

Quantile Factor-Augmented Prediction Model and Its Applications to Alpha-arbitrage Strategy in China's Stock Market

Siyu Wang, Xiaoxia Wang, Qingqing Zhao and Xiaorong Yang*

Collage of Statistics & Mathematics, Zhejiang Gongshang University, 18 Xuezheng Rd, Hangzhou, China

*Corresponding author

Abstract—Quantitative equity portfolio management has become a fundamental building block of the investment management. The development of general equilibrium asset pricing models enables statistical arbitrage strategies to capture the effect factors of the market returns. In empirical analysis, a crucial step in the model-building process is the selection of an essential factors which may contribute to the positive excess returns. However, it could be challenging since thousands of candidate factors can be obtained. In this study, we employed a factor-augmented model to identify the effect factors for excess returns, and rank the portfolios according the selected factors. A trading strategy of the combination of buying stock portfolio and stock index futures hedging is then perform for Alpha arbitrage.

Keywords—factor-augmented; predictors; quantile regression; arbitrage strategy

I. INTRODUCTION

The launch of stock index futures, margin financing and securities lending accelerate the development of financial derivative instruments. Arbitrage strategies with combination of buying stock portfolio and stock index futures hedging have been received more and more attentions. When referring to the arbitrage strategy, a key point is to select the optimal portfolio. Thus, the multi-factor models are ubiquitous in finance modeling. However, despite their apparent simplicity and widespread use, multi-factor models entail conceptual subtleties that are not immediate to grasp, as there are too many potential factors associated with stock returns can be provided. The classical Capital Asset Pricing Model (CAPM) assumed that the prices of asset portfolio are the reflection of the market information (see [1]). By hypothesis of rational man, the investors will change their investments to maximize the profits. Thus, the arbitrage models were proposed as an alternative to traditional CAPM. A representative work is the three-factor model by Fama and French in 1993 (see [2]). It was pointed out that the empirically determined average-return variables includes size, leverage, earnings/price, and book-to-market equity. Their models were validated by using different market data and were proved to be effective. Recently, a five-factor model directed at capturing the size, value, profitability, and investment patterns in average stock returns were proposed by Fama and French (see [3]). The new model performs better

than the three-factor model in terms of considering variation in average returns related to profitability and investment.

Enlightening by the classical CAPM model, today's quantitative investments use factor models as fundamental building block for trading strategies. Judging by different kinds of factors, investors determine securities to be bought and the time to sell. In financial applications, vast amounts of data could be collected, and realizing positive excess returns in the equity markets is challenging. To achieve the excess returns as higher as possible, we need to successfully design, research, and implement an investment strategy. The capabilities of computerized systems allow us to develop program trading, and thus modeling in portfolio management will be paid more and more attention in the near future.

The essential work for quantitative investments is the prediction of the portfolio returns. Forecasting stock prices has been regarded as one of the most important applications. The factor method is an attractive tool for constructing a proxy of the predictors, as we believe that variations in a set of variables are driven by a small number of latent primitive factors. Reviewing the thousands of the related works in the recent decade, we refer to the literature multi-factor models in [4-7].

In this paper, we propose to use the quantile factor-augmented model to study the effect of the potential predictors on the stock returns. We separate the factors into two groups, one consists of very few predetermined factors and the other may have large number of potential factors. By an information-theoretic criterion, some synthesized factors can be derived as the linear combination of the original factors, which well represent the common influence of each stock as well as the solution of dimension reduction. Stocks with higher predicted returns will be selected as the elements of the portfolio. Then the Alpha arbitrage strategy will be executed by consideration of buying stock portfolio and selling the stock index futures as hedge. Our empirical study shows that based on the proposed method, it work well in China's A-share market. Through 2007 to 2017, the average annualized return is 0.149, the Sharpe ratio is 2.681, and the max drawdown is only 0.069.

The rest of the paper is organized as follows. In section 2, we describe the quantile factor-augmented prediction model, and give the estimation and model selection criterion. In

section 3 is the empirical study of China's stock market. By applying the proposed method, we determine the latent primitive factors for each stock, and predict its returns. Then according to the predicted results, we perform the Alpha arbitrage strategy by buying stock portfolio and selling the stock index futures. The results indicate that our method works well in China's A-share market. Conclusions and discussions are made in section 4.

