The Annals of Applied Statistics
2016, Vol. 10, No. 2, 989–1018
DOI: 10.1214/16-AOAS924

ASYMMETRIC CONDITIONAL CORRELATIONS IN STOCK RETURNS

BY HUI JIANG*, PATRICK W. SAART† AND YINGCUN XIA*,‡,[1]

National University of Singapore, Newcastle University† and University of Electronic Science and Technology of China‡*

Modeling and estimation of correlation coefficients is a fundamental step in risk management, especially with the aftermath of the financial crisis in 2008, which challenged the traditional measuring of dependence in the financial market. Because of the serial dependence and small signal-to-noise ratio, patterns of the dependence in the data cannot be easily detected and modeled. This paper introduces a common factor analysis into the conditional correlation coefficients to extract the features of dependence. While statistical properties are thoroughly derived, extensive empirical analysis provides us with common patterns for the conditional correlation coefficients that give new insight into a number of important questions in financial data, especially the asymmetry of cross-correlations and the factors that drive the cross-correlations.

1. Introduction. In the financial world, often financial market participants must manage a large number of financial assets simultaneously. The obvious examples are equity investors who often face risks that affect assets in their portfolio in various ways and must therefore find a position to hedge against these risks. In practice, this is achieved by means of diversification across several stock markets or asset classes, for instance. However, constructing an efficient portfolio to benefit from diversification is not straightforward since it requires knowledge about comovements and associations, that is, correlations, of the assets in question. In addition, such knowledge about the correlations is required in a wide range of financial applications, especially asset pricing models, capital allocation, risk management and option pricing.

The main focus of the research in the current paper is on successful measurement and analysis of the comovements of returns for a portfolio consisting of a large number of assets. In particular, the research in this paper concentrates on a portfolio of thirty major American companies included in the Dow Jones Industrial Average, Dow30 hereafter. Index members for the Dow30 include public companies from various industries and should therefore be able to imitate those

Received May 2015; revised March 2016.

[1]Supported in part by the Ministry of Education of Singapore Grant MOE2014-T2-1-072 and the National Natural Science Foundation of China Grant 71371095.

Key words and phrases. Conditional cross-correlation coefficient, kernel smoothing, reduced rank model, semiparametric models.

of a well-diversified portfolio. To be able to conduct a fruitful analysis, we will develop in this paper a new method that is capable of thoroughly explaining what drives correlations between financial assets and how. The new method, the *reduced rank model for conditional correlation coefficients*, is designed for studying the pairwise conditional correlation structure of financial returns in a functional context of *a semiparametric factor model*. From the empirical point of view, the questions about the driving factor of the observed time-varying correlation structure in financial markets relate directly to a selection of the conditional variable used in the estimation of our semiparametric model. Here, we examine suggestions from two popular schools of thoughts that favor either market volatility or market return. By using measures of the S&P500 as proxy, we are able to establish empirical evidence in support of the well-known asymmetric effect of market return on correlations of the Dow30 returns, that is, a phenomenon whereby correlations of the Dow30 returns are higher during a crisis period than in a stable period. However, such evidence exists only when the possible leverage effect on the market is taken into consideration. Otherwise, the volatility effect of the market return leads to high correlations during the bull market so that the asymmetric effect is not statistically significant. A more detailed discussion about our empirical findings is presented in Section 5.

In the remainder of this section, let us present a brief review on the existing methods and discuss how our method fits among them. Traditionally, correlation was modeled as a constant and unconditional variable. However, over the years practitioners have come to realize that correlation actually varies through time. This motivated a continuously growing amount of research on a wide variety of conditional correlation models. The empirical evidence on the autocorrelation structure of correlation motivated researchers to investigate whether conditional volatility methods based on historical information, as in the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, can be extended for the purposes of modeling conditional correlation. However, new models established as the result of this investigation are either too restrictive that they are unable to explain the roles market variables (e.g. return or volatility) play in driving changes in the behavior of correlations between stock returns or too complex that the number of parameters required explodes with the dimension of the models. An example of models in the first category is the Constant Conditional Correlation GARCH model of Bollerslev (1990). In addition, there are other alternative dynamic conditional correlation GARCH models, which were discussed, for example, in Tse and Tsui (2002), Engle (2002) and Aielli (2013), namely, the VC-GARCH, DCC-GARCH and the cDCC models, respectively.

Even though they were introduced with some general specifications and do not suffer from the curse of the dimensionality problem, these models have quite limited capability. In particular, these models are not able to explain the roles market variables, such as return or volatility, play in driving changes in the behavior of correlations between stock returns, which are of particular interest to financial analysts

[see, for example, Ang and Chen (2002) and Amira, Taamouti and Tsafack (2011)]. As an alternative, Pelletier (2006) proposed a model with a regime-switching correlation structure so that the correlations remain constant in each regime while the change between the states was governed by transition probabilities. Silvennoinen and Teräsvirta (2015) introduced an alternative, which they referred to as the Smooth Transition Conditional Correlation GARCH (STCC-GARCH) model. The model allows the conditional correlations to change smoothly from one state to another as a function of a transition variable and so is associated to some extent with a prespecified model structure on the covariance (e.g., the GARCH-type evolution or regime-switching GARCH model, etc.). This leads to an important limitation which resides in the fact that the number of parameters required explodes with the dimension of the model [Kring et al. (2007) and Santos and Moura (2014)].

Since the ability to model comovements for portfolios with a large number of assets is essential in many areas of financial management, existence of the abovementioned drawbacks suggests that directly modeling the assets by a multivariate GARCH model might not be feasible. Instead, an asset manager must consider some form of factor-model strategies to reduce the overall dimension of the modeling problem. The use of factors to reduce the dimensionality of multivariate GARCH models was proposed in a seminal paper by Engle, Ng and Rothschild (1990), and further developed by Vrontos, Dellaportas and Politis (2003) and Lanne and Saikkonen (2007). More recently, Sheppard and Xu (2014) introduced the so-called Factor-HEAVY (F-HEAVY) model utilizing high frequency data, which has a deep root into the GARCH modeling of conditional volatility. Nonetheless, the purpose of most existing factor-based models, including the F-HEAVY, is to study the way in which covariance matrix changes while these changes are driven by the past information generated by the time series themselves. As for the results, the focus of the studies in multivariate factor GARCH is on predictive models rather than on nonparametric measurement of past volatility and correlations. On the contrary, the semiparametric factor model introduced in this paper enables examination of what and how exogenous forces drive the changes in the correlations of returns. We focus on exploring the asymmetric effect of the exogenous variable on pairwise correlations and identifying the main drivers of the asymmetry in pairwise correlations in a similar spirit to Ang and Chen (2002) and Amira, Taamouti and Tsafack (2011). The importance of the factor approach is to summarize the common patterns in the pairwise correlations. It will soon be clear that the method developed in this paper sits well within the well-known functional data analysis framework, and hence inherits the ability to deal with high-dimensional time series problems. Furthermore, it is based on nonprametric smoothing and is thus model- free, which makes it less likely to suffer modeling misspecification compared to the existing methods.

Our new technique begins with the empirical estimation of the pairwise correlation coefficients of the returns conditional on a particular variable that is of empirical interest, the selection of which is determined by the research problem

under consideration. For the sake of clarity, one can think of the above conditional variable as playing a similar role in our model to the transition variable in the STCC-GARCH model. Since the (pairwise) conditional correlation coefficients are derived based on unknown conditional mean and conditional variance, their estimators must be constructed using empirical estimates. Under the assumption that the conditional correlation coefficient functions share a finite number of common factors, we explore a method of common functional factor analysis along the line of the existing techniques of principal component analysis. To this end, we establish estimators of both the orthogonal functional factors and the corresponding loading coefficients. The theoretical analysis in this paper concentrates on the derivation of consistency and the asymptotic distribution of these estimators that are needed in order to perform statistical inference in the analysis.

The paper is organized as follows. Section 2 discusses the basic construction of our new method, including model assumptions, identification and estimation procedures. Section 3 presents the main asymptotic results of the paper, which focus on the consistency and asymptotic distribution of all the nonparametric estimators involved. These results are convincingly demonstrated by Monte Carlo simulations in Section 4. We then perform empirical analysis in Section 5, while all technical proofs are given in Jiang, Saart and Xia (2016).

2. Conditional correlation coefficients. In the current section and the next, the conditioning variable, denoted by U, plays a similar role in our model to the so-called transition variable in the STCC-GARCH model of Silvennoinen and Teräsvirta (2015). In practice, the choice of U can be made in accordance to the empirical question under investigation. Since the purpose here is to introduce the model in the general context, we will illustrate and discuss this process in more specific detail in Section 5. In this section, we first present the basic construction of our new method, the reduced rank model for conditional correlation coefficients, which includes model assumption and identification. Then we discuss the model's practical operation, which covers the estimation procedures and suggested methods of selecting the number of common factors.

2.1. *Definitions.* In the current paper, we first focus on the study of pairwise conditional correlations. Suppose r_1 and r_2 are returns of two stocks with $E(r_1) = E(r_2) = 0$ so that the unconditional correlation coefficient is defined as

$$\rho_{1,2} = \frac{E(r_1 r_2)}{\sqrt{Er_1^2 Er_2^2}}, \tag{2.1}$$

where the denominator, $E(r_1 r_2)$, measures the comovement of r_1 and r_2. We have, by conditioning upon U,

$$E(r_1 r_2|U) = \mu_1(U)\mu_2(U) + E\{(r_1 - \mu_1(U))(r_2 - \mu_2(U))|U\}, \tag{2.2}$$

where $\mu_k(U) = E(r_k|U), k = 1, 2$. In other words, the comovement between r_1 and r_2 is determined by U based on (i) the effect on the means of r_1 and r_2, and (ii) the effect through the conditional covariance after the effect due to the conditional mean is removed.

Expression (2.2) suggests that we need to consider these two effects separately. After standardization, we may define the correlation due to the effect passing through the conditional means as

$$\phi_{1,2}(U) = \frac{E(r_1|U)E(r_2|U)}{\sqrt{E(r_1^2|U)E(r_2^2|U)}}, \tag{2.3}$$

where $|\phi_{1,2}(U)| \le 1$ due to the Cauchy–Schwarz inequality. The quantity in (2.3) measures the comovement in the conditional mean, and hence it is referred to as the "conditional mean correlation." Similarly, we may define the correlation due to the effect passing through the conditional covariance as

$$\varrho_{1,2}(U) = \frac{E\{(r_1 - \mu_1(U))(r_2 - \mu_2(U))|U\}}{\sqrt{E((r_1 - \mu_1(U))^2|U)E((r_2 - \mu_2(U))^2|U)}}. \tag{2.4}$$

In (2.4), $\varrho_{1,2}(U)$ is the effect of U on the cross-correlation between $r_1 - \mu_1(U)$ and $r_2 - \mu_2(U)$ with the effect on the mean being removed, and is therefore referred to as the "conditional correlation coefficient."

Ang and Chen (2002) introduced a measure of conditional correlation, which was defined as $\mathrm{Corr}(r_1, r_2|c_1 \le U \le c_2)$. However, this definition can cause confusion. In this paper, we discuss the conditional correlation by considering $c_1 \to c_2$, that is, $\mathrm{Corr}(r_1, r_2|U)$. As an example, we consider the capital asset pricing model in financial analysis, which states that

$$r_k = \alpha_k + \beta_k U + e_k, \qquad k = 1, \ldots, m, \tag{2.5}$$

where U is the market return with $\mathrm{Var}(U) = \sigma_U^2$, and

$$E(e_k|U) = 0, \qquad \mathrm{Cov}(e_k, e_\ell|U) = \begin{cases} \sigma_k^2, & \text{if } \ell = k, \\ 0, & \text{otherwise.} \end{cases}$$

We have the unconditional correlation

$$\rho_{k,\ell} = \frac{\beta_k \beta_\ell \sigma_U^2}{(\beta_k^2 \sigma_U^2 + \sigma_k^2)^{1/2}(\beta_\ell^2 \sigma_U^2 + \sigma_\ell^2)^{1/2}},$$

but the conditional correlation

$$\varrho_{k,\ell}(U) = 0.$$

However, if the noises share some common innovations, for example, if

$$e_k = \rho_{k1}(U)\epsilon_1 + \rho_{k2}(U)\epsilon_2, \qquad k = 1, \ldots, m,$$

then

$$\varrho_{k,\ell}(U) = \frac{\rho_{k1}(U)\rho_{\ell 1}(U) + \rho_{k2}(U)\rho_{\ell 2}(U)}{\{\rho_{k1}^2(U) + \rho_{k2}^2(U)\}^{1/2}\{\rho_{\ell 1}^2(U) + \rho_{\ell 2}^2(U)\}^{1/2}}.$$

It is thus important to note that the conditional correlation coefficient defined above is not caused by the common factors in the conditional mean.

