ft}^2\sigma_{ei,t}^2] = \frac{\gamma_{f0}\gamma_{i0} + (\gamma_{f1}+\gamma_{f2})\gamma_{i0}\mathbb{E}[\sigma_{f,t-1}^2] + (\gamma_{i1}+\gamma_{i2})\gamma_{f0}\mathbb{E}[\sigma_{ei,t-1}^2]}{1-(\gamma_{f1}+\gamma_{f2})(\gamma_{i1}+\gamma_{i2}) - \gamma_{f1}\gamma_{i1}\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ei,t}]} \tag{46}$$

$$\mathbb{E}[\sigma_{i,t}^2\sigma_{j,t}^2] = \beta_i^2\beta_j^2\mathbb{E}[\sigma_f^4] + \beta_i^2\mathbb{E}[\sigma_{f,t}^2](\mathbb{E}[\sigma_{ei,t}^2] + \mathbb{E}[\sigma_{ej,t}^2]) + \mathbb{E}[\sigma_{ei,t}^2\sigma_{ej,t}^2] \tag{47}$$

$$\mathbb{E}[\sigma_{ei,t}^2\sigma_{ej,t}^2] = \frac{\gamma_{ei,0}\gamma_{ej,0} + (\gamma_{ej,1}+\gamma_{ej,2})\gamma_{ej,0}\mathbb{E}[\sigma_{ei,t-1}^2] + (\gamma_{ei,1}+\gamma_{ei,2})\gamma_{ei,0}\mathbb{E}[\sigma_{ej,t-1}^2]}{1-(\gamma_{ei,2}\gamma_{ej,2}+\gamma_{ei,1}\gamma_{ej,1}) - \gamma_{ei,1}\gamma_{ej,1}\mathbb{E}[\tilde{\nu}_{ei,t-1}\tilde{\nu}_{ej,t-1}]} \tag{48}$$

Therefore, the above results suggest that terms $\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ei,t}]$, $\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ej,t}]$, and $\mathbb{E}[\tilde{\nu}_{ei,t}\tilde{\nu}_{ej,t}]$ drive the cross-covariance. Furthermore, assume $\mathbb{E}[\tilde{\nu}_{ei,t}^2\tilde{\nu}_{ej,t}^2]$ stays the same, then positive values of $\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ei,t}]$ and $\mathbb{E}[\tilde{\nu}_{ft}\tilde{\nu}_{ej,t}]$ would increase $\mathbb{E}[\tilde{\nu}_{i,t}\tilde{\nu}_{j,t}]$. In other words, if volatility innovations of the risk factors in the mean equation are positively correlated with those of the individual securities, using the raw returns without removing risk factors would introduce higher correlations among individuals, and this explains the pattern we observe from Table 1.

Appendix B: Assumptions for Consistency and Assymptotic 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.

Appendix C: Industry Portfolio Classifications

Table A1: **Classifications of Industry Portfolios**
This table provides the 49 industry portfolios that we use in Subsection 3.3.2. These portfolios are classified into 5, 10, and 30 groups respectively based on CRSP SIC code. Industries and groups are defined according to Ken French's website[1].

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				