II. THE QUANTILE FACTOR-AUGMENTED PREDICTION MODEL AND STATISTICAL INFERENCES

A. Model and Assumptions

Consider the following model:

$$y_{it} = \boldsymbol{\alpha}'_i \mathbf{x}_{it} + \boldsymbol{\beta}'_i \boldsymbol{\omega}_{it} + \varepsilon_{it}, \quad (1)$$

where y_{it} is the i -th response (in our empirical study, y_{it} presents the return of stock i) at time t , $\mathbf{x}_{it} = (x_{i,1t}, x_{i,2t}, \dots, x_{i,Nt})'$ is the possible predictor vector with large dimension N , $\boldsymbol{\omega}_{it} = (\omega_{i,1t}, \omega_{i,2t}, \dots, \omega_{i,pt})'$ is another predetermined predictor vector with small dimension p , $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ are N - and p -dimensional coefficient parameters, and ε_{it} is the noise term. The above model depicts the response as a combination of few common important predictors ($\boldsymbol{\omega}_i$) and some featured predictors (\mathbf{x}_i). The advantage of model (1) is obvious that the more candidate factors can be involved. And clearly, all the factors are classified into two groups; one is predetermined, and the other is latent but can be estimated by some efficient methods. However, as we know, in regression setting, to avoid the over-fitting problem, some restrictions must be made on the number of the predictors. However, x_t usually contains much information about y_t , some dimension reduction techniques are needed to be involved in. We assume \mathbf{x}_i is generated by the following factor model:

$$\mathbf{x}_{it} = \boldsymbol{\lambda}_i \mathbf{f}_{it} + e_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, n. \quad (2)$$

where $\mathbf{f}_i = (f_{i1}, f_{i2}, \dots, f_{ir})'$ is the r -dimensional vector of latent factors. Equation (2) indicates that the potential predictors can be treated as the linear combination of a small set of latent variables, thus the difficulty of curse of dimensionality might be solved. By (1) and (2), we have

$$y_{it} = \boldsymbol{\alpha}'_i \mathbf{f}_{it} + \boldsymbol{\beta}'_i \boldsymbol{\omega}_{it} + \varepsilon_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, n. \quad (3)$$

In real data analysis, large number of predictors often leads to high dependence among the predictors and approximate multicollinearity. Moreover, if there are a few observations but too many predictors, dimension reduction techniques are required to adequately capture the main feature of the data

structure. If given \mathbf{x}_{it} and $\boldsymbol{\omega}_{it}$, the τ -th conditional quantile of y_{it} can be formulated as following:

$$Q_\tau(y_{it} | \mathbf{x}_{it}, \boldsymbol{\omega}_{it}) = \boldsymbol{\alpha}_i(\tau)' \mathbf{f}_{it} + \boldsymbol{\beta}_i(\tau)' \boldsymbol{\omega}_{it} =: \boldsymbol{\gamma}_i(\tau)' \mathbf{z}_{it}, \quad (4)$$

$$(t = 1, \dots, T, \quad i = 1, \dots, n.)$$

where $\boldsymbol{\gamma}_i(\tau)$ is the coefficient vector and $\mathbf{z}_{it} = (\mathbf{f}_{it}, \boldsymbol{\omega}_{it})'$. Thus, if we could appropriately estimate $\boldsymbol{\gamma}_i(\tau)$ and \mathbf{z}_{it} , any interested conditional quantile of y_{it} can be predicted.

