Applied Mathematics, 2017, 8, 1684-1702
http://www.scirp.org/journal/am
ISSN Online: 2152-7393
ISSN Print: 2152-7385

Analysis of 48 US Industry Portfolios with a New Fama-French 5-Factor Model

Liuling Li[1], Xiao Rao[2], Wentao Zhou[3], Bruce Mizrach[4]

[1]The Institute of Statistics and Econometrics, School of Economics, Nankai University, Tianjin, China
[2]School of Business, Nankai University, Tianjin, China
[3]Department of Economics, University of Wisconsin-Madison, Madison, WI, USA
[4]The Economics Department, Rutgers University, New Brunswick, NJ, USA
Email: liliuling@nankai.edu.cn, raoxiaonku@foxmail.com, zhoukaichengjr@163.com

How to cite this paper: Li, L.L., Rao, X., Zhou, W.T. and Mizrach, B. (2017) Analysis of 48 US Industry Portfolios with a New Fama-French 5-Factor Model. *Applied Mathematics*, **8**, 1684-1702.
https://doi.org/10.4236/am.2017.811122

Received: August 12, 2017
Accepted: November 27, 2017
Published: November 30, 2017

Abstract

In this paper, we analyze US stock market with a new 5-factor model in Zhou and Li (2016) [1]. Data we use are 48 industry portfolios (Jul. 1963-Jan. 2017). Parameters are estimated by MLE. LR and KS are used for model diagnostics. Model comparison is done with AIC. The results show Fama-French 5 factors are still alive. This new model in Zhou and Li (2016) [1] fits the data better than the one in Fama and French (2015) [2].

Keywords

Fama-French 5-Factor Model (FF5), Standardized Standard Asymmetric Exponential Power Distribution (SSAEPD), GARCH

1. Introduction

In 2015, Fama and French suggest a 5-factor model (denoted as FF5-Normal)[1] to capture the market, size, value, profitability and investment patterns in stock returns, which is found better than their 3-factor model in [3]. Since then, many researches about the 5-factor model are developed (see **Table 1**). These researches can be divided into following 2 groups. The 1st group of researches empirically tests the FF5-Normal model using different data. For example, [4] [5] find out that the FF5-Normal model works well in India.

The 2nd group is to extend Fama-French's 5-factor model. For example, [6] [7] choose Betting against Beta (BaB), Gross Profitability (GP) and other 9 factors to create a 14-factor model and find that the market factor is the most important factor for describing expected returns. [1] [8] [9] add the SSAEPD of [10] ([11]

[1]The FF5-Normal model of [2] is in **Appendix 2**.

Table 1. Researches about the 5-Factor model for stock market.

Author (Year)	Research Purpose	Model	Estimation	Data		
			Method	Country	Model factors	Frequency & Period
Hou *et al.* (2014)	Model Comparison	FF5-Normal, C, q-factor	-	USA	Mkt, SMB, RMW, CMA, WML, HMLO	M1972:1-2011:12
Fama *et al.* (2015)	FF4 Extension	FF5-Normal	-	USA	Mkt, SMB, HMLO, RMW, CMA	M1963:7-2013:12
Fama *et al.* (2014)	Model Comparison	CAPM, FF3, FF4, FF5-Normal, FF5+WML	-	USA	Mkt, SMB, HMLO, RMW, CMA, WML	M1963:7-2014:12
Hou *et al.* (2015)	Model Comparison	FF5-Normal, C, q-factor	-	USA	Mkt, SMB, HMLO, RMW, CMA, WML	M1967:1-2013:12
Harshita *et al.* (2015)	Model Comparison	CAPM, FF3, FF5-Normal	-	India	Mkt, SMB, HMLO, RMW, CMA	M1999:10-2014:9
Chiah *et al.* (2015)	Model Comparison	FF3, FF5-Normal	HAC-adj. OLS	Australia	Mkt, SMB, HMLO, RMW, CMA	M1982:12-2012:12
Fama *et al.* (2015)	Empirical Tests	FF5-Normal	-	Global	Mkt, SMB, HMLO, RMW, CMA	M1990:7-2014:9
Harvey *et al.* (2015)	Model Extension	14-factor	OLS	USA	Mkt, SMB, HMLO, RMW, CMA *et al.*	M1968:1-2012:12
Zhou & Li (2016)	Model Extension	FF5-SSAEPD-GARCH	MLE	USA	Mkt, SMB, HMLO, RMW, CMA	

[12] also mention some extension of normal distribution) and the GARCH-type volatilities of [13] into FF5-Normal model (Denote as FF5-SSAEPD-GARCH). They find out their new model has better in-sample fit than that of [2] and the non-normal error assumption of SSAEPD is capable of capturing many stylized facts in financial time series such as skewness and asymmetric fat-tailedness[2].

Based on the new model of [1], in this paper, we try to test following hypothesis: If different data such as 48 industry portfolios[3] are considered, can the new model of [1] still beat the 5-factor model in [2]? To find answers for above question, simulation is used to check the validity of [1]'s MatLab program[4]. Then, 48 industry portfolios are analyzed. Data are downloaded from the French's Data Library, and the sample period is from Jul. 1963 to Jan. 2017. Parameters are estimated by Method of Maximum Likelihood Estimation (MLE). Likelihood Ratio test (LR) and Kolmogorov-Smirnov test (KS) are used for model diagnostics. Model comparison is done with Akaike Information Criterion (AIC).

Simulation results show the MatLab program is valid and can be used for empirical analysis. Empirical results show the 5 factors in [2] are still alive! The GARCH-type volatility and SSAEPD can successfully capture the excess kurtosis. The new model of [1] fits the data well and has better in-sample fit than the 5-factors model of Fama and French.

The organization of this paper is as follows. Section 2 is the model and methodology. Section 3 presents the empirical results. Section 4 provides the conclusions and future extensions. The appendices contain additional information that may be helpful to understand our paper.

[2]The history of SSAEPD is displayed in **Appendix 3**.

[3][1] analyze 25 Fama-French portfolios, which is different from the dataset we use.

[4]Simulation results are listed in **Appendix 4**.

2. Model and Methodology

2.1. The FF5-SSAEPD-GARCH Model

[1] extend Fama-French's 5-factor model based on the GARCH-type volatility in [13] and non-Normal error distribution of SSAEPD in [10], and show their new model is better for 25 Fama-French portfolios. This new model in [1] is listed as follows (denoted as FF5-SSAEPD-GARCH).

$$\begin{aligned} R_t - R_{ft} &= \beta_0 + \beta_1\left(R_{mt} - R_{ft}\right) + \beta_2 SMB_t + \beta_3 HMLO_t + \beta_4 RMW_t \\ &\quad + \beta_5 CMA_t + u_t,\ t = 1, 2, \cdots, T, \end{aligned} \tag{1}$$

$$u_t = \sigma_t z_t,\ z_t \sim SSAEPD\left(\alpha, p_1, p_2\right), \tag{2}$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{r} a_i u_{t-i}^2 + \sum_{i=1}^{s} b_i \sigma_{t-i}^2. \tag{3}$$

$$\sum_{i=1}^{\max(r,s)} \left(a_i + b_i\right) < 1, a_0 > 0, a_i \geq 0, b_i \geq 0. \tag{4}$$

where $\theta = \left(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, a_0, \{a_i\}_{i=1}^{r}, \{b_i\}_{i=1}^{s}, \alpha, p_1, p_2\right)$ is the parameter vector to be estimated. T is the sample size. The error term z_t is distributed as the Standardized Standard Asymmetric Exponential Power Distribution (SSAEPD) proposed by Zhu and Zinde-Walsh. σ_t is the conditional standard deviation, *i.e.*, volatility.

R_t is the return on stock portfolio. R_{ft} is the risk-free return. R_{mt} is the value-weighted market return. SMB_t is the return of small minus big. RMW_t stands for the return of robust minus weak. CMA_t stands for the returns of conservative minus aggressive. $HMLO_t$ is the return of high minus low orthogonalized[5], which is the sum of the intercept and the residual from the regression of HML_t on $R_{mt} - R_{ft}, SMB_t, RMW_t, CMA_t$.