2.2. *Model assumption and identification.* Suppose there are m assets to be considered and the return of the kth asset is written as

$$r_k = \mu_k(U) + \sigma_k(U)\varepsilon_k, \qquad k = 1, \ldots, m, \tag{2.6}$$

where $E(\varepsilon_k^2|U) \equiv 1$ almost surely. When U is selected as the market return, it is not difficult to see that the CAPM model described in (2.5) can be taken as a special case. When dealing with sample correlations, it should be taken into account that the return of a given stock should be standardized before being used for estimation of the correlation. Hence, it is useful for the estimation purpose to consider the model

$$\big(r_k - \mu_k(U)\big)^2 = \sigma_k^2(U) + \sigma_k^2(U)\xi_k, \qquad k = 1, \ldots, m,$$

where $\xi_k = (\varepsilon_k^2 - 1)$, as done in Fan and Yao (1998), for example.

For the comovement of $\varepsilon_k, k = 1, \ldots, m$, we assume that the conditional correlation coefficient functions share $p \le m$ common functional factors based on

$$\begin{aligned} E(\varepsilon_k\varepsilon_\ell|U) &\equiv \varrho_{k,\ell}(U) = a_{k\ell} + G_{k\ell}(U) \\ &= a_{k\ell} + b_{k\ell}^{[1]}F_1(U) + \cdots + b_{k\ell}^{[p]}F_p(U), \end{aligned} \tag{2.7}$$

where, as usual, it is assumed that

$$\begin{aligned} &E\big\{F_j(U)\big\} = 0, \\ &E\big\{F_{j_1}(U)F_{j_2}(U)\big\} = 0, \qquad j, j_1, j_2 = 1, \ldots, p, j_1 \ne j_2; \\ &\operatorname{Var}(F_1) \ge \cdots \ge \operatorname{Var}(F_p) \end{aligned} \tag{2.8}$$

for identification purposes. In our analysis, we incorporate uncorrelated measurement errors to reflect additive measurement errors so that the model we consider is

$$\varepsilon_k\varepsilon_\ell = \varrho_{k,\ell}(U) + \epsilon_{k,\ell} = a_{k\ell} + b_{k\ell}^{[1]}F_1(U) + \cdots + b_{k\ell}^{[p]}F_p(U) + \epsilon_{k,\ell}, \tag{2.9}$$

where $\epsilon_{k,\ell}$ are conditionally uncorrelated with each other for all $1 \le k < \ell \le m$, that is, $E\{\epsilon_{k_1,\ell_1}\epsilon_{k_2,\ell_2}|U\} = 0$, if $\{k_1, \ell_1\} \ne \{k_2, \ell_2\}$. With observations at $\{(r_{k,t}, U_t) : t = 1, \ldots, n, k = 1, \ldots, m\}$, where t and k denote respectively the tth time point and the kth asset, our model of interest is thus (2.6) with

$$\varepsilon_{k,t}\varepsilon_{\ell,t} = a_{k\ell} + b_{k\ell}^{[1]}F_1(U_t) + \cdots + b_{k\ell}^{[p]}F_p(U_t) + \epsilon_{k,\ell,t}, \tag{2.10}$$

which we will refer to hereafter as the "reduced rank model."

The reduced rank model differs from existing models in longitudinal data or panel data since here the common factors, $F_1(U), \ldots, F_p(U)$, are unobservable. A similar model was considered in studies on semiparametric comparison of regression curves. A few well-known examples are Härdle and Marron (1990) and Munk and Dette (1998), who studied the comparison of two functions, and James, Hastie and Sugar (2000), who used a similar model but under a random effect setting. In addition, the semiparametric panel data model was also investigated by Boneva, Linton and Vogt (2015). They examined the common component structure of nonparametric functions, however, their dependent variables are observable. Under our model framework, $\varepsilon_{k,t}, k = 1, \ldots, m$, are latent variables and are designed to be estimated nonparametrically based on a GARCH framework. Naturally, the estimation error at the first stage will be inherited, which may increase the difficulty in identifying common factors and estimating corresponding loadings.

In addition, the reduced model differs in a number of ways from existing multivariate GARCH models that also use a factor-based structure. Engle, Ng and Rothschild (1990) generated the covariance structure by introducing a typical factor model for asset excess returns which allows the factors following GARCH processes, but leaves the constant part unrestricted. As a variant of this factor model, Vrontos, Dellaportas and Politis (2003) introduced the full-factor GARCH model by assuming a triangular structure of the parameter matrix. Lanne and Saikkonen (2007) considered a similar model in which some of the diagonal elements of the conditional covariance matrix are constants. For the sake of illustration, let us discuss in more detail one of the most recent and well-known models, namely, the F-HEAVY of Sheppard and Xu (2014). The model resembles the β-GARCH model and relies on the return process

$$r_{k,t} = \beta_{k,t} r_{f,t} + \epsilon_{i,t} = \beta_{k,t}(\sigma_{f,t}\xi_{f,t}) + \sigma_{k,t}\xi_{k,t}, \tag{2.11}$$

where $\sigma_{f,t}$ and $\sigma_{k,t}$ are the conditional variance of the factor and that of $r_{k,t}$ given the factor return, respectively. With this structure, the variance–covariance matrix is

$$\mathbf{H}_t = \begin{pmatrix} \beta_{1,t}^2\sigma_{f,t}^2 + \sigma_{1,t}^2 & \beta_{1,t}\beta_{2,t}\sigma_{f,t}^2 & \cdots & \beta_{1,t}\beta_{N,t}\sigma_{f,t}^2 \\ \beta_{1,t}\beta_{2,t}\sigma_{f,t}^2 & \beta_{2,t}^2\sigma_{f,t}^2 + \sigma_{2,t}^2 & \cdots & \beta_{2,t}\beta_{N,t}\sigma_{f,t}^2 \\ \vdots & \vdots & \cdots & \vdots \\ \beta_{1,t}\beta_{N,t}\sigma_{f,t}^2 & \beta_{N,t}\beta_{2,t}\sigma_{f,t}^2 & \cdots & \beta_{N,t}^2\sigma_{f,t}^2 + \sigma_{2,t}^2 \end{pmatrix}. \tag{2.12}$$

For the sake of comparison, let us take $U_t = \sigma_{f,t}^2$ as the factor (exogenous variable) so that $\mathbf{H}_t$ has a parametric form as shown previously. On the other hand, our model has a more general nonparametric form

$$\mathbf{H}_t = \begin{pmatrix} s_{11}(\sigma_{f,t}^2) & s_{12}(\sigma_{f,t}^2) & \cdots & s_{1N}(\sigma_{f,t}^2) \\ \vdots & \vdots & \cdots & \vdots \\ s_{N1}(\sigma_{f,t}^2) & s_{N2}(\sigma_{f,t}^2) & \cdots & s_{NN}(\sigma_{f,t}^2) \end{pmatrix}. \tag{2.13}$$

While the way by which the conditional covariance evolves with the variable of interest, such as $\sigma^2_{f,t}$, is assumed in F-HEAVY, our method in comparison allows it to be recovered by the data using a nonparametric setting.

In the remainder of this section, we discuss in detail the theoretical construction of our method. To do so, let us denote the vector of individual conditional correlation coefficient functions by $\varrho(u) = (\varrho_{1,2}(u), \ldots, \varrho_{1,m}(u), \varrho_{2,3}(u), \ldots, \varrho_{2,m}(u), \ldots, \varrho_{m-1,m}(u))^\top$. In addition, let $\mathrm{G}(U) = (G_{12}(U), \ldots, G_{1m}(U), G_{23}(U), \ldots, G_{2m}(U), \ldots, G_{m-1,m}(U))^\top$ and $a = (a_{12}, \ldots, a_{1m}, a_{23}, \ldots, a_{2m}, \ldots, a_{m-1,m})^\top$. Then write $\varrho(u) = a + \mathrm{G}(u)$ and $\mathrm{G}(U) = \mathbf{B}\mathrm{F}(U)$,

$$\mathbf{B} = (b_1, \ldots, b_p) \quad \text{and} \quad \mathrm{F}(U) = \big(F_1(U), \ldots, F_p(U)\big)^\top, \tag{2.14}$$

where $b_k = (b_{12}^{[k]}, \ldots, b_{1m}^{[k]}, b_{23}^{[k]}, \ldots, b_{2m}^{[k]}, \ldots, b_{m-1,m}^{[k]})^\top$, $k = 1, \ldots, p$.

With observations at $\{U_t : t = 1, \ldots, n\}$, we define the $m(m-1)/2 \times n$ matrices

$$\mathbf{G} = \big(\mathrm{G}(U_1), \ldots, \mathrm{G}(U_n)\big), \qquad \mathbf{F} = \big(\mathrm{F}(U_1), \ldots, \mathrm{F}(U_n)\big),$$
$$\boldsymbol{\varrho} = \big(\varrho(U_1), \ldots, \varrho(U_n)\big)$$

and write

$$\mathbf{G} = \mathbf{B}\mathbf{F} \quad \text{and} \quad \boldsymbol{\varrho} = a 1_n^\top + \mathbf{G},$$

where 1_n is a column vector of length n with all elements being 1. For ease of exposition, hereafter we let $M = m(m-1)/2$.

From (2.7), since it is reasonable to assume that the information of the pairwise conditional correlation coefficients could be fully captured by the p uncorrelated functional factors, our plan is to apply a similar technique used in principal component analysis to our problem. Let us denote the covariance matrix of $\mathrm{G}(U)$ by

$$\mathbf{\Lambda} = \mathrm{Cov}\big(\mathrm{G}(U)\big) = E\big\{\mathrm{G}(U)\mathrm{G}^\top(U)\big\}. \tag{2.15}$$

An immediate idea is to employ the eigenvalue–eigenvector decomposition. For simplicity, we assume that eigenvalues $\lambda_1, \ldots, \lambda_M$ of $\mathbf{\Lambda}$ satisfy $\lambda_1 > \cdots > \lambda_p > 0$ and $\lambda_{p+1} = \cdots = \lambda_M = 0$ and let $\mathrm{V}_1, \ldots, \mathrm{V}_M$ denote the corresponding orthonormal eigenvectors. Then $\mathbf{\Lambda}$ can be factorized as

$$\mathbf{\Lambda} = \mathbf{V}\mathbf{D}\mathbf{V}^\top = \mathbf{V}_1^*\mathbf{D}^*\mathbf{V}_1^{*\top}, \tag{2.16}$$

where $\mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p, 0, \ldots, 0)$ is a $M \times M$ diagonal matrix, $\mathbf{V} = (\mathrm{V}_1, \ldots, \mathrm{V}_M) = (\mathbf{V}_1^*, \mathbf{V}_2^*)$ is a $M \times M$ matrix, $\mathbf{D}^* = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, $\mathbf{V}_1^* = (\mathrm{V}_1, \ldots, \mathrm{V}_p)$, and $\mathbf{V}_2^* = (\mathrm{V}_{p+1}, \ldots, \mathrm{V}_M)$.

On the one hand, we have the eigenvalue–eigenvector decomposition stated in (2.16). But on the other hand, we indicated previously that $\mathrm{G}(U) = \mathbf{B}\mathrm{F}(U)$ so that

$$\mathbf{\Lambda} = \mathbf{B}E\big\{\mathrm{F}(U)\mathrm{F}^\top(U)\big\}\mathbf{B}^\top. \tag{2.17}$$

In order to proceed, we assume $E\{\mathrm{F}(U)\mathrm{F}^\top(U)\} = \mathbf{D}^*$, which is equivalent to suggesting that $\mathrm{F}(U) = \mathbf{V}_1^{*\top}\mathrm{G}(U)$. Another way of illustrating this point is to consider the matrix $E\{\mathrm{G}(U)\mathrm{F}^\top(U)\}$, which is

$$\begin{aligned} E\{\mathrm{G}(U)\mathrm{G}^\top(U)\mathbf{V}_1^*\} &= E\{\mathbf{B}\mathrm{F}(U)\mathrm{F}^\top(U)\}, \\ \mathbf{\Lambda}\mathbf{V}_1^* &= \mathbf{B}\mathbf{D}^*, \\ \mathbf{V}_1^*\mathbf{D}^* &= \mathbf{B}\mathbf{D}^*, \end{aligned}$$

since $\mathbf{D}^* = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, thus,

$$\mathbf{B} = \mathbf{V}_1^* \quad \text{or} \quad b_j = \mathrm{V}_j. \tag{2.18}$$

This will be essential when we introduce the estimation procedure in the next section.