B. The Two-stage Estimation

Our ultimate goal is to estimate the conditional quantile of y_{it} . Before we introduce the estimation procedure, we have to point out that although the data can be viewed as panel data, we treat it as the cross-section data. When we do estimation, we fix the time at time t . The estimation varies when t changes. We consider a statistical modeling methodology. It first estimates the common factors from the data by the principal component analysis and then applies the estimated factors to standard quantile regression. The inference procedure will be carried out by two stages. In the first stage, factor analysis is performed by asymptotic principal component analysis, and some common factors are found. Denote $\hat{\mathbf{f}}_{it}$ as the estimated common factor, and substitute it into (4), then we have

$$Q_\tau(y_{it} | \hat{\mathbf{x}}_{it}, \boldsymbol{\omega}_{it}) = \boldsymbol{\alpha}_i(\tau)' \hat{\mathbf{f}}_{it} + \boldsymbol{\beta}_i(\tau)' \boldsymbol{\omega}_{it} =: \boldsymbol{\gamma}_i(\tau)' \hat{\mathbf{z}}_{it} \quad (5)$$

$$(t = 1, \dots, T, \quad i = 1, \dots, n.)$$

In the second stage, we estimate $\hat{\boldsymbol{\gamma}}_i(\tau)$ by maximum likelihood method, i.e.

$$\hat{\boldsymbol{\gamma}}_i(\tau) = \underset{\boldsymbol{\gamma}_i(\tau)}{\operatorname{argmax}} L_\tau(y_{it}; \boldsymbol{\gamma}_i(\tau), \hat{\mathbf{z}}_{it})$$

$$= \underset{\boldsymbol{\gamma}_i(\tau)}{\operatorname{argmax}} \frac{1}{n} \log \left[\tau^T (1-\tau)^T \exp \left\{ - \sum_{i=1}^n \rho_\tau(y_{it} - \boldsymbol{\gamma}'_i(\tau) \hat{\mathbf{z}}_{it}) \right\} \right],$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ is the check loss function defined as in [8]. Then the estimated quantile regression model with factor-augmented predictor is

$$\hat{Q}_\tau(y_{it} | \hat{\mathbf{x}}_{it}, \boldsymbol{\omega}_{it}) = \hat{\boldsymbol{\gamma}}_i(\tau)' \hat{\mathbf{z}}_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, n.$$

C. Model Selection Criterion

Once the model is set up, a critical step is to determine the number of common factors in the first estimation step. We adopt the information criterion by [7]. Suppose $\mathbf{y}_n = (y_1, \dots, y_n)'$ is generated by an unknown true

distribution $G(\mathbf{y})$, whose density function is $g(\mathbf{y})$. Let $\mathbf{u}_n = (u_1, \dots, u_n)'$ be a replicates vector of \mathbf{y}_n drawn from $g(\mathbf{u})$. Define

$$\eta_r(G; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) := \int L_r(\mathbf{u}_n; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) dG(\mathbf{u}_n), \quad (6)$$

where $dG(\mathbf{u}_n)$ is the Lebesgue measure with respect to the true probability density $g(\mathbf{u})$. The best model is chosen by maximizing the expected log-likelihood function in (6) among different candidate models. Here the optimal number of factor is defined as the value r that maximizes the expected log-likelihood (see [7] for details). The sample based on log-likelihood estimation is given by

$$\begin{aligned} \eta_r(\hat{G}; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) &:= \int L_r(\mathbf{u}_n; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) d\hat{G}(\mathbf{y}_n) \\ &= \log[\tau(1-\tau)] - \frac{1}{n} \sum_{i=1}^n \rho_\tau(y_{it} - \hat{\mathbf{y}}_t'(\tau)\mathbf{z}_{it}'), \end{aligned} \quad (7)$$

where $\hat{G}(\cdot)$ is the empirical distribution of $G(\cdot)$. Then the bias of the log-likelihood is

$$b_r(G) := \int [\eta_r(\hat{G}; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) - \eta_r(G; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t)] dG(\mathbf{y}_n). \quad (8)$$

If $b_r(G)$ can be approximated appropriately, an analogous form of information criteria in [9] can be defined as

$$IC = -2l_r(y_T; \hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) + 2n\hat{b}_r(G),$$

which gives a model selection criterion for evaluating the estimated mode, and the number of common factors r can also be determined. The estimation procedure for $\hat{b}_r(G)$ is quite complicated and should follow Theorem 3.1 in [7].