Especially, with

$$\{a_i = 0\}_{i=1}^{r}, \{b_i = 0\}_{i=1}^{s}, \alpha = 0.5, p_1 = p_2 = 2$$

the FF5-SSAEPD-GARCH model is reduced to Fama-French's 5-factor model and in the following section we will compare these two models.

2.2. MLE

Maximum Likelihood Estimation (MLE) is used to estimate previous model. The likelihood function is

$$L\left(\left\{R_t - R_{ft}, R_{mt} - R_{ft}, SMB_t, HMLO_t, RMW_t, CMA_t\right\}_{t=1}^{T}; \theta\right) = \prod_{t=1}^{T} f\left(R_t - R_{ft}\right) \tag{5}$$

$$= \prod_{t=1}^{T} \begin{cases} \dfrac{\delta}{\eta}\left(\dfrac{\alpha}{\alpha^*}\right) K\left(p_1\right) \exp\left(-\dfrac{1}{p_1}\left|\dfrac{\omega + \delta z_t}{2\alpha^*}\right|^{p_1}\right), & z_t \leq -\dfrac{\omega}{\delta}, \\ \dfrac{\delta}{\eta}\left(\dfrac{1-\alpha}{1-\alpha^*}\right) K\left(p_2\right) \exp\left(-\dfrac{1}{p_2}\left|\dfrac{\omega + \delta z_t}{2\left(1-\alpha^*\right)}\right|^{p_2}\right), & z_t > -\dfrac{\omega}{\delta}. \end{cases} \tag{6}$$

[5]The reason of using $HMLO_t$ instead of HML_t can be found in **Appendix 2**.

where

$$z_t = \frac{R_t - R_{ft} - \beta_0 - \beta_1\left(R_{mt} - R_{ft}\right) - \beta_2 SMB_t - \beta_3 HMLO_t - \beta_4 RMW_t - \beta_5 CMA_t}{\sigma_t}, \quad (7)$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{r} a_i u_{t-i}^2 + \sum_{i=1}^{s} b_i \sigma_{t-i}^2. \quad (8)$$

3. Empirical Analysis

3.1. Data

Different from [1], the data we analyze are the monthly returns of 48 industry portfolios for US stock market downloaded from French's Data Library, which include Agriculture, Food, Real Estate, Finance *et al.* The sample period is from 1963:07 to 2017:01. The descriptive statistics of sample data are calculated by MatLab and listed in **Table 2**. For each observation, the skewness (except one

Table 2. Descriptive Statistics (1963:07-2017:01).

	Mea.	Med.	Max.	Min.	St Dev.	Ske.	Kur.	*P*		Mea.	Med.	Max.	Min.	St Dev.	Ske.	Kur.	*P*
Agric	1.02	0.86	22.88	−28.79	6.46	0.01	4.59	0	Guns	1.19	1.33	32.64	−30.08	6.57	−0.14	5.04	0
Food	1.08	1.00	19.59	−17.88	4.39	0.15	5.20	0	Gold	1.01	0.94	78.68	−33.57	10.58	0.74	7.78	0
Soda	1.16	1.48	38.27	−26.26	6.33	0.13	7.28	0	Mines	1.12	0.70	26.91	−34.75	7.50	−0.31	4.87	0
Beer	1.13	1.11	26.09	−19.76	5.15	−0.02	5.49	0	Coal	1.11	0.67	46.06	−37.94	10.43	0.17	5.02	0
Smoke	1.40	1.70	32.47	−24.93	6.08	−0.05	5.63	0	Oil	1.01	0.96	24.66	−18.21	5.41	0.06	4.30	0
Toys	0.84	0.97	26.88	−34.41	7.14	−0.18	4.26	0	Util	0.83	0.90	18.84	−12.65	4.01	−0.08	4.13	0
Fun	1.35	1.38	38.77	−31.86	7.68	−0.21	5.94	0	Telcm	0.88	1.02	21.34	−16.22	4.61	−0.20	4.31	0
Books	0.94	0.68	30.74	−25.19	5.83	0.02	5.04	0	PerSv	0.70	0.86	27.00	−28.25	6.90	−0.15	4.47	0
Hshld	0.90	1.07	18.54	−21.64	4.64	−0.31	4.95	0	BusSv	1.13	1.37	25.43	−27.54	6.48	−0.15	4.40	0
Clths	1.08	1.10	32.37	−30.90	6.46	−0.06	5.65	0	Comps	0.96	0.78	24.24	−32.36	6.95	−0.18	4.74	0
MedEq	1.14	1.34	21.03	−20.56	5.27	−0.35	4.27	0	Chips	1.04	1.52	27.27	−32.23	7.36	−0.37	4.79	0
Drugs	1.09	1.10	31.80	−19.11	4.97	0.15	5.71	0	LabEq	1.08	1.14	22.04	−30.15	7.06	−0.19	4.25	0
Chems	0.95	0.96	22.05	−28.00	5.55	−0.11	5.33	0	Paper	0.98	1.00	24.27	−26.35	5.48	0.11	5.31	0
Rubbr	1.09	1.32	31.95	−30.57	6.06	−0.22	5.67	0	Boxes	1.02	1.05	20.87	−28.24	5.61	−0.39	5.06	0
Txtls	1.07	1.05	59.04	−32.51	7.13	0.48	12.58	0	Trans	0.97	1.11	19.02	−27.90	5.79	−0.24	4.19	0
BldMt	1.02	1.18	35.51	−30.67	6.12	−0.03	7.11	0	Whlsl	1.04	1.34	18.12	−28.64	5.62	−0.34	5.31	0
Cnstr	0.98	0.90	24.01	−31.10	7.24	−0.08	3.90	0	Rtail	1.05	0.77	27.07	−29.17	5.38	−0.16	5.28	0
Steel	0.75	0.67	30.67	−32.91	7.37	−0.19	5.29	0	Meal	1.18	1.52	27.88	−31.28	6.12	−0.47	5.58	0
FabPr	0.71	0.71	30.38	−26.67	7.27	−0.09	4.21	0	Banks	0.94	0.96	25.03	−27.25	5.97	−0.24	5.08	0
Mach	0.99	1.25	23.05	−31.19	6.20	−0.39	5.49	0	Insur	1.02	1.14	26.84	−26.46	5.63	−0.04	5.36	0
ElcEq	1.16	0.89	23.22	−32.20	6.23	−0.19	4.69	0	RlEst	0.67	0.73	59.60	−37.73	7.56	0.54	11.89	0
Autos	0.84	0.67	49.57	−36.41	6.78	0.24	8.99	0	Fin	1.12	1.39	19.56	−25.91	6.15	−0.44	4.34	0
Aero	1.18	1.31	25.33	−30.33	6.68	−0.29	4.77	0	Hlth*	1.02	1.16	36.47	−39.11	8.20	−0.04	5.67	0
Ships	1.09	1.08	29.15	−32.27	7.20	0.00	4.60	0	Other	0.65	0.81	21.40	−26.37	6.86	−0.38	4.48	0

Notes: The sample period of Hlth is 1969:7-2017:01 due to the data availability; Mea. = mean, Med. = median, Max. = maximum, Min. = minmum St Dev. = standard deviation, Ske. = skewness Kur. = kurtosis, *P* = *P*-value of Jarque-Bera Test.

portfolio, the "Ships" industry) is not 0 and the kurtosis is more than 3. The *p*-value of Jarque-Bera test for each portfolio is 0, which is smaller than 5% significance level. Hence, we can reject the null hypothesis and conclude that the asset returns do not follow Normal distribution. Thus, non-Normal error assumption of SSAEPD might be able to fit the data better.