2.3. *Estimator of conditional correlation coefficients.* Let $\hat{\mu}_k(u)$, $\hat{\mu}_\ell(u)$, $\hat{\sigma}_k^2(u)$ and $\hat{\sigma}_\ell^2(u)$ be local linear estimators of $\mu_k(u)$, $\mu_\ell(u)$, $\sigma_k^2(u)$ and $\sigma_\ell^2(u)$, respectively. Note that ε_k and ε_ℓ are unobservable in practice, but can be estimated by $\hat{\varepsilon}_{k,t} = (r_{k,t} - \hat{\mu}_k(U_t))/\hat{\sigma}_k(U_t)$ and $\hat{\varepsilon}_{\ell,t} = (r_{\ell,t} - \hat{\mu}_\ell(U_t))/\hat{\sigma}_\ell(U_t)$. We can then write

$$\hat{\varepsilon}_{k,t}\hat{\varepsilon}_{\ell,t} = \varrho_{k,\ell}(U_t) + \epsilon_{k,\ell,t} + \hat{\varepsilon}_{k,t}\hat{\varepsilon}_{\ell,t} - \varepsilon_{k,t}\varepsilon_{\ell,t}.$$

By applying the local linear method, an alternative estimator of $\varrho_{k,\ell}(u)$ can be constructed as

$$\hat{\varrho}_{k,\ell}(u) = \frac{\sum_{t=1}^n W_{n,h}(U_t - u)\hat{\varepsilon}_{k,t}\hat{\varepsilon}_{\ell,t}}{\sum_{t=1}^n W_{n,h}(U_t - u)}, \tag{2.19}$$

where $W_{n,h}(U_t - u) = s_{n,h,2}K_h(U_t - u) - s_{n,h,1}K_h(U_t - u)(U_t - u)$, $K(\cdot)$ is a kernel function, $K_h(U_t - u) = K(\frac{U_t - u}{h})/h$, and $s_{n,h,r} = \sum_{t=1}^n K_h(U_t - u)(U_t - u)^r$ for $r = 0, 1, 2$. Moreover, by letting

$$\varrho_{k,\ell}^*(u) = \frac{\sum_{t=1}^n W_{n,h}(U_t - u)\varepsilon_{k,t}\varepsilon_{\ell,t}}{\sum_{t=1}^n W_{n,h}(U_t - u)},$$

then we are able to write

$$\hat{\varrho}_{k,\ell}(u) = \varrho_{k,\ell}^*(u) + \frac{\sum_{t=1}^n W_{n,h}(U_t - u)(\hat{\varepsilon}_{k,t}\hat{\varepsilon}_{\ell,t} - \varepsilon_{k,t}\varepsilon_{\ell,t})}{\sum_{t=1}^n W_{n,h}(U_t - u)}. \tag{2.20}$$

We will present in Section 3 the asymptotic properties of $\hat{\varrho}_{k,\ell}(u)$.

2.4. *Estimators of common functional factors and coefficients.* The basic construction of the model discussed in Section 2.2 suggests that we can make use of the eigenvalue–eigenvector decomposition to estimate the common functional factors and loading coefficients. To do so, we must first obtain an empirical estimate

of the covariance matrix $\mathbf{\Lambda}$, which will take the following approximation in the current paper:

$$\mathbf{\Lambda}_{\mathbf{G}} = \frac{1}{n}\mathbf{G}\mathbf{G}^{\top}.$$

Once the empirical estimate of the conditional correlation coefficients are obtained, then we may estimate $a_{k\ell}$ by

$$\hat{a}_{k\ell} = n^{-1}\sum_{t=1}^{n}\hat{\varrho}_{k,\ell}(U_t).$$

We then estimate each function $G_{k\ell}(u)$ separately by

$$\hat{G}_{k\ell}(u) = \frac{\sum_{t=1}^{n}(\hat{\varepsilon}_{k,t}\hat{\varepsilon}_{\ell,t} - \hat{a}_{k\ell})W_{n,h}(U_t - u)}{\sum_{t=1}^{n} W_{n,h}(U_t - u)}, \tag{2.21}$$

so that we may form $\hat{\mathrm{G}}(U) = (\hat{G}_{12}(U), \ldots, \hat{G}_{1m}(U), \hat{G}_{23}(U), \ldots, \hat{G}_{2m}(U), \ldots, \hat{G}_{m-1,m}(U))^{\top}$. With observations at $\{U_t : t = 1, \ldots, n\}$, the $M \times n$ matrices $\mathbf{G}$ can be estimated by $\hat{\mathbf{G}} = (\hat{\mathrm{G}}(U_1), \ldots, \hat{\mathrm{G}}(U_n))$. Accordingly, an estimate of $\mathbf{\Lambda}_{\mathbf{G}}$ can be constructed as

$$\mathbf{\Lambda}_{\hat{\mathbf{G}}} = \frac{1}{n}\hat{\mathbf{G}}\hat{\mathbf{G}}^{\top}.$$

Second, we obtain the empirical estimates of the eigenvalues $\lambda_1, \ldots, \lambda_M$ and the corresponding orthonormal eigenvectors $\mathrm{V}_1, \ldots, \mathrm{V}_M$ of $\mathbf{\Lambda}$. The asymptotic results presented in Section 3 suggest that we can do so through computing the eigenvalues and the corresponding orthonormal eigenvectors of $\mathbf{\Lambda}_{\hat{\mathbf{G}}}$, which are defined in this paper as $\hat{\lambda}_1, \ldots, \hat{\lambda}_M$ and $\hat{\mathrm{V}}_1, \ldots, \hat{\mathrm{V}}_M$, respectively. Recall that our goal is to obtain

$$\hat{\mathbf{B}} = (\hat{b}_1, \ldots, \hat{b}_p) \quad \text{and} \quad \hat{\mathrm{F}}(U) = \big(\hat{F}_1(U), \ldots, \hat{F}_p(U)\big)^{\top}, \tag{2.22}$$

where $\hat{b}_k = (\hat{b}_{12}^{[k]} \cdots \hat{b}_{1m}^{[k]}, \hat{b}_{23}^{[k]}, \ldots, \hat{b}_{2m}^{[k]}, \ldots, \hat{b}_{m-1,m}^{[k]})^{\top}, k = 1, \ldots, p$. They are the estimates of $\mathbf{B}$ and $\mathrm{F}(U)$ as defined in (2.14), respectively. The first p component functions can be obtained by $\hat{F}_j(u) = \hat{\mathrm{V}}_j^{\top}\hat{\mathrm{G}}(u)$ for $j = 1, \ldots, p$. Finally, based on (2.18) we can directly estimate b_j by $\hat{b}_j = \hat{\mathrm{V}}_j$.

Next, we present the estimators of the common functional factors and loading coefficients under the assumption that there exist a number of common factors $p \le m$ such that $\lambda_1 > \cdots > \lambda_p > 0$, $\lambda_{p+1} = \cdots = \lambda_M = 0$. However, this quantity is unknown in practice. Furthermore, previous experience of functional principal component analysis shows that statistical inference is more difficult for higher-order principal components. Estimation of the new reduced rank model does share a similar difficulty, and so selecting the number of common factors is also an important model selection problem.

To this end, Li, Wang and Carroll (2013) introduced a number of information criteria which are useful in selecting the number of principal components within the context of functional data analysis. In principle, these criteria should also be useful for selecting the common factors in our context. Inspired by Bai and Ng (2002), we consider the following class of information criteria:

$$\mathrm{IC}(p) = \log\big[\hat{\sigma}^2_{[p]}\big] + p g_{M,n}, \tag{2.23}$$

where

$$\hat{\sigma}^2_{[p]} = \frac{1}{nM}\sum_{t=1}^{n}\sum_{k=1}^{m}\sum_{\ell\neq k}^{m}\big(\hat{\varepsilon}_{k,t}\hat{\varepsilon}_{\ell,t} - \hat{a}_{k\ell} - \hat{b}^{[1]}_{k\ell}\hat{F}_1(U_t) - \cdots - \hat{b}^{[p]}_{k\ell}\hat{F}_p(U_t)\big)^2$$

is defined similarly to the estimated variance in Bai and Ng (2002) and

$$g_{M,n} = \left(\frac{M+n}{nM}\right)\log\left(\frac{nM}{M+n}\right)$$

is a penalty function. Finally, we select the number of components as

$$\hat{p} = \min_{p}\mathrm{IC}(p).$$

3. Asymptotics. We first present the asymptotic properties of the estimators for $\varrho_{k,\ell}(u)$. For the estimator $\hat{\varrho}_{k,\ell}(u)$ defined by (2.19), the following asymptotic results are provided.

THEOREM 3.1. *Suppose that the regularity conditions* (C1)–(C6) *in Jiang, Saart and Xia* (2016) *hold. Then for particular* k *and* ℓ, *as* $n \to \infty$, *we have*

$$(nh)^{1/2}\left\{\hat{\varrho}_{k,\ell}(u) - \varrho_{k,\ell}(u) - \frac{1}{2}w_2^K B_{\hat{\varrho}_{k,\ell}}(u)h^2\right\} \to N\big(0,\, f_U^{-1}(u)\omega_{2,k,\ell}(u)\big), \tag{3.1}$$

where

$$B_{\hat{\varrho}}(u) = \varrho''_{k,\ell}(u) - \frac{\varrho_{k,\ell}(u)(\sigma_k^2(u))''}{2\sigma_k^2(u)} - \frac{\varrho_{k,\ell}(u)(\sigma_\ell^2(u))''}{2\sigma_\ell^2(u)},$$

$$\omega_{2,k,\ell}(u) = \nu_K^2\zeta_\epsilon^{k,\ell}(u) + \frac{1}{4}\nu_{K*K}^2\varrho_{k,\ell}^2(u)\zeta_\xi^{k,\ell}(u) - \varrho_{k,\ell}(u)\nu_{K,K*K}\zeta_{\epsilon,\xi}^{k,\ell}(u),$$

with

$$\zeta_\epsilon^{k,\ell}(u) = E\big\{\epsilon_{k,\ell,t}^2|U_t = u\big\}, \qquad \zeta_\xi^{k,\ell}(u) = E\big\{(\xi_{k,t}+\xi_{\ell,t})^2|U_t = u\big\},$$

$$\zeta_{\epsilon,\xi}^{k,\ell}(u) = E\big\{\epsilon_{k,\ell,t}(\xi_{k,t}+\xi_{\ell,t})|U_t = u\big\}.$$

Next, we present asymptotic results for estimators of $\hat{F}_j(u)$ and $\hat{b}^{[j]}_{k\ell}$. Let

$\tilde{\epsilon}_t = (\epsilon_{1,2,t}, \ldots, \epsilon_{1,m,t}, \epsilon_{2,3,t}, \ldots, \epsilon_{2,m,t}, \ldots, \epsilon_{m-1,m,t})^\top$,

$\tilde{\xi}_t = (\xi_{1,t}+\xi_{2,t}, \ldots, \xi_{1,t}+\xi_{m,t}, \xi_{2,t}+\xi_{3,t}, \ldots, \xi_{2,t}+\xi_{m,t}, \ldots, \xi_{m-1,t}+\xi_{m,t})^\top$,

and $\boldsymbol{\epsilon} = (\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_n)$, $\boldsymbol{\xi} = (\tilde{\xi}_1, \ldots, \tilde{\xi}_n)$.