III. THE EMPIRICAL STUDY OF CHINA'S STOCK MARKET

A. Data and Their Statistical Properties

In this section, we select some sample stocks of Shanghai 180 index (SH180) to illustrate our methods. The special treatment (ST) stocks and some other shares with financial abnormality are excluded, and finally, there are 49 stocks in the following analysis. The data are from Jan 2, 2007 to Dec 29, 2017. We use the logarithm of close price to obtain the return series, i.e. $y_{it} = \ln P_{it} - \ln P_{i,t-1}$, with P_t denoting the close price of the stock at time t . Then we construct the following model:

$$\begin{cases} \mathbf{x}_{it} = \boldsymbol{\lambda}_r \mathbf{f}_{it} + e_t \\ Q_\tau(y_{it} | \mathbf{x}_{it}, \boldsymbol{\omega}_{it}) = \beta_{0,\tau}(\tau) + \boldsymbol{\beta}'_\tau(\tau)\boldsymbol{\omega}_{it} + \boldsymbol{\alpha}_\tau^x(\tau)\mathbf{f}_{it} \end{cases},$$

where $\boldsymbol{\omega}_{it} = (mv_{it}, bm_{it}, beta_{it}, mon_{it}, rev_{it})'$ is the predetermined variables with five elements considered: mv_{it} is the size of firm i , bm_{it} is the book-to-market value, $beta_{it}$ is the system risk measured by Scholes and Williams in 1977 (see [10]), mon_{it} is the momentum effect, $mon_{it} = \sum_{j=2}^{12} y_{it-j}$, and rev_{it} is the reversal effect, $rev_{it} = \sum_{j=1}^3 y_{it-j}$. $\mathbf{x}_{it} = (x_{1,it}, x_{2,it}, \dots, x_{36,it})$ is the potential factors that might influence the returns. Here \mathbf{x}_{it} contains 36 elements, which fall into 6 categories: the enterprise's financial data (C1), China's stock market composite index (C2), macroeconomic influences (C3), commodity prices (C4), exchange rate (C5), international market composite index (C6). Please see Table I for details.

B. Determination of Common Factors

Let $\tau = 0.05, 0.1, 0.5, 0.9, 0.95$, and we use the information criterion described in section 2.3 to determine the number of common factors. Table II shows the results.

From Table II we know that: (1) the optimal number of factors r depends on the quantile levels; (2) at lower (0.05) or higher (0.95) quantiles, lots of the series can be fitted well (R^2 is large), which will provide explanations of the extremely low or high returns; (3) if the number of factors $r < 3$, almost all R^2 are small.

For space limitation, we could not list the results for all the time, therefore, we just use $t = 1$ (Mar, 2007) for instance. For $t = 1$, four common factors are selected by our proposed method. According to the factor loading matrix, they are

$$\begin{cases} f_1 = -3.42x_{11} - 3.47x_{12} - 3.34x_{13} \\ f_2 = 2.26x_{14} + 4.92x_{15} \\ f_3 = 1.14x_1 + 2.11x_2 + 1.21x_3 - 1.92x_4 - 2.72x_6 \\ f_4 = 1.25x_5 + 2.16x_7 - 1.92x_{26} \end{cases}$$

The first factor is the linear combination of three elements in terms of Economic Sentiment Index, which considers the macroeconomic influences. The second factor considers the inflation and industrial price, which represents the consumption. The third factor concerns the fundamentals of stock. It is natural that good financial health will bring higher returns. And the fourth factor is associated with domestic and international markets, which provides reasonable information.

C. The Alpha-arbitrage Strategy

The selected 49 stocks in SH180 include gas, steel, electronics, transportation products, agriculture, insurance, securities, manufacture, food, textile, apparel and other financial service, they have a good representation of China's stock market. According to Table I, for higher returns ($\tau = 0.9, 0.95$), we propose to use fewer synthesize factors for further analysis. For different time t , the components of the synthesize factors may not be the same. However, once we

have the four synthesized factors, we rank each synthesized factor and the predetermined factor (decreasing or increasing order will depend on the sign of the coefficients). Then we score the stocks by using the average rank of the synthesized and predetermined factors. The score will be finally used in the decision of the portfolio for arbitrage strategy.