3.2. Estimation Results

The estimates for our new model are displayed in Table 3. We find out that our model can successfully capture the skewness, fat-tailness and excess kurtosis of the data. More specifically, the skewness parameter α of 46 out of 48 estimates are not equal to 0.5, which captures the skewness in the data. 84 out of 96 estimates for the tail parameters $p_i(i=1,2)$ are smaller than 2, which suggests that portfolio returns are fat-tailed distributed. Besides, all the tail parameters p_1 and p_2 (except one potfolio, the "Other" industry) are not equal to each other, which documents the asymmetric fat-tailedness. And 28 out of 48 portfolios have bigger estimates for the left tail parameter P_1 which means that these returns tend to have thinner left tails.

3.3. Model Diagnostics

To test the significance of coefficients in FF5-SSAEPD-EGARCH, Likelihood Ratio test (LR) is applied[6], which is calculated using Equation (9).

$$LR = -2\ln(\text{likelihood for null}) + 2\ln(\text{likelihood for alternative}) \quad (9)$$

3.3.1. Tests for Parameter Restrictions

- **Tests for Parameters in the Mean Equation**

The *P*-values of LR are listed in **Table 4**. The null hypothesis of the joint significance test is $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$. The *P*-values of the joint significance test for all the 48 portfolios are 0, which means $\beta_1, \beta_2, \beta_3, \beta_4$ and β_5 are statistically jointly significant under 5% significance level.

The individual significance tests show that under 5% significance level the coefficient β_1 in all 48 portfolios are statistically significant; 40/48, 28/48, 38/48 and 32/48 portfolios have a statistically coefficient $\beta_2, \beta_3, \beta_4$ and β_5, respectively. As for coefficient β_0 (*i.e.*, the Alpha return), 31 out of the 48 portfolios are statistically significant under 5% significance level. Thus, most of the 48 portfolios seem to be able to earn the *Alpha* returns.

As a whole, since the 5 factors are significant in most of the 48 portfolios, therefore we can conclude that with non-Normal errors such as SSAEPD and GARCH-type volatilities, the Fama-French 5 factors are still alive.

- **Tests for Parameters in the GARCH Equation**

In this part, some restrictions on the parameters in the GARCH equation are tested with Likelihood Ratio test (LR). And the results are listed in Table 5. Results show the GARCH-type volatility should be included in Fama-French

[6]LR formula is from Neyman and Pearson (1993).

Table 3. Estimates for FF5-SSAEPD-GARCH (Monthly, 1963:07-2017:01).

	β_0	β_1	β_2	β_3	β_4	β_5	α	p_1	p_2	a	b	c
Agric	0.03	0.86	0.48	−0.01	0.22	0.03	0.48	1.50	1.37	18.91	0.25	0.00
Food	−0.19	0.83	0.01	−0.05	0.45	0.45	0.71	2.01	0.91	0.27	0.09	0.88
Soda	0.31	0.83	−0.04	−0.15	0.68	0.09	0.63	1.36	0.89	13.44	0.47	0.09
Beer	−1.59	0.97	0.15	−0.14	0.59	0.70	0.56	0.98	1.41	2.72	0.07	0.84
Smoke	0.34	0.83	−0.25	−0.22	0.61	0.42	0.68	1.36	0.86	22.72	0.10	0.00
Toys	−0.56	1.20	0.66	0.11	0.72	0.21	0.42	1.37	1.77	16.45	0.16	0.00
Fun	−0.32	1.30	0.56	0.10	0.28	0.10	0.40	1.57	1.93	0.73	0.07	0.89
Books	−0.07	1.06	0.35	0.19	0.49	0.09	0.31	1.13	2.00	1.08	0.09	0.80
Hshld	−0.34	0.97	0.97	−0.13	0.57	0.10	0.49	1.16	1.43	0.44	0.08	0.86
Clths	−0.14	1.05	0.65	0.05	0.42	0.14	0.28	0.99	1.91	0.85	0.15	0.79
MedEq	0.34	0.86	0.10	−0.31	0.24	−0.33	0.51	1.69	1.89	8.53	0.21	0.00
Drugs	0.39	0.88	−0.25	−0.43	0.16	−0.26	0.56	1.87	1.60	0.69	0.10	0.83
Chems	−0.44	1.17	0.01	0.15	0.27	0.52	0.40	1.39	1.82	0.49	0.11	0.83
Rubbr	−0.25	1.07	0.70	0.16	0.41	0.27	0.51	1.61	1.42	0.45	0.07	0.88
Txtls	−0.41	1.08	0.68	0.29	0.58	0.74	0.51	1.35	1.28	10.96	0.33	0.00
BldMt	−0.31	1.19	0.41	0.27	0.41	0.46	0.50	1.62	1.48	0.22	0.08	0.90
Cnstr	−0.43	1.27	0.62	0.21	0.43	0.20	0.23	1.02	2.22	15.41	0.03	0.00
Steel	−0.51	1.23	0.25	0.32	−0.47	0.43	0.48	1.73	1.54	0.68	0.09	0.88
FabPr	−0.42	1.01	0.72	0.25	0.29	−0.05	0.46	1.37	1.23	21.06	0.16	0.00
Mach	−0.11	1.20	0.31	0.18	0.00	0.10	0.14	0.64	2.78	7.25	0.20	0.00
ElcEq	0.06	1.21	0.17	−0.03	0.12	0.03	0.68	2.79	1.25	7.72	0.23	0.00
Autos	0.19	1.10	0.13	0.54	0.15	0.31	0.54	1.60	1.17	0.82	0.09	0.87
Aero	−0.12	1.18	0.30	0.25	0.59	0.20	0.41	1.16	1.39	14.92	0.14	0.00
Ships	−1.36	1.29	0.42	0.18	0.37	1.06	0.60	2.02	1.36	1.01	0.08	0.88
Estimates for FF5-SSAEPD-GARCH (Monthly, 1963:07-2017:01)												
Guns	−0.08	0.98	0.37	0.48	0.72	0.36	0.24	0.88	1.84	23.20	0.08	0.00
Gold	1.04	0.46	0.53	−0.06	−0.19	−0.20	0.48	1.75	1.18	2.01	0.03	0.95
Mines	−0.10	1.09	0.42	0.26	0.19	0.41	0.43	1.48	1.83	25.74	0.14	0.00
Coal	0.37	1.03	0.29	0.05	0.05	−0.15	0.67	2.59	1.10	1.38	0.14	0.86
Oil	0.15	0.92	−0.16	0.12	0.12	0.53	0.53	1.81	1.35	11.37	0.30	0.00
Util	0.66	0.61	−0.29	0.27	0.04	0.10	0.45	1.65	1.84	0.47	0.11	0.84
Telcm	0.13	0.85	−0.27	0.05	−0.10	0.32	0.33	1.25	2.09	0.33	0.09	0.87
PerSv	−0.57	1.09	0.69	0.14	0.70	0.07	0.56	1.33	1.27	15.03	0.18	0.00
BusSv	0.39	1.08	0.34	−0.30	−0.26	−0.69	0.70	2.39	0.93	4.50	0.23	0.00
Comps	−0.28	1.03	0.13	−0.38	−0.42	−0.53	0.65	2.02	1.25	0.80	0.06	0.89
Chips	−0.56	1.18	0.37	−0.37	−0.34	−0.28	0.33	1.11	2.34	0.28	0.09	0.89
LabEq	0.34	1.09	0.48	−0.50	−0.24	−0.48	0.46	1.27	1.26	9.47	0.24	0.00
Paper	−0.28	1.10	0.11	0.15	0.44	0.53	0.58	1.54	0.96	8.79	0.07	0.00

Continued

Boxes	0.24	0.97	−0.02	0.08	0.28	−0.05	0.43	1.54	2.29	0.37	0.08	0.89
Trans	−0.28	1.12	0.32	0.31	0.48	0.28	0.55	1.97	1.39	9.43	0.00	0.00
Whlsl	0.10	0.98	0.55	−0.05	0.21	−0.14	0.61	1.55	1.06	0.18	0.13	0.84
Rtail	0.00	1.03	0.17	−0.04	0.45	−0.04	0.49	1.47	1.50	7.79	0.10	0.00
Meal	−0.16	1.07	0.47	−0.03	0.85	0.10	0.68	2.17	1.20	9.22	0.24	0.00
Banks	−0.44	1.23	−0.07	0.75	0.20	0.55	0.51	1.56	1.40	0.76	0.12	0.80
Insur	−0.10	1.07	−0.04	0.34	0.07	0.31	0.51	1.27	1.24	7.11	0.46	0.00
RlEst	−0.75	1.14	0.95	0.53	0.49	0.42	0.60	1.49	0.99	15.41	0.09	0.00
Fin	0.12	1.18	0.11	0.44	−0.24	0.02	0.60	1.51	0.98	4.10	0.39	0.00
Hlth*	−0.32	1.09	0.76	−0.18	0.56	0.10	0.38	0.91	1.36	28.29	0.13	0.00
Other	−0.22	1.11	0.40	0.04	0.06	0.03	0.50	1.01	1.01	4.63	0.65	0.29

Notes: The data period of Hlth is 1969:7-2017:01 due to the data availability.