THEOREM 3.2. *Suppose that the eigenvalues of* $\mathbf{\Lambda}$ *satisfy* $\lambda_1 > \cdots > \lambda_p > 0$, $\lambda_{p+1} = \cdots = \lambda_M = 0$. *Let* $\mathbf{I}$ *be the identity matrix of size* M, *and let* $(\lambda_j \mathbf{I} - \mathbf{\Lambda})^+$ *be the Moore–Penrose inverse of* $\lambda_j \mathbf{I} - \mathbf{\Lambda}$. *Under conditions* (C1)–(C6), *as* $n \to \infty$, *for* $j = 1, \ldots, p$,

$$\begin{aligned}\sqrt{n}\Big(\hat{\lambda}_j - \lambda_j - \Big(\frac{1}{2} w_2^K h^2\Big) E\big\{2F_j(U)F_j''(U) - b_j^\top F_j(U)\big(\varrho(U) \circ \sigma(U)\big)\big\}\Big) \\ \xrightarrow{d} N\big(0, \sigma_{\lambda_j}^2\big),\end{aligned} \tag{3.2}$$

where $\circ$ *denotes the Hadamard product of two matrices having the same dimensions*, *and*

$$\begin{aligned}\sigma_{\lambda_j}^2 &= E\big\{I_{j,1}^2\big\} + 2\sum_{s=1}^{\infty} E\{I_{j,1}, I_{j,s+1}\} \\ &= E\big\{F_j^2(U_1) b_j^\top \operatorname{Cov}\big(2\tilde{\epsilon}_1 - \varrho(U_1) \circ \tilde{\xi}_1 | U_1\big) b_j\big\} + E\big\{F_j^4(U_1)\big\} - \lambda_j^2 \\ &\quad + 2\sum_{s=1}^{\infty} E\big\{F_j(U_1) F_j(U_{s+1}) b_j^\top \\ &\quad \times \operatorname{Cov}\big(2\tilde{\epsilon}_1 - \varrho(U_1) \circ \tilde{\xi}_1, 2\tilde{\epsilon}_{s+1} - \varrho(U_{s+1}) \circ \tilde{\xi}_{s+1} | U_1, U_{s+1}\big) b_j\big\} \\ &\quad + 2\sum_{s=1}^{\infty} E\big\{\big(F_j^2(U_1) - \lambda_j\big)\big(F_j^2(U_{s+1}) - \lambda_j\big)\big\},\end{aligned}$$

with

$$I_{j,t} = 2b_j^\top \tilde{\epsilon}_t F_j(U_t) - b_j^\top \big(\varrho(U_t) \circ \tilde{\xi}_t\big) F_j(U_t) + F_j^2(U_t) - E F_j^2(U).$$

Moreover, for the corresponding estimated eigenvectors $\hat{\mathrm{V}}_1, \ldots, \hat{\mathrm{V}}_p$, *under conditions* (C1)–(C6), *as* $n \to \infty$, *for* $j = 1, \ldots, p$,

$$\sqrt{n}\Big(\hat{\mathrm{V}}_j - \mathrm{V}_j - \Big(\frac{1}{2} w_2^K h^2\Big) E\mathrm{W}_{j,1}\Big) \xrightarrow{d} \mathbf{N}_M(0, \mathbf{\Sigma}_{\mathrm{V}_j}), \tag{3.3}$$

where

$$\begin{aligned}E\mathrm{W}_{j,1} = E\Bigg\{(\lambda_j \mathbf{I} - \mathbf{\Lambda})^+ \Bigg[\sum_{i=1}^{p} \mathrm{V}_i F_i(U_t) F_j''(U_t) + \sum_{i=1}^{p} \mathrm{V}_i F_j(U_t) F_i''(U_t) \\ - \frac{1}{2}\sum_{i=1}^{p} \mathrm{V}_i F_i(U_t) \mathrm{V}_j^\top \big(\varrho(U_t) \circ \sigma(U_t)\big) - \frac{1}{2} F_j(U_t)\big(\varrho(U_t) \circ \sigma(U_t)\big)\Bigg]\Bigg\},\end{aligned}$$

$$\mathbf{\Sigma}_{\mathrm{V}_j} = \operatorname{Cov}(\mathrm{H}_{j,1}) + 2\sum_{s=1}^{\infty} \operatorname{Cov}(\mathrm{H}_{j,1}, \mathrm{H}_{j,s+1})$$

$$
\begin{aligned}
&= (\lambda_j \mathbf{I} - \mathbf{\Lambda})^{+} \Bigg[\Bigg(\sum_{i=1}^{p} \mathrm{V}_i \mathrm{V}_j^{\top} F_i(U_1) + F_j(U_1) \mathbf{I} \Bigg) \\
&\quad \times \mathrm{Cov}\Bigg(\tilde{\epsilon}_1 - \frac{1}{2} \varrho(U_1) \circ \tilde{\xi}_1 | U_1 \Bigg) \Bigg(\sum_{i=1}^{p} \mathrm{V}_j \mathrm{V}_i^{\top} F_i(U_1) + F_j(U_1) \mathbf{I} \Bigg) \\
&\quad + 2 \sum_{s=1}^{\infty} \mathrm{Cov}\Bigg(\Bigg(\sum_{i=1}^{p} \mathrm{V}_i \mathrm{V}_j^{\top} F_i(U_1) + F_j(U_1) \mathbf{I} \Bigg) \Bigg(\tilde{\epsilon}_1 - \frac{1}{2} \varrho(U_1) \circ \tilde{\xi}_1 \Bigg), \\
&\quad \Bigg(\sum_{i=1}^{p} \mathrm{V}_i \mathrm{V}_j^{\top} F_i(U_{s+1}) + F_j(U_{s+1}) \mathbf{I} \Bigg) \Bigg(\tilde{\epsilon}_{s+1} - \frac{1}{2} \varrho(U_{s+1}) \circ \tilde{\xi}_{s+1} \Bigg) \Bigg) \\
&\quad + \Bigg(\sum_{i=1}^{p} \mathrm{V}_i F_i(U_1) F_j(U_1) \Bigg) \Bigg(\sum_{i=1}^{p} \mathrm{V}_i^{\top} F_i(U_1) F_j(U_1) \Bigg) \\
&\quad + 2 \sum_{s=1}^{\infty} \mathrm{Cov}\Bigg(\sum_{i=1}^{p} \mathrm{V}_i F_i(U_1) F_j(U_1), \sum_{i=1}^{p} \mathrm{V}_i F_i(U_{s+1}) F_j(U_{s+1}) \Bigg) \Bigg] \\
&\quad \times (\lambda_j \mathbf{I} - \mathbf{\Lambda})^{+},
\end{aligned}
$$

with

$$
\begin{aligned}
\mathrm{W}_{j,t} &= (\lambda_j \mathbf{I} - \mathbf{\Lambda})^{+} \Bigg[\sum_{i=1}^{p} \mathrm{V}_i F_i(U_t) F_j''(U_t) + \sum_{i=1}^{p} \mathrm{V}_i F_j(U_t) F_i''(U_t) \\
&\quad - \frac{1}{2} \sum_{i=1}^{p} \mathrm{V}_i F_i(U_t) \mathrm{V}_j^{\top} \big(\varrho(U_t) \circ \sigma(U_t) \big) - \frac{1}{2} F_j(U_t) \big(\varrho(U_t) \circ \sigma(U_t) \big) \Bigg], \\
\mathrm{H}_{j,t} &= (\lambda_j \mathbf{I} - \mathbf{\Lambda})^{+} \Bigg[\Bigg(\sum_{i=1}^{p} \mathrm{V}_i \mathrm{V}_j^{\top} F_i(U_t) + F_j(U_t) \mathbf{I} \Bigg) \Bigg(\tilde{\epsilon}_t - \frac{1}{2} \varrho(U_t) \circ \tilde{\xi}_t \Bigg) \\
&\quad + \sum_{i=1}^{p} \mathrm{V}_i F_i(U_t) F_j(U_t) \Bigg].
\end{aligned}
$$

Because $b_j = \mathrm{V}_j$ by (2.18), and $b_j = (b_{12}^{[j]}, \ldots, b_{1m}^{[j]}, b_{23}^{[j]}, \ldots, b_{2m}^{[j]}, \ldots, b_{m-1,m}^{[j]})$, the asymptotic results for the estimated coefficients vector $\hat{b}_j$ are equivalent to results for $\hat{\mathrm{V}}_j$. In this case, the following corollary could be obtained directly from the above theorem.

COROLLARY 3.1. *Suppose that all assumptions in Jiang, Saart and Xia* (2016) *are fulfilled. Then for a particular estimated vector* $\hat{b}_j$, *as* $n \to \infty$, *for*

$j = 1, \ldots, p$,

$$\sqrt{n}\left(\hat{b}_j - b_j - \left(\frac{1}{2}w_2^K h^2\right) E\mathrm{W}_{j,1}\right) \xrightarrow{d} \mathbf{N}_M(0, \mathbf{\Sigma}_{\mathrm{V}_j}), \tag{3.4}$$

where $E\mathrm{W}_{j,1}$ *and* $\mathbf{\Sigma}_{\mathrm{V}_j}$ *are the same as which have been given in Theorem* 3.2.

THEOREM 3.3. *Assume that conditions* (C1)–(C6) *in Jiang, Saart and Xia* (2016) *hold, and Section* 2.4 *shows that* $F_j(u) = \mathrm{V}_j^\top \mathrm{G}(u)$, $\hat{F}_j(u) = \hat{\mathrm{V}}_j^\top \hat{\mathrm{G}}(u)$, *as* $n \to \infty$, *we have*

$$\sqrt{nh}\left(\hat{F}_j(u) - F_j(u) - \left(\frac{1}{2}w_2^K h^2\right) EA_1(u)\right) \xrightarrow{d} N\big(0, \sigma_{F_j}^2\big), \tag{3.5}$$

where

$$\begin{aligned} EA_1(u) = {} & \big[F_j''(u) - EF_j''(U)\big] - \frac{1}{2}\mathrm{V}_j^\top\big[\big(\varrho(u) \circ \sigma(u)\big) - E\big(\varrho(U) \circ \sigma(U)\big)\big] \\ & + E\bigg[F_i(U)F_j''(U)\mathrm{V}_i^\top + F_j(U)F_i''(U)\mathrm{V}_i^\top \\ & - \frac{1}{2}\mathrm{V}_j^\top\big(\varrho(U) \circ \sigma(U)\big)F_i(U)\mathrm{V}_i^\top \\ & - \frac{1}{2}\big(\varrho^\top(U) \circ \sigma^\top(U)\big)F_j(U)\bigg](\lambda_j - \mathbf{\Lambda})^+ \mathbf{V}_1^* \mathrm{F}(u), \\ \sigma_{F_j}^2 = {} & \bigg(V_j^\top\bigg[\nu_K^2 \operatorname{Var}(\tilde{\epsilon}_1) + \frac{1}{4}\nu_{K*K}^2\big(\big(\varrho(u)\varrho^\top(u)\big) \circ \operatorname{Var}(\tilde{\xi}_1)\big) \\ & - \nu_{K,K*K} E\big\{\tilde{\epsilon}_1\big(\varrho^\top(u) \circ \tilde{\xi}_1^\top\big)\big\}\bigg]V_j\bigg) / f_U(u). \end{aligned}$$

Finally, we present the asymptotic consistency of $\hat{p}$, which is selected as the minimizer of the above-introduced information criterion, to the true number of common factors. Assume that the true value of p is p_0. For $p \le p_0$, denote $\mathbf{V}_{1,[p]}^* = (\mathrm{V}_1, \ldots, \mathrm{V}_p)$, $\mathbf{V}_{1,[p+1:p_0]}^* = (\mathrm{V}_{p+1}, \ldots, \mathrm{V}_{p_0})$, $\mathbf{D}_{[p]}^* = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$, and $\mathbf{D}_{[p+1:p_0]}^* = \operatorname{diag}(\lambda_{p+1}, \ldots, \lambda_{p_0})$.

THEOREM 3.4. *Let* $\hat{p}$ *be the minimizer of the information criteria defined in* (2.23) *among* $0 \le p \le p_{\max}$ *with* $p_{\max} > p_0$ *being a fixed search limit, and the regularity conditions* (C1)–(C6) *hold. If the penalty function* g_n *satisfies* (i) $g_{m,n} \xrightarrow{P} 0$, (ii) $g_{m,n}/(h^2 + (\frac{\log n}{nh})^{\frac{1}{2}}) \xrightarrow{P} \infty$ *as* $n \to \infty$, *then* $\lim_{n\to\infty} P(\hat{p} = p_0) = 1$.