In Alpha-arbitrage strategy designing, the research object will be all the stocks (special treatment stocks are excluded) in China's A-share market (through Jan 2, 2007 to Dec 29, 2017). By using the proposed method, we give a score to each stock by ranking the synthesized factors and predetermined factors associated with the predicted returns \hat{y}_t at time t . Obviously, at different time t , the ranking results are different. We will buy a stock portfolio which contains 20 stocks whose scores are in the top 20, and meanwhile, a short selling of stock index futures hedging with same amount of money is performed. The trading rule is that, we use equal money to buy 20 stocks, and will change the portfolio every three months by updating the new data and new predictors (as the financial data usually can only be reported quarterly). Then and if the returns of the portfolio are higher than that of the market, we could have arbitrage. The trading commission is set as 1%. We report some frequently used statistics for arbitrage strategy in Table III (AR: Annualized Return; CR: Cumulative Return; SPR: Sharpe Ratio; SOR: Sortino Ratio; MDD: Max Drawdown).

TABLE I. LIST OF THE POTENTIAL FACTORS ("*" MEANS THE DATA IS TAKEN LOGARITHM).

	factor name	category		factor name	category
x_1	return on equity	C1	x_{19}	growth rate of aluminum price*	C4
x_2	earnings per share	C1	x_{20}	growth rate of lead price*	C4
x_3	price to book value ratio	C1	x_{21}	growth rate of zinc price*	C4
x_4	price earning ratio	C1	x_{22}	growth rate of tin price*	C4
x_5	turnover	C1	x_{23}	growth rate of nickel price*	C4
x_6	price change ratio	C1	x_{24}	growth rate of natural gas price*	C4
x_7	return of Shanghai composite index*	C2	x_{25}	growth rate of WIT crude oil price*	C4
x_8	return of Shenzhen composite index*	C2	x_{26}	growth rate of exchange rate: USD to RMB*	C5
x_9	growth rate of broad money	C3	x_{27}	growth rate of exchange rate: EUR to RMB*	C5
x_{10}	growth rate of narrow money	C3	x_{28}	growth rate of exchange rate: HKD to RMB*	C5
x_{11}	coincidence index	C3	x_{29}	growth rate of exchange rate: 100YEN to RMB*	C5
x_{12}	leading index	C3	x_{30}	growth rate of exchange rate:	C5

				Pound to RMB*	
x_{13}	lagging index	C3	x_{31}	return of Dow Jones industrial average*	C6
x_{14}	inflation rate	C3	x_{32}	return of NASDAQ*	C6
x_{15}	industrial production price index	C3	x_{33}	return of S&P 500*	C6
x_{16}	growth rate of gold price*	C4	x_{34}	return of FTSE 100*	C6
x_{17}	growth rate of silver price	C4	x_{35}	return of HSI	C6
x_{18}	growth rate of copper price	C4	x_{36}	return of Nikkei 225	C6

From Table III, we know that the strategy works well when selecting the portfolio by the proposed factor-augmented prediction method. Through 2007 to 2017, the average annualized return is high and the max drawdown is only 0.069. Also, that Sharpe ratio is 2.681 which allows risk manager to have diversification benefits by adjusting the allocation of his investment.

IV. CONCLUSION AND DISCUSSION

Quantitative equity investing has been attracted a lot. As a fundamental tool, multi-factor model is frequently used to asset pricing problem. In financial markets, with the developments of data compilation and computation technique, much more information can be derived. When depicting the stock returns, a large amount of potential predictors can be found, and it will cause the model selection problem. Under the framework of quantile regression, we propose to use the quantile factor-augmented prediction model to forecast the returns. Essentially, the candidate predictors are classified into two subgroups, one contains very few predetermined elements, and the other allows much more potential factors. The dimension reduction will benefit from information-theoretic criterion by [7]. The empirical study shows that our method works very well for Alpha arbitrage strategy.