Table 4. *P*-values of Likelihood Ratio Test (LR).

	TJ	T0	T1	T2	T3	T4	T5		TJ	T0	T1	T2	T3	T4	T5
Agric	0*	0.89	0*	0*	0.93	0.01*	0.76	Guns	0*	0.42	0*	0*	0*	0*	0*
Food	0*	0*	0*	0.62	0.17	0*	0*	Gold	0*	0*	0*	0*	0.79	0.37	0.25
Soda	0*	1.00	0*	1.00	0.06	0*	0.36	Mines	0*	0.67	0*	0*	0.04*	0.05*	0*
Beer	0*	1.00	0*	1.00	1.00	1.00	1.00	Coal	0*	0.02*	0*	0*	0.74	0.81	0.27
Smoke	0*	0.09	0*	0*	0.04*	0*	0*	Oil	0*	0.17	0*	0.01*	0.11	0.12	0*
Toys	0*	0*	0*	0*	0.21	0*	0.04*	Util	0*	0*	0*	0*	0*	0.94	0.11
Fun	0*	0*	0*	0*	0.34	0*	0.21	Telcm	0*	0.03*	0*	0*	0.33	0.15	0*
Books	0*	0.31	0*	0*	0*	0*	0.07	PerSv	0*	0*	0*	0*	0.09	0*	0.42
Hshld	0*	0*	0*	0.19	0.01*	0*	0.04*	BusSv	0*	0*	0*	0*	0*	0*	0*
Clths	0*	0.03*	0*	0*	0.15	0*	0.02*	Comps	0*	0*	0*	0.3*	0*	0*	0*
MedEq	0*	0.01*	0*	0.04*	0*	0*	0*	Chips	0*	0*	0*	0*	0*	0*	0*
Drugs	0*	0*	0*	0*	0*	0.03*	0*	LabEq	0*	0.01*	0*	0*	0*	0*	0*
Chems	0*	0*	0*	0.79	0.01*	0*	0*	Paper	0*	0.03*	0*	0.01*	0.01*	0*	0*
Rubbr	0*	0*	0*	0*	0.01*	0*	0*	Boxes	0*	0*	0*	0.67	0.25	0*	0.38
Txtls	0*	0.01*	0*	0*	0*	0*	0*	Trans	0*	0.03*	0*	0*	0*	0*	0*
BldMt	0*	0*	0*	0*	0*	0*	0*	Whlsl	0*	0.02*	0*	0*	0.21	0*	0*
Cnstr	0*	0.01*	0*	0*	0*	0*	0.02*	Rtail	0*	0.98	0*	0*	0.51	0*	0*
Steel	0*	0*	0*	0*	0*	0*	0*	Meal	0*	0.25	0*	0*	0.68	0*	0.17
FabPr	0*	0.03*	0*	0*	0.02*	0*	0.66	Banks	0*	0*	0*	0.09	0*	0*	0*
Mach	0*	0.03*	0*	0*	0*	0.02*	0.01*	Insur	0*	0.41	0*	0.40	0*	0.46	0*
ElcEq	0*	0.62	0*	0*	0.56	0.04*	0.61	RlEst	0*	0*	0*	0*	0*	0*	0*
Autos	0*	0.04*	0*	0.04*	0*	0.12	0*	Fin	0*	0.20	0*	0*	0*	0*	0.71
Aero	0*	0.46	0*	0*	0*	0*	0.03*	Hlth*	0*	0.18	0*	0*	0.47	0*	0.35
Ships	0*	1.00	0*	0*	1.00	0.17	0*	Other	0*	0.13	0*	0*	0.06	0.43	0.73

Notes: Sample of Hlth is from 1969:7 to 2017:01 due to the availability of data. TJ means test of joint hypothesis of $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$. T0 means $H_0: \beta_0 = 0$. T1 means $H_0: \beta_1 = 0$, T2 means $H_0: \beta_2 = 0$. T3 means $H_0: \beta_3 = 0$. T4 means $H_0: \beta_4 = 0$. T5 means $H_0: \beta_5 = 0$.

Table 5. *P*-values of likelihood ratio test (LR).

	T8	T9	T10	T11	T12	T13	T14	T15	T16		T8	T9	T10	T11	T12	T13	T14	T15	T16
Agric	0*	0*	0*	0.83	0*	0.71	0*	0.13	0*	Guns	0*	0*	0.10	0.73	0*	0.05*	0*	0*	0*
Food	0*	0*	1.00	0*	0*	0.02*	0*	0.96	0*	Gold	0*	0*	0.39	0*	0*	0.74	0*	0.63	0.20
Soda	0*	0*	0.37	1.00	0*	0.03*	0*	0.06	0*	Mines	0*	0*	0*	1.00	0.10	0.35	0.04*	0.07	0.59
Beer	0*	0*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Coal	0*	0*	0*	0*	0.01*	0.07	0.02*	0.11	0.01*
Smoke	0*	0*	0.03*	0.79	0*	0.12	0*	0*	0*	Oil	0*	0*	0*	1.00	0*	0.73	0*	0.59	0.01*
Toys	0.08	0*	0.02*	1.00	0.01*	0.31	0*	0*	0.49	Util	0*	0*	0*	0*	0.51	0.68	0.30	0.24	0.71
Fun	0*	0*	0*	0*	0.20	0.10	0.32	0.29	0.97	Telcm	0*	0*	0*	0*	0.02*	0.04*	0.01*	0.01*	0.72
Books	0*	0*	0*	0*	0*	0.05*	0*	0.01*	1.00	PerSv	0*	0*	0.01*	0*	0*	0.24	0*	0*	0*
Hshld	0*	0*	0*	0*	0*	0.88	0*	0*	0.06	BusSv	0*	0*	0*	0*	0*	0.05*	0*	0.20	0*
Clths	0*	0*	0*	0*	0*	0.07	0*	0*	0.61	Comps	0*	0*	0*	1.00	0.03*	0.04*	0.01*	0.92	0.01*
MedEq	0*	0*	0*	0.56	0.26	0.90	0.39	0.26	0.76	Chips	0*	0*	0*	0*	0.15	0.36	0.12	0.16	0.25
Drugs	0*	0*	0*	0*	0.34	0.68	0.19	0.72	0.58	LabEq	0*	0*	0*	1.00	0*	0.47	0*	0*	0*
Chems	0*	0*	0*	0*	0.05*	0.16	0.03*	0.04*	0.47	Paper	0*	0*	0*	0*	0*	0.03*	0*	0.04*	0*
Rubbr	0*	0*	0*	0*	0*	0.86	0*	0.33	0.04*	Boxes	0*	0*	0*	0*	0.24	0.57	0.38	0.18	0.59
Txtls	0*	0*	0*	1.00	0*	0.82	0*	0.01*	0*	Trans	0*	0*	1.00	0*	0.02*	0.61	0.06	0.89	0.13
BldMt	0*	0*	0*	0*	0*	0.96	0*	0.22	0.46	Whlsl	0*	0*	0*	0*	0*	0.15	0*	0.03*	0*
Cnstr	0*	0*	0.57	0*	0*	0*	0*	0*	0.06	Rtail	0*	0*	0.04*	1.00	0.01*	0.83	0*	0*	0.05*
Steel	0*	0*	0*	0*	0.01*	0.85	0.02*	0.72	0*	Meal	0*	0*	0*	0.30	0.02*	0.02*	0.01*	0.57	0*
FabPr	0.01*	0*	0*	0.33	0*	0.53	0*	0.06	0*	Banks	0*	0*	0*	1.00	0*	0.63	0*	0.08	0.02*
Mach	0*	0*	0*	0.03*	0*	0*	0*	0*	0.03*	Insur	0*	0*	0*	0*	0*	0.77	0*	0.01*	0*
ElcEq	0*	0*	0*	1.00	0.09	0.09	0.02*	0.12	0.03*	RlEst	0.31	0*	0.13	0.68	0*	0.05*	0*	0.11	0*
Autos	0*	0*	0*	0*	0*	0.69	0*	0.25	0*	Fin	0*	0*	0*	0.96	0*	0.03*	0*	0.01*	0*
Aero	0*	0*	0.03*	0*	0*	0.30	0*	0*	0.02*	Hlth*	0*	0*	0*	0.51	0*	0.13	0*	0*	0*
Ships	0*	0*	1.00	0*	0.03*	0.17	0.02*	0.97	0.57	Other	0*	0*	0*	1.00	0*	0*	0*	0*	0*