4. Simulation studies. Although this section focuses mainly on experimental studies that examine the finite sample performance of the proposed framework, a complementary discussion is also presented about the performance comparison between the existing models previously discussed and ours.

4.1. *Monte Carlo studies of the finite sample performance.* The objective of the studies is twofold. First, it is to examine the finite sample performance of (i) the local linear estimator for the conditional comovement of returns, (ii) the newly proposed estimators for the common factors, (iii) the information criterion for selecting the number of the common factors, and (iv) the newly proposed estimators for the common factors. Second, it is to conduct a robustness analysis of the finite sample performance under features which are common in finance. To achieve these objectives, the studies are conducted based on simulated data from a known data-generating process, specifically the return process

$$r_k = a_k + b_k\mu(U) + c_{k0}\epsilon_0 + c_{k1}\epsilon_1 f_1(U) + c_{k2}\epsilon_2 f_2(U), \qquad k = 1, \ldots, m, \tag{4.1}$$

where $a_k, b_k, c_{k0}, c_{k1}, c_{k2}$ are constant coefficients and $\epsilon_0, \epsilon_1, \epsilon_2$ are random renovations with zero mean. For the model in (4.1), it is clear that $E(r_k|U) = a_k + b_k\mu(U)$. For all simulation studies in this section, let us define $\mu(U) = U$ with $U \sim \text{Uniform}(0, 1)$, while the required parameters are generated from independent normal distributions, that is, $a_k, b_k, c_{k0}, c_{k1}, c_{k2} \sim \text{Normal}(0, 0.2)$. In order to demonstrate the robustness of our method, we consider two illustrative scenarios as follows:

Scenario 1: Let $\epsilon_0, \epsilon_1, \epsilon_2 \overset{\text{IID}}{\sim} \text{Normal}(0, 1)$. In addition, let

$$f_1(U) = \sqrt{1 + \cos(v\pi U)} \quad \text{and} \quad f_2(U) = \sqrt{1 + \sin(2\pi U)}.$$

The above specifications suggest that we have

$$\text{Cov}(r_k, r_\ell|U) = \text{Corr}(r_k, r_\ell|U) \equiv C_{k\ell}(U) = \alpha_{k\ell} + \beta_{k\ell}F_1(U) + \gamma_{k\ell}F_2(U), \tag{4.2}$$

where $\alpha_{k\ell} = c_{k0}c_{\ell 0} + c_{k1}c_{\ell 1} + c_{k2}c_{\ell 2}$, $\beta_{k\ell} = c_{k1}c_{\ell 1}$ and $\gamma_{k\ell} = c_{k2}c_{\ell 2}$. In other words, $C_{k\ell}(U)$ involves two common factors defined by

$$F_1(U) = \cos(v\pi U) \quad \text{and} \quad F_2(U) = \sin(2\pi U). \tag{4.3}$$

In the simulation study that follows, we set the value of parameter v in (4.3) as either 2 or 3. Note that the latter introduces a rougher first common factor compared to the former, and hence the resulting conditional correlation functions are less smooth as the results. These functions can be considered as representing structural breaks in the conditional comovements of returns.

Scenario 2: Let $f_1(U)$ and $f_2(U)$ be defined as in Scenario 1, where $v = 2$, but let $\epsilon_0, \epsilon_1, \epsilon_2 \overset{\text{IID}}{\sim} t_\nu$. Such specifications suggest that we have instead the $C_{k\ell}(U)$ with parameters $\alpha_{k\ell} = c_{k0}c_{\ell 0}\sigma_\epsilon^2 + c_{k1}c_{\ell 1}\sigma_\epsilon^2 + c_{k2}c_{\ell 2}\sigma_\epsilon^2$, $\beta_{k\ell} = c_{k1}c_{\ell 1}\sigma_\epsilon^2$ and $\gamma_{k\ell} = c_{k2}c_{\ell 2}\sigma_\epsilon^2$, where $\sigma_\epsilon^2 = \nu/(\nu - 2)$ is the unconditional variance of ϵ_j, for $j = 1, 2, 3$. In the simulation study that follows, we set the parameter ν to 20, 15, 10 or 5. In the probability theory, it is well known that the Student's t-distribution has heavier tails than those of the normal distribution. Hence, from the finance point of view, Scenario 2 simulates return processes with a heavy-tailed behavior. The first three

TABLE 1
Finite Sample Performance of the Estimation Procedure

		$m = 15$			$m = 30$		
v	**ASE**	**100**	**300**	**600**	**100**	**300**	**600**
2	ASE_{F_1}	0.1187	0.0467	0.0321	0.1046	0.0419	0.0279
	ASE_{F_2}	0.1145	0.0592	0.0347	0.1011	0.0496	0.0329
	ASE_C	0.0057	0.0021	0.0009	0.0065	0.0018	0.0008
3	ASE_{F_1}	0.1834	0.0877	0.0552	0.1709	0.0838	0.0529
	ASE_{F_2}	0.1070	0.0603	0.0318	0.1104	0.0472	0.0321
	ASE_C	0.0052	0.0019	0.0009	0.0063	0.0018	0.0009

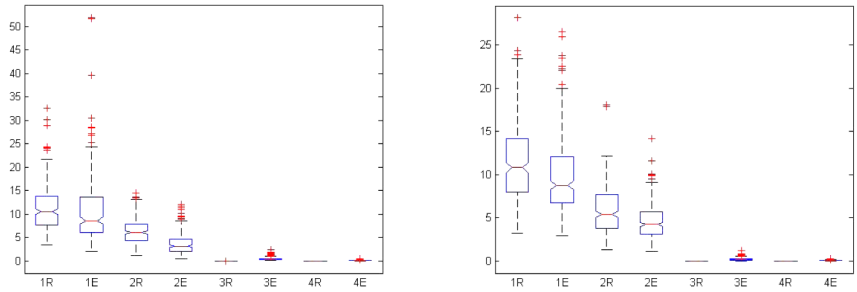
parameter values, namely 20, 15 and 10, reflect the range of values we obtain by fitting the Student's t-distribution with the MLE to the empirically estimated standardized returns of the Dow30, which is denoted in Section 5 by $\hat{\varepsilon}_{k,t}$. To this end, it seems to be the case that multiple estimation and smoothing steps, which are required, lead to confidence intervals that include point-estimates which are relatively close to normality (see Section 5.1 for details). In addition, $\nu = 5$ is included as a benchmark.

We will concentrate first on the simulation work done based on Scenario 1. For the first set of simulation results in Tables 1, 2 and Figure 1, we set the number of

TABLE 2
Finite Sample Performance of the Information Criteria

v	m	n	$\widehat{p} = 0$	$\widehat{p} = 1$	$\widehat{p} = 2$	$\widehat{p} = 3$	$\widehat{p} = 4$
2	15	100	0.2560	0.5960	0.1480	0.0000	0.0000
		300	0.0640	0.3760	0.5600	0.0000	0.0000
		600	0.0160	0.2520	0.7320	0.0000	0.0000
		1000	0.0080	0.0800	0.9120	0.0000	0.0000
	30	100	0.2360	0.5760	0.1680	0.0000	0.0000
		300	0.0440	0.3760	0.5800	0.0000	0.0000
		600	0.0000	0.0000	0.9960	0.0040	0.0000
		1000	0.0000	0.0000	1.0000	0.0000	0.0000
3	15	100	0.2480	0.4520	0.3000	0.0000	0.0000
		300	0.0360	0.1280	0.8360	0.0000	0.0000
		600	0.0120	0.0800	0.9080	0.0000	0.0000
		1000	0.0560	0.0160	0.9280	0.0000	0.0000
	30	100	0.2400	0.4360	0.3240	0.0000	0.0000
		300	0.0200	0.1000	0.8800	0.0000	0.0000
		600	0.0120	0.0680	0.9200	0.0000	0.0000
		1000	0.0000	0.0000	1.0000	0.0000	0.0000

(a) Scenario 1 for $v = 2$ with $n = 100$ and 600 (left and right panels, respectively)



(b) Scenario 1 for $v = 3$ with $n = 100$ and 600 (left and right panels, respectively)

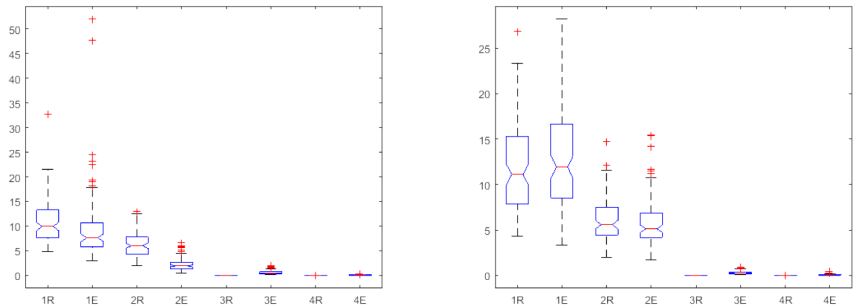


FIG. 1. *Boxplots for eigenvalues calculated based on* $C_{k\ell}(\cdot)$ *and* $\hat{C}_{k\ell}(\cdot)$ *at* $m = 30$.

observations on the time series dimension as $n = 100, 300, 600$ or 1000. We would like to also investigate the importance of the number of assets in the portfolio on the finite sample performance, and therefore set the parameter m as either 15 or 30. The number of simulation replications is 250. We focus first on the finite-sample performance of the local linear estimator for the conditional comovement and the proposed estimators for the common factors. The relevant simulation results are summarized in Table 1 and Figure 1. In the table, the short abbreviation "ASE" stands for the "average squared errors." For $j = 1, 2$,

$$\mathrm{ASE}_{F_j} = \frac{1}{n}\sum_{t=1}^{n}\{\hat{F}_j(U_t) - F_j(U_t)\}^2$$

and

$$\mathrm{ASE}_C = \frac{1}{nM}\sum_{t=1}^{n}\sum_{k=1}^{m}\sum_{\ell\neq k}^{m}\{\hat{C}_{k\ell}(U_t) - C_{k\ell}(U_t)\}^2$$

measure the finite-sample performance of the proposed estimator for the jth common factor and for the estimator of the conditional comovement of returns for any one simulation replication, respectively. For given values of m and n, Table 1

reports the averages of ASE_{F_j} over the simulation replications. In all cases, the estimation errors have a strong tendency to converge to zero as the number of observations increases. An interesting point to make is the fact that increasing the number of assets from $m = 15$ to $m = 30$ is able to slightly improve the overall finite sample performance. In addition, the short abbreviations "R" and "E" (for example, as in "1R" and "1E") in Figure 1 indicate that the eigenvalues are computed based on $C_{k\ell}(\cdot)$ [as defined in (4.2)] and $\hat{C}_{k\ell}(\cdot)$, respectively. Since there are two common factors, that is, $p_0 = 2$, in our model example, 3R and 4R in Figure 1 are appropriately equal to zero. From the figures, it is apparent that the estimation of the eigenvalues performs well, especially since 3E and 4E in the figures are virtually zero across all simulation replications and since the distributions of the estimates tend to follow closely those of the true eigenvalues. Therefore, we have convincing evidence that the proposed esimation procedure for the common factors perform well especially for the number of observations of above 500, i.e. about two years of sample for daily return data.

The important factor contributing to this success is the ability of our method to accurately estimate the conditional comovement of the simulated returns. In Table 1, this is demonstrated by the small magnitude and the tendency of the averaged ASE_C to converge to zero. Let us also point out that specifying the conditional variable U as in Section 4.1 of the current paper contains a special case, which is consistent to taking $\tau = \frac{t}{n} \in (0, 1)$, for $t = 1, \ldots, n$. When such a special case is considered, our experimental design is of a similar nature to that of Engle (2002), which was also used in Aslanidis and Casas (2013), CS hereafter, to illustrate the finite sample performance of the local-linear estimator introduced in their paper. Note that for this special case the CS estimator is merely a simplified version of the local linear estimator introduced in the current paper. On the one hand, this suggests that satisfactory simulation results in Section 4.1 can be interpreted as the ability of our method to nonparametrically model the conditional covariance matrix of returns under misspecification. On the other hand, it also means that the finite-sample superiority of the nonparametric estimator found in CS over the DCC and cDCC models should also hold for the local linear estimator introduced in the current paper.