Due to the limitation of space, we could not list all the empirical study results. However, two more aspects are deserved to discuss. First of all, we need to mention that in our study, only cross-section data is considered. We estimate the primitive factors at different time t as we believe that the primitive factors may be allowed to change since there are too many potential factors exist. The data can also be treated as a panel data. Under the framework of panel data, it is a totally different story. Secondly, as quantile regression is considered in our paper, it is easy to apply our method to financial risk management. From a statistical view, we could calculate the Value-at-Risk (VaR), as VaR is estimated by either lower (1% or 5%) or upper (99% or 95%) percentile of the returns. Once the predetermined factors given, by proposed method, some comprehensive factors are generated, then we can estimate the

TABLE II. THE NUMBER OF COMMON FACTORS FOR DIFFERENT TIME (MONTHLY DATA ARE USED)

Time	$\tau = 0.05$		$\tau = 0.1$		$\tau = 0.5$		$\tau = 0.9$		$\tau = 0.95$	
	r	R^2	r	R^2	r	R^2	r	R^2	r	R^2
Mar, 2007	4	0.89	4	0.83	4	0.74	4	0.8	4	0.87
Jun, 2007	4	0.86	6	0.81	4	0.68	4	0.75	4	0.83
Sep, 2007	4	0.88	2	0.81	2	0.77	3	0.77	3	0.78
Dec, 2007	6	0.81	6	0.78	5	0.81	4	0.79	4	0.87
Mar, 2008	2	0.64	5	0.73	5	0.63	5	0.71	5	0.79
Jun, 2008	4	0.57	2	0.67	6	0.63	6	0.73	6	0.83
Sep, 2008	2	0.74	2	0.69	4	0.63	4	0.72	4	0.81
Dec, 2008	2	0.69	4	0.59	4	0.67	4	0.76	4	0.81
Mar, 2009	5	0.77	5	0.72	5	0.68	5	0.75	5	0.84
Jun, 2009	3	0.68	3	0.66	5	0.58	5	0.69	5	0.77
Sep, 2009	4	0.66	5	0.66	4	0.82	4	0.66	4	0.76
Dec, 2009	4	0.55	4	0.60	5	0.59	5	0.67	5	0.77
Mar, 2010	3	0.65	3	0.53	4	0.54	4	0.66	4	0.77
Jun, 2010	5	0.57	6	0.71	5	0.69	5	0.76	5	0.82
Sep, 2010	5	0.69	5	0.64	5	0.67	5	0.75	4	0.78
Dec, 2010	4	0.74	4	0.68	4	0.56	4	0.59	4	0.72
Mar, 2011	3	0.62	4	0.55	4	0.65	4	0.71	4	0.79
Jun, 2011	4	0.81	4	0.73	6	0.66	6	0.75	6	0.86
Sep, 2011	3	0.48	4	0.60	4	0.58	6	0.73	4	0.81
Dec, 2011	5	0.89	5	0.84	5	0.75	5	0.82	5	0.89
Mar, 2012	4	0.69	4	0.64	6	0.67	4	0.72	4	0.82
Jun, 2012	2	0.53	3	0.60	4	0.58	4	0.68	4	0.74
Sep, 2012	4	0.69	4	0.67	4	0.65	4	0.71	4	0.77
Dec, 2012	2	0.64	3	0.61	4	0.70	4	0.78	4	0.85
Mar, 2013	3	0.61	3	0.61	4	0.65	6	0.65	5	0.71
Jun, 2013	4	0.68	6	0.73	6	0.66	4	0.73	6	0.83
Sep, 2013	2	0.61	4	0.52	2	0.59	2	0.60	2	0.62
Dec, 2013	3	0.79	2	0.66	6	0.78	6	0.64	6	0.73
Mar, 2014	4	0.79	4	0.70	4	0.66	4	0.74	4	0.82
Jun, 2014	3	0.63	2	0.75	5	0.69	4	0.70	5	0.76
Sep, 2014	5	0.84	4	0.75	5	0.58	5	0.67	4	0.75
Dec, 2014	6	0.79	6	0.70	6	0.62	4	0.45	5	0.79
Mar, 2015	2	0.65	2	0.73	2	0.68	3	0.76	3	0.71
Jun, 2015	2	0.65	2	0.87	4	0.69	4	0.78	4	0.84
Sep, 2015	4	0.69	6	0.61	6	0.62	5	0.64	5	0.76
Dec, 2015	4	0.78	4	0.71	4	0.65	4	0.72	4	0.80
Mar, 2016	4	0.79	4	0.74	5	0.77	4	0.82	4	0.86
Jun, 2016	6	0.78	6	0.72	5	0.81	4	0.66	5	0.78
Sep, 2016	4	0.61	3	0.63	4	0.70	5	0.72	4	0.68
Dec, 2016	3	0.58	2	0.66	4	0.73	4	0.65	3	0.70
Mar, 2017	5	0.82	4	0.77	4	0.75	3	0.81	3	0.69
Jun, 2017	5	0.63	3	0.65	3	0.72	5	0.79	3	0.68
Sep, 2017	6	0.72	4	0.80	4	0.81	6	0.69	6	0.75
Dec, 2017	4	0.65	3	0.73	2	0.68	2	0.79	4	0.66