Notes: The data period of Hlth is 1969:7-2017:01 due to the lack of data from 1963:7-1969:6. T8 means $H_0: b = c = 0$. T9 means $H_0: a = 0$. T10 means $H_0: b = 0$. T11 means $H_0: c = 0$. T12 means $H_0: \alpha = 0.5, p_1 = p_2 = 2$. T13 means $H_0: \alpha = 0.5$. T14 means $H_0: p_1 = p_2 = 2$. T15 means $H_0: p_1 = 2$. T16 means $H_0: p_2 = 2$.

5-factor model. For instance, we do the joint significance test for hypothesis $H_0: b = c = 0$. For 46 out of the 48 portfolios, the *p*-value of the LR are smaller than the significance level 5%, which means our GARCH-type volatilities are quite necessary. As for individual hypotheses, we discover that most *P*-values of LR are smaller than the significance level 5%. And to be specific, ARCH term ($H_0: b = 0$) is significant in 39 out of 48 portfolios and GARCH term ($H_0: c = 0$) is significant in 27 out of 48 portfolios.

- **Tests for Parameters in SSAEPD**

We also run significance tests for the parameters in the SSAEPD and the results of parameter restrictions show strong non-Normality. For example, for the

Hypothesis $H_0: \alpha = 0.5, p_1 = p_2 = 2$, 39 out of 48 *p*-values are smaller than the significance level 5%, which means that Normal error assumption is not supported by most of our data. Besides, Asymmetry is documented ($H_0: \alpha = 0.5$ is rejected by 14 out of 48 portfolios). And non-normality is found ($H_0: p_1 = 2$ is rejected by 21 out of 48 portfolios and 29 out of 48 portfolios reject the null $H_0: p_2 = 2$).

3.3.2. Residual Check

In this subsection, the residuals for previous models are checked with both Kolmogorov-Smirnov test and graphs. Our results show 41 out of the 48 portfolios have residuals which do follow SSAEPD. That means, the FF5-SSAEPD-GARCH is adequate for the 48 industry portfolios. But the FF5-Normal model is not adequate for the data, since all the 48 portfolios have residuals which do not follow the Normal error distribution.

- **Kolmogorov-Smirnov Test for Residuals**

To check the residuals, the Kolmogorov-Smirov test (KS)[7] is employed. The *p*-value of KS test is displayed in Table 6. The *p*-values of KS test[8] show the residuals from the new model do follow SSAEPD. For instance, the *p*-value of the portfolio of Agriculture industry is 0.79, greater than 5%, which means under 5% significance level, the null hypothesis is not rejected and the residuals from FF5-SSAEPD-EGARCH do follow the SSAEPD. Similarly, the null hypothesis cannot be rejected for all other 40 portfolios.

Then, we apply the KS test for the residuals from the FF5-Normal model[9]. The *p*-values of the KS test are also listed in Table 6. All of the 48 portfolios have smaller *p*-values than 0.05, which means these 48 industry portfolios reject the nulls. Hence, the error terms of the portfolios do not follow Normal distribution. And the FF-Normal model is not adequate for the data.

- **PDFs of Residuals**

By method of "eye-rolling", the PDF of residuals is compared with theoretical PDFs. Taking the portfolio of Agriculture industry for example, in Figure 1(a), the probability density function (PDF) for the estimated residuals $\hat{z}_t$ in FF5-SSAEPD-EGARCH and that of $SSAEPD(\hat{\alpha}, \hat{p}_1, \hat{p}_2)$ are plotted. These curves are very close to each other, which means the residuals are distributed as SSAEPD. Hence, the FF5-SSAEPD-GARCH model fits the data well.

Similarly, the probability density function (PDF) for the estimated residuals $\hat{u}_t$ in FF5-Normal and that of $Normal(\hat{\mu}, \hat{\sigma}^2)$ are shown in Figure 1(b). And there are big differences between these two curves, which means the residuals do not follow Normal distribution.

[7]The null hypothesis of KS test is H_0: Data follows a specified distribution. If the *P*-value of KS test is bigger than 5% significance level, the null hypothesis is not rejected. Otherwise, the null hypothesis is rejected.

[8]The null hypothesis is H_0: FF5-SSAEPD-GARCH residuals are distributed as $SSAEPD(\hat{\alpha}, \hat{p}_1, \hat{p}_2)$.

[9]The FF5-Normal model is listed in Appendix 2. The null hypothesis H_0: FF5-Normal residuals are distributed as $Normal(\hat{\mu}, \hat{\sigma}^2)$.

Table 6. *P*-values of KS Test for Residuals.

	M1	M2		M1	M2
Agric	0.79	0*	Guns	0.40	0*
Food	0.59	0*	Gold	0.29	0*
Soda	0.80	0*	Mines	0.98	0*
Beer	0*	0*	Coal	0.42	0*
Smoke	0.58	0*	Oil	0.96	0*
Toys	0.76	0*	Util	0*	0*
Fun	0.04*	0*	Telcm	0.91	0*
Books	0.51	0*	PerSv	0.68	0*
Hshld	0.25	0*	BusSv	0.95	0*
Clths	0.86	0*	Comps	0*	0*
MedEq	0.99	0*	Chips	0*	0*
Drugs	0.67	0*	LabEq	0.88	0*
Chems	0.22	0*	Paper	0.98	0*
Rubbr	0.94	0*	Boxes	0.49	0*
Txtls	0.93	0*	Trans	0.87	0*
BldMt	0.93	0*	Whlsl	0.17	0*
Cnstr	1.00	0*	Rtail	0.99	0*
Steel	0.66	0*	Meal	0.96	0*
FabPr	1.00	0*	Banks	0.65	0*
Mach	0.40	0*	Insur	0.97	0*
ElcEq	0.99	0*	RlEst	0.96	0*
Autos	0*	0*	Fin	0.81	0*
Aero	0.67	0*	Hlth*	0.72	0*
Ships	0*	0*	Other	0.96	0*

Notes: 1. The data period of Hlth is 1969:7-2017:01 due to the data availability. 2.* means the data doesn't follow the specified distribution under 5% significance level. M1 = FF5-SSAEPD-GARCH, M2 = FF5-Normal.

3.4. Model Comparison

In this subsection, we compare the model in [1] with the 5-factor model of Fama and French. The Akaike Information Criterion (AIC) is used as the model selection criterion. Table 7 lists the AIC values. We find that 47 out of 48 AIC values of the FF5-SSAEPD-GARCH model are smaller than those of the FF5-Normal model. Hence, we conclude that the new model we used (FF5-SSAEPD-GARCH) is better than the 5-factor model of Fama and French.