Our attention is now shifted to the finite sample performance of the above-introduced information criterion for selecting the number of common factors. Note that the error terms, which are required in the calculation, are estimated based on $\varepsilon_k = (r_k - a_k - b_k U)/\sqrt{\sigma_k^2(U)}$, where $\sigma_k^2(U) = c_{k0}^2 + c_{k1}^2 f_1^2(U) + c_{k2}^2 f_2^2(U)$. The empirical distribution of the selected number of components summarized in Table 2 is obtained by setting $p_{\max} = 4$ with $p_0 = 2$, which should be obvious from the specification of (4.2). In Table 2, it is clear that lower numbers of common factors than p_0 are often wrongly selected when $n = 100$. However, the results improve substantially as we increase the number of observations to $n = 300$. Further improvement is made when $n = 600$ and 1000, where the right number of common factors is selected up to 100% of the replications for $m = 30$.

TABLE 3
Finite-Sample Performance with Non-normal Renovations at $m = 30$

	$\nu = 20$		$\nu = 10$		$\nu = 5$	
ASE	**100**	**600**	**100**	**600**	**100**	**600**
ASE_{F_1}	0.1377	0.0283	0.1591	0.0404	0.3558	0.2777
ASE_{F_2}	0.1564	0.0382	0.1985	0.0642	0.4443	0.2719
ASE_C	0.0072	0.0012	0.0116	0.0017	0.0314	0.0054

We will now concentrate on the simulation work done based on Scenario 2. Since the importance of the size of the portfolio has already been examined previously, it is sufficient for our purpose to set the number of observations, n, to either 100 or 600 with $m = 30$. The simulation results are presented in Tables 3 and 4. In Table 3, the effects of the deviation from the normality assumption within the range found in our empirical data, that is, ν is between 20 to 10, seem to be minimal. Significance changes in the results only become apparent by a reduction of the degree of freedom to $\nu = 5$, that is, a level by which data transformation might be required for an application where empirical support for the Student's t-distribution and the degree of freedom can be established. Nonetheless, such negative changes are not apparent in Table 4, which shows the finite sample performance of the information criteria. The information criteria seem to have performed consistently well across the degree of freedom in question.

4.2. *Models performance comparisons*. The objective of this section is to discuss the evidence that our method, which is model-free, is able to provide a more accurate estimation of the covariance matrix as the true data-generating process deviates further away from the prespecified parametric specification. Since it has already been discussed in the previous section, the most convenient experimental example is based on the F-HEAVY framework as specified by (2.11) and (2.12).

TABLE 4
Finite Sample Performance with Non-normal Renovations at $m = 30$

ν	n	$\widehat{p} = 0$	$\widehat{p} = 1$	$\widehat{p} = 2$	$\widehat{p} = 3$	$\widehat{p} = 4$
20	100	0.2600	0.6200	0.1200	0.0000	0.0000
	600	0.0040	0.0080	0.9800	0.0000	0.0000
10	100	0.1050	0.7100	0.1850	0.0000	0.0000
	600	0.0000	0.0150	0.9850	0.0000	0.0000
5	100	0.1950	0.6100	0.1950	0.0000	0.0000
	600	0.0000	0.0750	0.9250	0.0000	0.0000

TABLE 5
Models performance comparisons

	$n = 500$		$n = 1000$		$n = 2000$	
c	**MLE**	**NP**	**MLE**	**NP**	**MLE**	**NP**
0.0	0.0556	0.0932	0.0325	0.0757	0.0272	0.0582
0.1	0.0783	0.1001	0.0575	0.0819	0.0486	0.0626
0.2	0.1142	0.1073	0.0943	0.0885	0.0808	0.0676
0.3	0.1549	0.1149	0.1344	0.0954	0.1169	0.0728
0.4	0.1995	0.1228	0.1783	0.1026	0.1543	0.0783
0.5	0.2464	0.1309	0.2223	0.1099	0.1953	0.0841

In particular, we design the experimental model to check the accuracy as follows:

$$s_{ij}(\sigma_{f,t}^2) = a_i a_j \sigma_{f,t}^2 + c(\sigma_{f,t}^2 - \sigma_0)_+$$

and

$$s_{ii}(\sigma_{f,t}^2) = a_i a_i \sigma_{f,t}^2 + c(\sigma_{f,t}^2 - \sigma_0)_+ + \sigma_i^2,$$

where $x_+ = 0$ if $x < 0$ (or $x_+ = x$ otherwise) and $\sigma_{f,t}^2$ is the conditional variance under the GARCH(1,1) specification, that is,

$$\xi_t = \sigma_{f,t}\epsilon_t, \qquad \sigma_{f,t}^2 = 0.01 + 0.1\xi_{t-1}^2 + 0.89\sigma_{f,t-1}^2$$

with ϵ_t IID $N(0,1)$. We estimate $\sigma_{f,t}^2$ by the maximum likelihood estimation within the model as a F-HEAVY framework, which actually focuses on estimating the parameters a and $\sigma_1^2, \ldots, \sigma_p^2$, while the alternative is to follow our nonparametric estimation by letting $U_t = \sigma_{f,t}^2$. In addition, we calculate the estimation error as

$$\frac{1}{np^2} \sum_{t=1}^{n} \sum_{i=1}^{p} \sum_{j=1}^{p} |s_{ij}(\sigma_{f,t}^2) - \hat{s}_{ij}(\sigma_{f,t}^2)|.$$

Table 5 presents the average of the estimation errors of the maximum likelihood and the nonparametric method, which are denoted respectively by MLE and NP over 100 simulation replications. The results suggested that as c becomes bigger (in other words, as the misspecification becomes more serious), our approach has better accuracy.

5. Effects of market variables on the correlation structure. The empirical study in this section focuses on estimation and analysis of the conditional correlation coefficients for returns of a portfolio of the Dow30 for the observation period between 1 July 1990 to 31 July 2014. Important questions that will be the subject of main interest are what drives the observed time-varying correlation structure of

the Dow30 portfolio and how. In the literature, while there is broad agreement that the correlation structure in financial markets is not constant over time, an outstanding issue of concern is on the driving factor (or factors) behind the observed time variation. Generally, there are two schools of thought which are contradictorily in favor of the market volatility and the market return, respectively. The following paragraphs provide a brief review of these.

A number of previous studies have found that the cross-correlations estimated during volatile periods are significantly larger compared to those computed during calm periods. Using multivariate GARCH models, Longin and Solnik (1995) show that cross-correlations between international markets tended to increase especially in periods of high volatility. Similarly, Ramchand and Susmel (1998) examined the relation between variance and correlation in a conditional time and state-varying framework, and found that the correlations are much higher when the U.S. market is in a high variance state. Furthermore, Chesnay and Jondeau (2001) applied a multivariate Markov-switching model, where the correlation matrix is varied across regimes, to investigate the relationship between international correlation and stock market turbulence, and found that correlation significantly increased during the turbulent periods. In addition, there were also other studies based on the Markov-switching models which have also found that correlation was generally higher in a high-volatility regime [see, e.g., Ang and Bekaert (2002)].

On the other hand, Longin and Solnik (2001) established a pattern of asymmetric dependence using extreme value theory, which implied that international stock markets were more highly correlated during extreme market downturns than during extreme market upturns. Later, Ang and Chen (2002) developed a statistic for testing the asymmetries in conditional correlations based on the exceedance correlation and established evidence in support of Longin and Solnik (2001). Another strand of literature attempted to connect the variability of stock return correlations to the overall economic condition, which was represented by a proxy of market return. Erb, Harvey and Viskanta (1994), for example, suggested that correlations were time varying and dependent on the state of the economy. More importantly, they found a strong tendency for correlation to rise during periods of recession.

It is noteworthy that these schools of thought often consider the market return and volatility as two separate and competing entities. Hence, in order to perform the empirical analysis of interest, we may select the conditional variable, U, as either a measure of the market return or that of the market volatility. However, in the literature it has long been discussed the observed tendency of an asset's volatility to be negatively correlated with the asset's return, that is, what is commonly referred to as the "leverage effect." Furthermore, it has also been documented that the leverage effect is generally asymmetric, that is, other things' equal declines in stock prices are accompanied by larger increases in volatility than the decline in volatility that accompanies rising stock markets. Hence, it is also the main interest of the research in this section to also examine if and how the presence of the leverage effect affects our investigation on the driving factor behind the observed

time variation of stock correlations. For the sake of clarity, we will present first in Section 5.1 relevant methodological details and estimation results, while a thorough discussion on the financial implications and interpretation will be given in Section 5.2.

5.1. *Relevant methodological remarks and estimation results*. Let us begin with the following empirical details: (i) The data used, which consist of the daily close prices (adjusted for dividends and splits) of the Dow30 components and S&P500, and the Chicago Board Options Exchange Market Volatility Index (VIX) between 1 July 1990 to 31 July 2014, are retrieved from Yahoo Finance. (ii) As usual, the closing prices are transformed into returns by taking natural logarithms and differencing. These leads, therefore, to $m = 30$ with $M = 30 \times (30-1)/2 = 435$ conditional correlation coefficients and $n = 6068$ number of observations. (iii) The market volatility is represented in our study by the VIX, which is a popular measure of the implied volatility of S&P 500 index options. (iv) The market return is represented in our study by the return of the S&P500. Furthermore, it is assumed that the return follows an AR(1) + GARCH(1, 1) process. Intuitively, this assumption implies that the leverage effect may influence the market return through both volatility and its persistence that leads to temporally dependence of the market return, that is, autocorrelation. As for the results, the leverage effect for the market can be excluded by first modeling the conditional mean and volatility using the AR(1) + GARCH(1, 1) model, then devolatilizing the raw market return using the resulting conditional variances. Hereafter, let us refer to the resulting process as the *devolatilized* market return such that raw the S&P500 return counterpart is referred to as the *nondevolatilized* market return. (v) We also apply a similar devolatilization to the Dow30 returns.

For the sake of clarity, let us also collect a list of methodological remarks here: (vi) The estimation procedure employed can be summarized as the following steps. Step 1: For a given selection of U (either as the nondevolatilized/devolatilized market return or the VIX for market volatility), the first step in our estimation procedure is to obtain the local linear estimates of $\mu_k(U)$, $\mu_\ell(U)$, $\sigma_k^2(U)$ and $\sigma_\ell^2(U)$. Step 2: These are then used in the calculation of the estimates for the conditional correlation functions, that is, $\hat{\varrho}_{k,\ell}(u)$ in (2.19). Step 3: The asymptotic results in Section 3 suggest that we can calculate $\hat{G}_{k,\ell}(u)$ as $\hat{\varrho}_{k,\ell}(u) - \hat{a}_{k\ell}$, then construct $\hat{\mathbf{G}}$ in order to obtain the covariance matrix $\mathbf{\Lambda}_{\hat{\mathbf{G}}} = \frac{1}{n}\hat{\mathbf{G}}\hat{\mathbf{G}}^\top$. Step 4: We are then able to calculate $\mathbf{V}_{1p}^*$ for each value of $p \le m$ so that the common factor analysis can be conducted based on the IC(p) criterion defined in (2.23). Step 5: Once the number of common factors is selected, we are then able to obtain the empirical estimate of the common factor based on $\hat{F}_1(u) = \hat{V}_1^\top \hat{G}(u)$. The 99% point-wise confidence bands are computed based on the asymptotic variance formula, $\sigma_{F_1}^2$, which was defined in Theorem 3.3. This calculation requires the use of $\hat{V}_1$, which is calculated under the condition $\| V_1 \| = 1$, where $\hat{\varepsilon}_{k,t} = (r_{k,t} - \hat{\mu}_k(U_t))/\hat{\sigma}_k(U_t)$,

$\hat{\varepsilon}_{\ell,t} = (r_{\ell,t} - \hat{\mu}_\ell(U_t))/\hat{\sigma}_\ell(U_t)$ and $\hat{\xi}_{k,t} = \hat{\varepsilon}^2_{k,t} - 1$. Step 6: To compute the nonparametric estimators involved, we choose the normal kernel function given by $K(x) = \frac{1}{\sqrt{2\pi}} e^{-(x^2/2)}$ with $h = std(U)/n^{0.2}$. The above choice of kernel function leads to $\nu_K^2 = 1/2\sqrt{\pi}$, $\nu_{K*K}^2 = 1/2\sqrt{2\pi}$ and $\nu_{K,K*K} = 1/(\sqrt{3}\sqrt{2\pi})$. (vii) The methods and associated results introduced in the current paper are readily applicable to higher frequency financial data. For example, we should be able to employ, as conveniently as in our empirical analysis, the intraday return at the one-minute (or five-minute, ten-minute, etc.) sampling frequency. Nonetheless, it is important to note that the main motivation of the current study is on the identification and estimation of the asymmetry of the overall cross-correlations. This differs significantly from other studies that motivate the use of higher frequency financial data such as Sheppard and Xu (2014). (viii) We have also tried different specifications on the conditional mean and conditional variance equations. However, the functional-based nature of the method and use of the smooth technique mean that they do not bring about significant changes to the results. (ix) We have also attempted to incorporate the asymmetry in the leverage effect into our analysis. This was done by modeling the volatility based on the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). Although the asymmetric effect of market variables was felt more strongly in magnitude, the differences in the results were not statistically significant. (x) Comparing to, for example, the cDCC model, where $O(m^3)$ [alternatively, $O(m^2)$ or smaller] calculations are required for the full likelihood function (for the composite likelihood function) because of the computation of the inverse matrix and constant matrix, our proposed factor approach is based on a nonparametric model structure in which m conditional variance functions are estimated at the first stage, and then $m(m-1)/2$ conditional correlations are estimated nonparametrically. In addition, the eigenvectors of a $m(m-1)/2$ by $m(m-1)/2$ matrix need to be computed to obtain the common functional factors.