TABLE III. TRADING RESULTS FOR ALPHA ARBITRAGE STRATEGY

Year	AR	CR	SPR	SOR	MDD
2007	0.095	0.095	1.767	2.541	0.042
2008	0.168	0.279	3.288	5.218	0.014
2009	0.294	0.655	4.358	6.341	0.031
2010	0.169	0.933	3.737	5.209	0.027
2011	0.105	1.136	3.453	5.002	0.027
2012	0.122	1.396	3.904	5.467	0.019
2013	0.123	1.692	3.136	4.505	0.022
2014	0.217	2.276	3.322	5.552	0.027
2015	0.138	2.728	1.479	2.247	0.069
2016	0.089	3.062	1.798	2.735	0.028
2017	0.143	3.643	2.407	3.538	0.045
2007-2017	0.149	3.643	2.681	4.031	0.069

VaR, even the CAViAR (the conditional autoregressive Value-at-Risk) by considering as more aspects as possible, see also [11] for reference.

ACKNOWLEDGMENT

This work is supported by Natural Science Foundation of Zhejiang Province (No. LY17A010002); National Social Science Foundation of China (No. 17BTJ027); First Class Discipline of Zhejiang - A (Zhejiang Gongshang University - Statistics).

REFERENCES

- [1] W. F. Sharpe, "Capital asset prices: a theory of market equilibrium under conditions of risk," *The Journal of Finance*, Vol. 19(3), pp. 425-442, 1964.
- [2] E. F. Fama, and K. R. French, "Common risk factors in the returns on stock and bonds," *Journals of Financial Economics*, Vol. 33, pp. 3-56, 1993.
- [3] E. F. Fama, and K. R. French, "A five-factor asset pricing model," *Journal of Financial Economics*, Vol.116(1), pp. 1-22, 2015.
- [4] B. Bernanke, J. Boivin, and P. Elias, "Factor augmented vector autoregressions (FVARs) and the analysis of monetary policy," *Quarterly Journal of Economics*, Vol. 120, pp. 387-422, 2005.
- [5] A. Banerjee, and M. Marcellino, "Are there any reliable leading indicators for US inflation and GDP growth?" *International Journal of Forecasting*, Vol. 22, pp. 137-151, 2006.
- [6] J. Bai, and S. Ng, "Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions," *Econometrica*, Vol. 74, pp. 1133-1150, 2006.
- [7] T. Ando, and R. S. Tsay, "Quantile regression models with factor-augmented predictors and information criterion," *Econometrics Journal*, Vol. 14, pp. 1-24, 2011.
- [8] R. Koenker, and G. Bassett, "Regression quantile," *Econometrica*, Vol. 46(1), pp. 33-50, 1978.
- [9] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, Vol. 19, pp. 716-723, 1974.
- [10] M. Scholes, and J. William, "Estimating betas from nonsynchronous data," *Journal of Financial Economics*, Vol 5(3), pp. 309-327, 1977.
- [11] X. Yang, C. He, and J. Chen, "Several extended CAViaR models and their applications to the VaR forecasting of the security markets," *Journal of Advanced Computational Intelligent and Intelligent Informatics*, Vol. 20(4), pp. 590-596, 2016.