4. Conclusions

In this paper, we empirically test the new 5-factor model suggested in [1]. Their new model generalizes the 5-factor model in [2] by introducing a non-normal

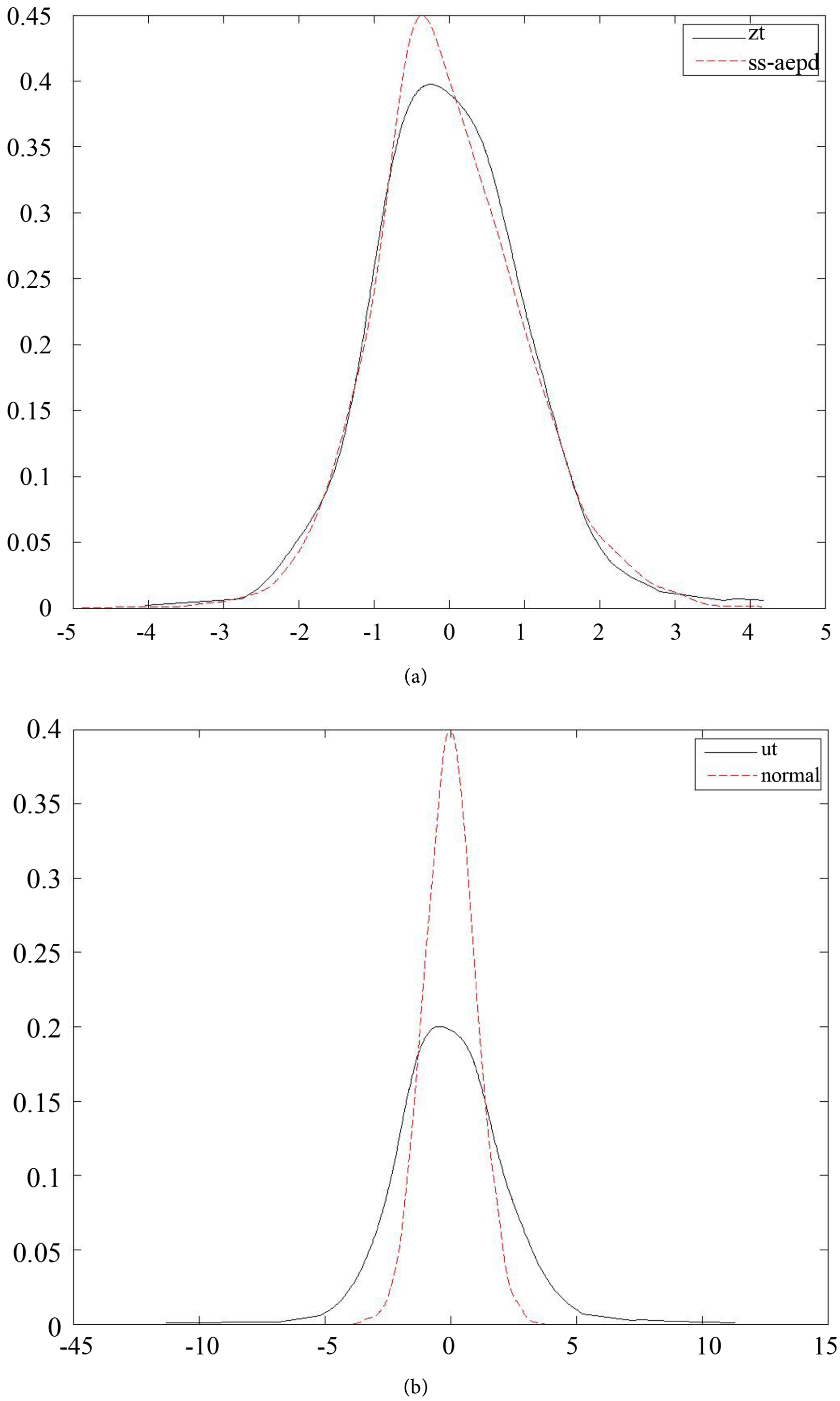

Figure 1. Comparison of PDFs. (a) PDFs of the Residuals (FF5-SSAEPD-GARCH) and $SSAEPD(\hat{\alpha}, \hat{p}_1, \hat{p}_2)$; (b) PDFs of the Residuals (FF5-Normal) and $Normal(\hat{\mu}, \hat{\sigma}^2)$.

error term and time-varying volatilities. The non-normal error assumption their used is the SSAEPD of [10] and the time-varying volatilities is the GARCH model of [13]. For comparison, monthly US stock returns of 48 industry portfolios (1963:07-2017:01) are analyzed.

Table 7. AIC Values (Monthly, 1963:07-2017:01).

	M1	M2		M1	M2
Agric	6.00*	6.08	Guns	6.03*	6.10
Food	4.81*	4.97	Gold	7.35*	7.50
Soda	5.91*	6.07	Mines	6.24*	6.27
Beer	5.58	5.46*	Coal	6.92*	7.26
Smoke	5.96*	6.09	Oil	5.54*	5.62
Toys	5.81*	5.84	Util	5.06*	5.09
Fun	5.81*	5.87	Telcm	4.91*	5.01
Books	5.06*	5.12	PerSv	5.69*	5.75
Hshld	4.64*	4.83	BusSv	4.52*	4.61
Clths	5.31*	5.48	Comps	5.58*	5.60
MedEq	5.21*	5.25	Chips	5.20*	5.35
Drugs	5.05*	5.15	LabEq	5.26*	5.35
Chems	4.85*	4.96	Paper	5.01*	5.12
Rubbr	4.98*	5.09	Boxes	5.26*	5.41
Txtls	5.51*	5.77	Trans	5.10*	5.11
BldMt	4.75*	4.90	Whlsl	4.39*	4.67
Cnstr	5.60*	5.63	Rtail	5.00*	5.02
Steel	5.63*	5.76	Meal	5.30*	5.35
FabPr	6.00*	6.07	Banks	5.03*	5.15
Mach	4.96*	5.01	Insur	5.17*	5.34
ElcEq	5.12*	5.16	RlEst	5.60*	5.75
Autos	5.60*	5.73	Fin	4.53*	4.75
Aero	5.64*	5.71	Hlth*	5.61*	5.79
Ships	6.06*	6.09	Other	6.22*	6.32

Notes: 1. The data period of Hlth is 1969:7-2017:01 due to the data availability. 2. Numbers with * are smaller AIC values. M1 = FF5-SSAEPD-GARCH, M2 = FF5-Normal.

Method of Maximum Likelihood (MLE) is used for parameters estimation. Likelihood Ratio Test (LR) is used to test the hypotheses of parameter restrictions. Kolmogorov-Smirnov test (KS) is used to check residuals. Akaike Information Criterion (AIC) is used to compare models.

Simulation results show the MatLab program for the new 5-factor model in [1] is valid. And empirical results show 1) this new model can capture the skewness, fat tails and asymmetric fat-tailedness in the data. 2) the Fama-French 5 factors are still alive even if the non-normal errors and GARCH-type volatilities are considered. Since we find out the 5 factors are statistically significant in most of the 48 portfolios. And 3) FF5-SSAEPD-GARCH model can fit the data much better than the 5-factor model in [2].

Future extensions will include but not limited to follows. First, we can exam our results with different data. Second, we can compare our results with those from other models such as ARIMA model. Last but not the least, other factors can be introduced into this model to explain the *Alpha* returns of industry portfolios.