The first picture in each panel in Figure 2 displays empirical estimates of 435 correlation functions of the Dow30 components conditioned on a given selection of U, that is, U_{Dv}, U_{Nv} and U_V which denote the devolatilized, nondevolatilized market return and the market volatility, respectively. Although the correlation functions in each of these pictures seem to have their own pattern, overall they tend to share some essential common features. Let us take the first picture of panel (a), which represents the case for U_{Dv}, as an example. In most cases, large negative or positive returns on the S&P500 index imply high correlations, that is, a convex v-shaped conditional correlation function. The common feature is even more apparent in the first picture of panel (c), which represents the case of U_V, where we witness (almost linearly) positive correlation functions with a low degree of variation.

Next we perform the common factor analysis based on the information criterion presented in (2.23). The relevant $\mathrm{IC}(\hat{p})$ values are shown in Table 6. For each of the rows, minimization of these values suggests that a single common

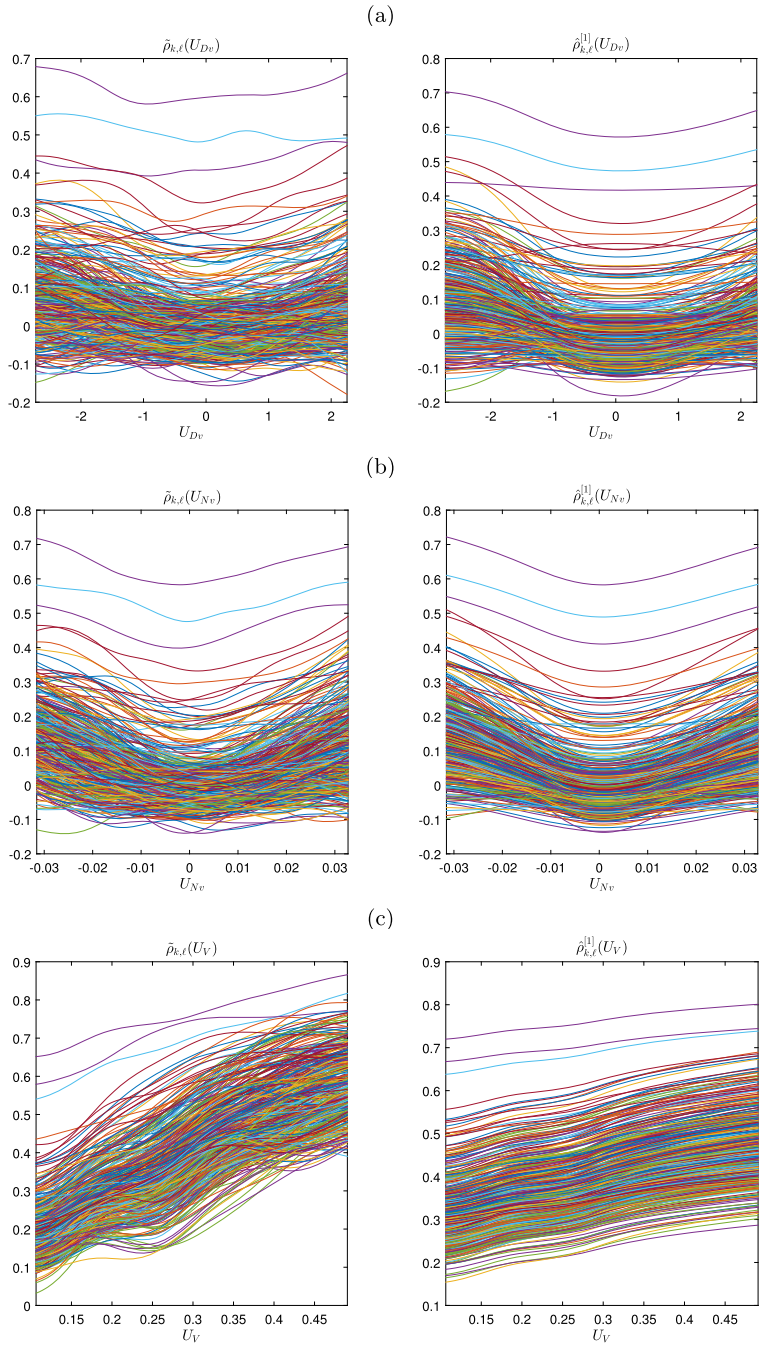


FIG. 2. *Empirical estimates of* $\varrho_{k,\ell}(U)$ *based on* $\hat{\varrho}_{k,\ell}(U)$ *and* $\hat{\varrho}^{[j]}_{k,\ell}(U)$ *for* $j = 1$.

TABLE 6
Information Criterion for Common Factor Analysis

U	IC($\hat{p}=1$)	IC($\hat{p}=2$)	IC($\hat{p}=3$)	IC($\hat{p}=4$)	IC($\hat{p}=5$)	IC($\hat{p}=6$)
U_{Dv}	0.7893	0.8435	0.8981	0.9521	1.0063	1.0598
U_{Nv}	1.0833	1.1370	1.1910	1.2450	1.2988	1.3527
U_V	0.4641	0.5194	0.5730	0.6271	0.6800	0.7350

factor, $p = 1$, should be selected for all cases. The second pictures in panels (a) to (c) of Figure 2 present the empirical estimates of the conditional correlation coefficient functions calculated according to the suggestion made by the information criterion that there exists only one common factor, that is, $\varrho_{k,\ell}(U) = a_{k\ell} + G_{k\ell}(U) = a_{k\ell} + b_{k\ell}^{[1]} F_1(U)$. Hereafter, let us denote these estimates by $\hat{\varrho}_{k,\ell}^{[1]}(U) = \hat{a}_{k\ell} + \hat{b}_{k\ell}^{[1]} \hat{F}_1(U)$, where the upper subscript [1] indicates an involvement of a single common factor. In all cases of U, the graphs seem to provide graphical evidence in support of a single common factor, that is, a conclusion reached due to the fact that the shape of $\hat{\varrho}_{k,\ell}^{[1]}(U)$ closely follows that of $\hat{\varrho}_{k,\ell}(U)$. As for the results, the financial discussion in the next section will focus heavily on $F_1(U)$. For the sake of completion, we present in panels (a), (b) and (c) of Figure 4 empirical estimates of $F_1(U)$ computed based on U_V, U_{Nv} and U_{Dv}, respectively. The 99% point-wise confidence bands were calculated as discussed in Step 4. The red solid curve in each of the figures will be discussed in detail in the next section.

We will now focus on the coefficients $b_{k\ell}^{[j]}$. In a sense, $b_{k\ell}^{[j]}$ should quantify the contribution of the jth common factor on the $k\ell$ conditional correlation function, that is, a role which is usually played by the so-called functional principal component scores in the functional data analysis literature. This is not the case in our model, however, due to the necessity of the assumption $\mathbf{B} = \mathbf{V}_1^*$, which is stated in (2.18). Nonetheless, since a single common factor was selected, the shape of $\hat{\varrho}_{k,\ell}^{[1]}(U)$ depends on $b_{k\ell}^{[1]}$, and so it is important that we perform inferences for $\hat{b}_{k\ell}^{[1]}$. To do so, we calculate the standard errors and, consequently, the 90% confidence intervals of $\hat{b}_{k\ell}^{[1]}$. Figure 3 presents $\hat{b}_{k\ell}^{[1]}$ in ascending order together with the associated 90% confidence intervals for the cases of U_{Dv}, U_{Nv} and U_V [see panels (a), (b) and (c), respectively]. In panel (a), the fact that most of the $\hat{b}_{k\ell}^{[1]}$ presented are positive further suggests that the shape of the common factor is well taken by the (pairwise) conditional correlation functions under consideration. In addition, a similar conclusion can also be obtained in panels (b) and (c) but with stronger statistical significance. Observe, however, that the confidence bands in (c) seem to be smaller than those in panels (a) and (b). This is due mostly to the empirical estimate of $\mathbf{\Sigma}_{V_1}$, which is quite small compared to those for cases of the market returns. Such a result was influenced by $\hat{\varrho}_{k,\ell}(U_V)$, in which we witnessed in Figure 2(c) that they were (almost linearly) positive correlation functions with a

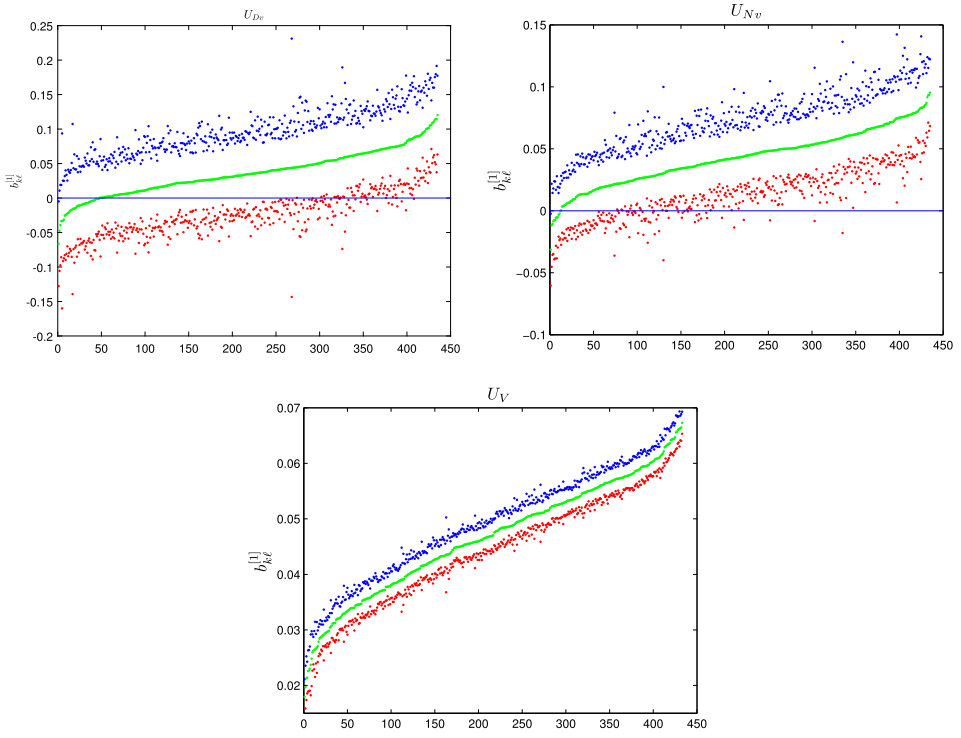


FIG. 3. $\hat{b}_{k\ell}^{[1]}$ *presented in ascending order and the* 90%.

relatively low degree of variation. In this case, higher correlation leads to a larger value of the largest eigenvalue, but also the eigenvector with lower variance. In addition, the first common factor explains up to 97% of the total variations compared to only 70% and 77% in panels (a) and (b), respectively.