References

[1] Zhou, W. and Li, L. (2016) A New Fama-French 5-Factor Model Based on SSAEPD Error and GARCH-Type Volatility. *Journal of Mathematical Finance*, **6**, 711-727. https://doi.org/10.4236/jmf.2016.65050

[2] Fama, E.F. and French, K.R. (2015) A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, **116**, 1-22. https://doi.org/10.1016/j.jfineco.2014.10.010

[3] Fama, E.F. and French, K.R. (1993) Common Risk Factors in Returns on Stocks and Bonds. *Journal of Financial Economics*, **33**, 3-56. https://doi.org/10.1016/0304-405X(93)90023-5

[4] Harshita, S.S. and Yadav, S.S. (2015) Indian Stock Market and the Asset Pricing Models. *Procedia Economics* & *Finance*, **30**, 294-304. https://doi.org/10.1016/S2212-5671(15)01297-6

[5] Gharghori, P., Chan, H. and Faff, R. (2007) Are the Fama-French Factors Proxying Default Risk? *Australian Journal of Management*, **32**, 223-249. https://doi.org/10.1177/031289620703200204

[6] Hou, K., Xue, C. and Zhang, L. (2015) Digesting Anomalies: An Investment Approach. *Review of Financial Studies*, **28**, 650-705.

[7] Harvey, C.R. and Liu, Y. (2015) Lucky Factors. Social Science Electronic Publishing, San Francisco.

[8] Li, L., Zhu, Q.Y. and Mu, Y. (2017) Analysis of the Sector of Software & Computer Services with a New Carhart 4-Factor Model. *Journal of Mathematical Finance*, **7**, 65-82.

[9] Li, L., Zhu, Q. and Yang, Y. (2017) A New 5-factor Model Based on the EGARCH-Typed Volatilities and the SSAEPD Errors. *Journal of Multidisciplinary*, **4**, 83-98.

[10] Zhu, D. and Zinde-Walsh, V. (2009) Properties and Estimation of Asymmetric Exponential Power Distribution. *Journal of Econometrics*, **148**, 86-99. https://doi.org/10.1016/j.jeconom.2008.09.038

[11] Subbotin, M.T. (1923) On the Law of Frequency of Error. *Матем. сб.*, **31**, 296-301.

[12] Azzalini, A.A. (1985) Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics*, **12**, 171-178.

[13] Engle, R.F. and Bollerslev, T. (1986) Modeling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 1-50. https://doi.org/10.1080/07474938608800095

Appendix 1.

Four-digit SIC codes are used to assign firms to 48 industries. The variables defined in the 1st column in **Table 8** are used as the dependent variables in this paper.

Appendix 2. Fama-French 5-Factor Model (FF5-Normal)

Equation (10) is the new 5-factor model (denoted as FF5-Normal) suggested by Fama and French (2015). And they show this model empirically outperforms Fama-French (1993)'s 3-factor model.

$$\begin{aligned} R_t - R_{ft} &= \beta_0 + \beta_1 * \left(R_{mt} - R_{ft}\right) + \beta_2 * SMB_t + \beta_3 * HMLO_t \\ &\quad + \beta_4 * RMW_t + \beta_5 * CMA_t + u_t, u_t \sim Normal\left(\mu, \sigma^2\right). \end{aligned} \tag{10}$$

where $\theta = \left(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \mu, \sigma\right)$ are parameters in this model. $t = 1, 2, \cdots, T$ R_t is the return on stock portfolio. R_{ft} is the risk-free return. R_{mt} is the value-weighted market return. SMB_t is the return of small minus big. RMW_t stands for the return of robust minus weak. CMA_t stands for the returns of conservative minus aggressive.

$HMLO_t$ is the return of high minus low orthogonalized, which is the sum of the intercept and the residual from the regression of HML_t on $R_{mt} - R_{ft}, SMB_t, RMW_t, CMA_t$. The reason of using $HMLO_t$ instead of HML_t is that Fama and French (2015) show HML_t (the high minus low book-to-market ratio) is redundant in following 5-factor model.

$$\begin{aligned} R_t - R_{ft} &= \beta_0 + \beta_1 * \left(R_{mt} - R_{ft}\right) + \beta_2 * SMB_t + \beta_3 * HML_t \\ &\quad + \beta_4 * RMW_t + \beta_5 * CMA_t + u_t, \\ u_t &\sim Normal\left(\mu, \sigma^2\right), t = 1, 2, \cdots, T. \end{aligned} \tag{11}$$

Appendix 3. SSAEPD

If a random variable X is distributed as AEPD, we denote $X \sim AEPD\left(\mu, \sigma, \alpha, p_1, p_2\right)$ If a random variable X is distributed as standard AEPD, we denote $X \sim SAEPD\left(\mu = 0, \sigma = 1, \alpha, p_1, p_2\right)$ or in short, $X \sim SAEPD\left(\alpha, p_1, p_2\right)$ If a random variable Z is distributed as standardized standard AEPD, we denote $Z \sim SSAEPD\left(\mu = 0, \sigma = 1, \alpha, p_1, p_2\right)$ or $Z \sim SSAEPD\left(\alpha, p_1, p_2\right)$ with mean zero and the variance 1. That is, $E\left(Z\right) = 0$, $Var\left(Z\right) = 1$ The brief history of SSAEPD is listed in **Table 9**.

The probability density function (PDF) of the Standardized Standard AEPD (SSAEPD) proposed by Zhu and Zinde-Walsh (2009)[10] is

$$f\left(z_t \mid \beta\right) = \begin{cases} \delta\left(\dfrac{\alpha}{\alpha^*}\right) K\left(p_1\right) \exp\left(-\dfrac{1}{p_1}\left|\dfrac{w + z_t\delta}{2\alpha^*}\right|^{p_1}\right), & \text{if } z_t \le -\dfrac{w}{\delta}, \\ \delta\left(\dfrac{1-\alpha}{1-\alpha^*}\right) K\left(p_2\right) \exp\left(-\dfrac{1}{p_2}\left|\dfrac{w + z_t\delta}{2\left(1-\alpha^*\right)}\right|^{p_2}\right), & \text{if } z_t > -\dfrac{w}{\delta}. \end{cases} \tag{12}$$

[10]For more information about SSAEPD, one can refer to **Appendix 3**.

Table 8. Variable definitions for 48 industries.

Variable	Name	SIC codes
Agric	Agriculture	0100-0199, 0200-0299, 0700-0799, 0910-0919, 2048-2048
Food	Food Products	2000-2046, 2050-2063, 2070-2079, 2090-2092, 2095-2095, 2098-2099
Soda	Candy & Soda	2064-2068, 2086-2087, 2096-2097
Beer	Beer & Liquor	2080-2080, 2082-2085
Smoke	Tobacco Products	2100-2199
Toys	Recreation	0920-0999, 3650-3652, 3732-3732, 3930-3931, 3940-3949
Fun	Entertainment	7800-7833, 7840-7841, 7900-7900, 7910-7911, 7920-7933, 7940-4949, 7980-7980, 7990-7999
Books	Printing and Publishing	2700-2749, 2770-2771, 2780-2799
Hshld	Consumer Goods	2047-2047, 2391-2392, 2510-2519, 2590-2599, 2840-2844, 3160-3161, 3170-3172, 3190-3199,
		3229-3229, 3260-3260, 3262-3263, 3269-3269, 3230-3231, 3630-3639, 3750-3751,
		3800-3800-3860-3861, 3870-3873, 3910-3911, 3914-3915, 3960-3962, 3991-3991, 3995-3995
Clths	Apparel	2300-2390, 3020-3021, 3100-3111, 3130-3131, 3140-3151, 3963-3965
MedEq	Medical Equipment	3693-3693, 3840-3851
Drugs	Pharmaceutical Products	2830-2831, 2833-2836
Chems	Chemicals	2800-2829, 2850-2879, 2890-2899
Rubbr	Rubber and Plastic Products	3031-3031, 3041-3041, 3050-3053, 3060-3099
Txtls	Textiles	2200-2284, 2290-2295, 2297-2299, 2393-2395, 2397-2399
BldMt	Construction Materials	0800-0899, 2400-2439, 2450-2459, 2490-2499, 2660-2661, 2950-2952, 3200-3200, 3210-3211,
		3240-3241, 3250-3259, 3261-3261, 3264-3264, 3270-3275, 3280-3281, 3290-3293, 3295-3299,
		3420-3433, 3440-3442, 3446-3448
Cnstr	Construction	1500-1511, 1520-1549, 1600-1799
Steel	Steel Works Etc	3300-3300, 3310-3317, 3320-3325, 3330-3341, 3350-3357, 3360-3379, 3390-3399
FabPr	Fabricated Products	3400-3400, 3443-3444, 3460-3479
Mach	Machinery	3510-3536, 3538-3538, 3540-3569, 3580-3582, 3585-3586, 3589-3599
ElcEq	Electrical Equipment	3600-3600, 3610-3613, 3620-3621, 3623-3629, 3640-3646, 3648-3649, 3660-3660, 3690-3692,
		3699-3699
Autos	Automobiles and Trucks	2296-2296, 2396-2396, 3010-3011, 3537-3537, 3647-3647, 3694-3694, 3700-3700, 3710-3711,
		3713-3716, 3790-3792, 3799-3799
Aero	Aircraft	3720-3721, 3723-3725, 3728-3729
Ships	Shipbuilding	2*3730-3731, 3740-3743
	Railroad Equipment	
Guns	Defense	3760-3769, 3795-3795, 3480-3489
Gold	Precious Metals	1040-1049
Mines	Non-Metallic and	2*1000-1039, 1050-1119, 1400-1499
	Industrial Metal Mining	
Coal	Coal	1200-1299