5.2. *Financial implications and interpretations*. In this section, we will first discuss important implications of the above results about the effects of the market variables on the correlation structure of the Dow30 portfolio. We will then focus more specifically on the asymmetric effect of the market return.

Let us begin with the kind of effect that market volatility has on the correlations of the returns of the Dow30. Here, the VIX is used as a proxy for the market volatility. The estimation result in Figure 4(a) suggests that correlation significantly increases during volatile periods. This finding is in agreement with the conclusion made by many existing studies (some studies of which are mentioned in the paragraph just above Section 5.1). We consider next the empirical estimates of the common factors presented in Figure 4(b) and (c) which are associated with the nondevolatilized and devoatilized market returns, respectively. In these cases, the first common factor provides a strong evidence against the constant correlation

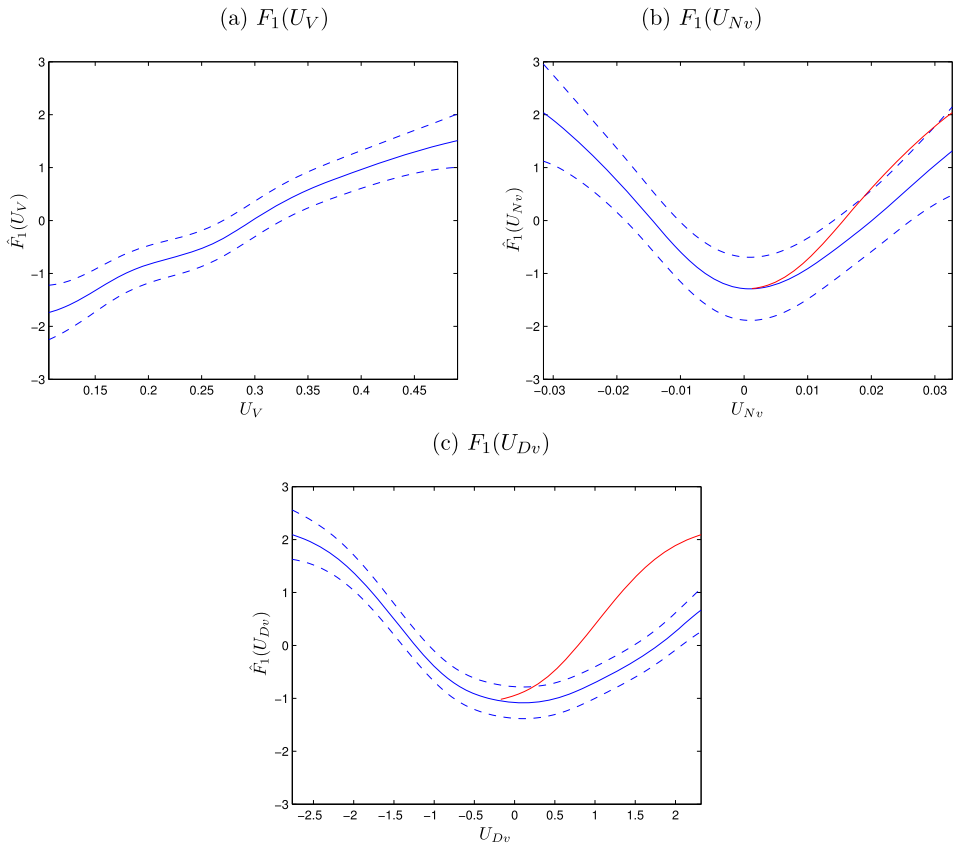


FIG. 4. *Empirical estimate of common factors.*

hypothesis, which was championed by a number of earlier studies [see Kaplanis (1988), e.g.].

An important question often investigated in the literature is whether comovements in the returns are stronger during general market recession than they are during boom periods [see Andersen et al. (2001) and Chesnay and Jondeau (2001), e.g.]. In order to shed some light on this issue, we draw in Figure 4(b) a solid red line, which represents the exact replication of the blue estimate that runs across the negative region of U_{Nv}. The fact that the solid red line lays almost everywhere in between the pairwise confidence bands provides empirical evidence (at least at the 1% significance level) against such an asymmetry. In the next step, we perform a similar analysis to the above, but this time based on U_{Dv}, that is, the devolatilized market return, and the result is reported in Figure 4(c). We find that the correlations decrease quite significantly in the positive region of the market return compared to those presented in Figure 4(b). The fact that the solid red line lays almost everywhere outside the pairwise confidence bands provides empirical evidence in

support of the asymmetric effect of market return on the conditional correlations of the stock returns. Such a finding can be interpreted as follows. Once the leverage effect in the market is disentangled and the volatility effect is removed, correlations of the stock returns drop significantly during the bull while remaining unchanged in the bear market. In effect, the tailing off in the correlations leads to the apparent asymmetric effect of the market return, which is clearly apparent in Figure 4(c).

The above discussion considered two extreme cases, where the conditional variable is either the devolatilized, U_{Dv}, or nondevolatilized, U_{Nv}. For the sake of comparison, we also consider a case by which devolatilization is done based on AR(0) + GARCH(1, 1). This practice reflects the point we have made that the leverage effect does influence market return not only through volatility, but also through volatility persistence, which leads to temporally dependence of the return, that is, autocorrelation. However, we have found the result to be closely similar to that in Figure 4(c), and so it is not reported.

6. Conclusions. In this paper, we examined the comovements of returns for a portfolio which comprised thirty major American companies included in the Dow30. Such a portfolio was of particular interest since it should be able to represent a well-diversified portfolio. To be able to thoroughly investigate the factors that drive correlations between returns of financial assets and how, we introduced in this paper a new semiparametric factor model. We first derived and provided theoretical discussion of an alternative local linear smoothing estimator for the (pairwise) conditional correlation coefficients of asset returns. The new method was then developed along the lines of tools in principal component analysis, which consist of selecting the number and estimation of the common factors together with the corresponding loadings. In the empirical analysis, we followed suggestions provided in a number of existing studies and specified market volatility and market return as the driving factors of the comovements of the Dow30 returns, where the corresponding measures of the S&P500 were used as proxy. We were able to establish empirical evidence in support of the well-known asymmetric effect of market return on the correlations of the Dow30 returns. Specifically, returns correlations were higher during the extreme market downturns compared to those during the extreme market upturns. Nonetheless, this was the case only when the possible leverage effect on the market was taken into consideration. It was apparent that the volatility effect of the market return led to high correlations of the Dow30 returns during the bull market so that the asymmetric effect was not evidenced. Nonetheless, once the leverage effect in the market was disentangled and the volatility effect was removed, the correlations dropped significantly, which then led to the apparent asymmetric effect of the market return.

SUPPLEMENTARY MATERIAL

suskaldyti doi **Supplement to "Asymmetric conditional correlations in stock returns."** (DOI: 10.1214/16-AOAS924SUPP; .pdf). Assumptions and proofs.

REFERENCES

AIELLI, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *J. Bus. Econom. Statist.* **31** 282–299. MR3173682

AMIRA, K., TAAMOUTI, A. and TSAFACK, G. (2011). What drives international equity correlations? Volatility or market direction? *J. Int. Money Financ.* **30** 1234–1263.

ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. and LABYS, P. (2001). The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.* **96** 42–55. MR1952727

ANG, A. and BEKAERT, G. (2002). International asset allocation with regime shifts. *Rev. Financ. Stud.* **15** 1137–1187.

ANG, A. and CHEN, J. (2002). Asymmetric correlations of equity portfolios. *J. Financ. Econ.* **63** 443–494.

ASLANIDIS, N. and CASAS, I. (2013). Noparametric correlation models for portfolio allocation. *J. Bank. Financ.* **37** 2268–2283.

BAI, J. and NG, S. (2002). Determining the number of factors in approximate factor models. *Econometrica* **70** 191–221. MR1926259

BOLLERSLEV, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Rev. Econ. Stat.* **72** 498–505.

BONEVA, L., LINTON, O. and VOGT, M. (2015). A semiparametric model for heterogeneous panel data with fixed effects. *J. Econometrics* **188** 327–345. MR3383213

CHESNAY, F. and JONDEAU, E. (2001). Does correlation between stock returns really increase during turbulent periods? *Economic Notes by Banca Monte dei Paschi di Siena SpA* **30** 53–80.

ENGLE, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.* **20** 339–350. MR1939905

GLOSTEN, L. R., JAGANNATHAN, R. and RUNKLE, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *J. Finance* **48** 1779–1801.

ENGLE, R. F., NG, V. K. and ROTHSCHILD, M. (1990). Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. *J. Econometrics* **45** 213–237.

ERB, C. B., HARVEY, C. R. and VISKANTA, T. E. (1994). Forecasting international equity correlations. *Financ. Anal. J.* **50** 32–45.

FAN, J. and YAO, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika* **85** 645–660. MR1665822

HÄRDLE, W. and MARRON, J. S. (1990). Semiparametric comparison of regression curves. *Ann. Statist.* **18** 63–89. MR1041386

JAMES, G. M., HASTIE, T. J. and SUGAR, C. A. (2000). Principal component models for sparse functional data. *Biometrika* **87** 587–602. MR1789811

JIANG, H., SAART, P. W. and XIA Y. (2016). Supplement to "Asymmetric conditional correlations in stock returns." DOI:10.1214/16-AOAS924SUPP.

KAPLANIS, E. (1988). Stability and forecasting of the comovement measures of international stock market returns. *J. Int. Money Financ.* **7** 63–75.

KRING, S., RACHEV, S. T., HOCHSTOTTER, M. and FABOZZI, F. J. (2007). Composed and factor composed multivariate GARCH models. Technical report, University of Karlsruhe.

LANNE, M. and SAIKKONEN, P. (2007). A multivariate generalized orthogonal factor GARCH model. *J. Bus. Econom. Statist.* **25** 61–75. MR2338871

LI, Y., WANG, N. and CARROLL, R. J. (2013). Selecting the number of principal components in functional data. *J. Amer. Statist. Assoc.* **108** 1284–1294. MR3174708

LONGIN, F. and SOLNIK, B. (1995). Is the correlation in international equity returns constant: 1960–1990? *J. Int. Money Financ.* **14** 3–26.

LONGIN, F. and SOLNIK, B. (2001). Extreme correlations in international equity markets. *J. Finance* **56** 649–676.

MUNK, A. and DETTE, H. (1998). Nonparametric comparison of several regression functions: Exact and asymptotic theory. *Ann. Statist.* **26** 2339–2368. MR1700235

PELLETIER, D. (2006). Regime switching for dynamic correlations. *J. Econometrics* **131** 445–473. MR2276007

RAMCHAND, L. and SUSMEL, R. (1998). Volatility and cross correlation across major stock markets. *J. Empir. Finance* **5** 397–416.

SANTOS, A. A. P. and MOURA, G. V. (2014). Dynamic factor multivariate GARCH model. *Comput. Statist. Data Anal.* **76** 606–617. MR3209460

SHEPPARD, K. and XU, W. (2014). Factor high-frequency based volatility (heavy) models. Working paper, Univ. Oxford.

SILVENNOINEN, A. and TERÄSVIRTA, T. (2015). Modeling conditional correlations of asset returns: A smooth transition approach. *Econometric Rev.* **34** 174–197. MR3268917

TSE, Y. K. and TSUI, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *J. Bus. Econom. Statist.* **20** 351–362. MR1939906

VRONTOS, I. D., DELLAPORTAS, P. and POLITIS, D. N. (2003). A full-factor multivariate GARCH model. *Econom. J.* **6** 312–334. MR2028238

H. JIANG
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
SCIENCE DRIVE, KENT RIDGE
117543, SINGAPORE

P. W. SAART
NEWCASTLE UNIVERSITY BUSINESS SCHOOL
NEWCASTLE UNIVERSITY
5 BARRACK ROAD
NEWCASTLE UPON TYNE, NE1 7RU
UNITED KINGDOM

Y. XIA
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
SCIENCE DRIVE, KENT RIDGE
117543, SINGAPORE
AND
UNIVERSITY OF ELECTRONIC SCIENCE
AND TECHNOLOGY OF CHINA
NO.2006, XIYUAN AVE, WEST HI-TECH ZONE
611731, CHINA
E-MAIL: staxyc@nus.edu.sg