Continued

Oil	Petroleum and Natural Gas	1300-1300, 1310-1339, 1370-1382, 1389-1389, 2900-2912, 2990-2999
Variable	Name	SIC codes
Util	Utilities	4900-4900, 4910-4911, 4920-4925, 4930-4932, 4939-4942
Telcm	Communication	4800-4800, 4810-4813, 4820-4822, 4830-4841, 4880-4892, 4899-4899
PerSv	Personal Services	7020-7021, 7030-7033, 7200-7200, 7210-7212, 7214-7217, 7219-7221, 7230-7231, 7240-7241, 7250-7251, 7260-7299, 7395-7395, 7500-7500, 7520-7549, 7600-7600, 7620-7620, 7622-7623, 7629-7631, 7640-7641, 7690-7699, 8100-8499, 8600-8699, 8800-8899, 7510-7115
BusSv	Business Services	2750-2759, 3993-3993, 7218-7218, 7300-7300, 7310-7342, 7349-7353, 7359-7372, 7374-7385, 7389-7394, 7396-7397, 7399-7399, 7519-7519, 8700-8700, 8710-8713, 8720-8721, 8730-8734, 8740-8748, 8900-8911, 8920-8999, 4220-4229
Comps	Computers	3570-3579, 3680-3689, 3695-3695, 7373-7373
Chips	Electronic Equipment	3622-3622, 3661-3666, 3669, 3670-3679, 3810-3810, 3812-3812
LabEq	Measuring and Control Equipment	2*3811-3811, 3820-3820, 3821-3827, 3829-3839
Paper	Business Supplies	2520-2549, 2600-2639, 2670-2699, 2760-2761, 3950-3955
Boxes	Shipping Containers	2440-2449, 2640-2659, 3220-3221, 3410-3412
Trans	Transportation	4000-4013, 4040-4049, 4100-4100, 4110-4121, 4130-4131, 4140-4142, 4150-4151, 4170-4173, 4190-4200, 4210-4219, 4230-4231, 4240-4249, 4400-4700, 4710-4712, 4720-4749, 4780-4780, 4782-4785, 4789-4789
Whlsl	Wholesale	5000-5000, 5010-5015, 5020-5023, 5030-5060, 5063-5065, 5070-5078, 5080-5088, 5090-5094, 5099-5100, 5110-5113, 5120-5122, 5130-5172, 5180-5182, 5190-5199
Rtail	Retail	5200-5200, 5210-5231, 5250-5251, 5260-5261, 5270-5271, 5300-5300, 5310-5311, 5320-5320, 5330-5331, 5334-5334, 5340-5349, 5390-5400, 5410-5412, 5420-5469, 5490-5500, 5510-5579, 5590-5700, 5710-5722, 5730-5736, 5750-5799, 5900-5900, 5910-5912, 5920-5932, 5940-5990, 5992-5995, 5999-5999
Meals	Restaurants, Hotels, Motels	2*5800-5829, 5890-5899, 7000-7000, 7010-7019, 7040-7049, 7213-7213
Banks	Banking	6000-6000, 6010-6036, 6040-6062, 6080-6082, 6090-6100, 6110-6113, 6120-6179, 6190-6199
Insur	Insurance	6300-6300, 6310-6331, 6350-6351, 6360-6361, 6370-6379, 6390-6411
RlEst	Real Estate	6500-6500, 6510-6510, 6512-6515, 6517-6532, 6540-6541, 6550-6553, 6590-6599, 6610-6611
Fin	Trading	6200-6299, 6700-6700, 6710-6726, 6730-6733, 6740-6779, 6790-6795, 6798-6799
Hlth	Healthcare	8000-8099
Other	Almost Nothing	4950-4961, 4970-4971, 4990-4991

where

$$\alpha^{*}=\frac{\alpha K\left(p_{1}\right)}{\alpha K\left(p_{1}\right)+\left(1-\alpha\right)K\left(p_{2}\right)}, \tag{13}$$

Table 9. The History of the SSAEPD distribution.

Author(s)	Distribution(s)
De Moivre (1738)	Normal distribution
Gauss (1809)	Normal applied in astronomy
Subbotin (1923)	EPD
Aitchison and Brown (1957)	LogNormal distribution
Azzalini (1986)	SEPD
Zolotarev V.M. (1986)	Stable distribution
Bolleslev (1987)	Student–t distribution
Fernandez et al.(1995)	Modified SEPD
Swamee P.K. (2002)	Near Lognormal distribution
Ayebo and Kozubowski (2004)	SEPD applied in finance
DiCiccio and Monti (2004)	Properties of MLE of the SEPD
Zhu and Zinde-Walsh (2009)	SSAEPD

Notes: EPD = Exponential Power Distribution; SEPD = Skewed Exponential Power Distribution; SSAEPD = Standardized Standard Asymmetric Exponential Power Distribution. This table is a revision of the one in Jin (2011).

$$K(p)=\frac{1}{2p^{1/p}\Gamma(1+1/p)}, \tag{14}$$

$$\Gamma(x)=\int_0^\infty y^{x-1}\mathrm{e}^{-y}\mathrm{d}y, \tag{15}$$

$$w=\frac{1}{B}\left[(1-\alpha)^2\frac{p_2\Gamma(2/p_2)}{\Gamma^2(1/p_2)}-\alpha^2\frac{p_1\Gamma(2/p_1)}{\Gamma^2(1/p_1)}\right], \tag{16}$$

$$\begin{aligned}\delta^2=\frac{1}{B^2}\Bigg\{&(1-\alpha)^3\frac{p_2^2\Gamma(3/p_2)}{\Gamma^3(1/p_2)}+\alpha^3\frac{p_1^2\Gamma(3/p_1)}{\Gamma^3(1/p_1)}\\&-\left[(1-\alpha)^2\frac{p_2\Gamma(2/p_2)}{\Gamma^2(1/p_2)}-\alpha^2\frac{p_1\Gamma(2/p_1)}{\Gamma^2(1/p_1)}\right]^2\Bigg\},\end{aligned} \tag{17}$$

$$B=\alpha K(p_1)+(1-\alpha)K(p_2). \tag{18}$$

$\mu\in R$, $\sigma>0$, $p_1>0$, $p_2>0$, $\alpha\in(0,1)$. p_1 (or p_2) is the parameter controlling the left (or right) tail. α controls the skewness. The mean of z_t is zero and its variance is 1. When $\alpha=0.5$, $p_1=p_2=2$, SSAEPD can be reduced to Normal (0, 1).

Appendix 4. Simulation Results

We check the MatLab program written by Zhou and Li (2016) by following simulation and find out the program is valid and can be used to analyze our empirical data. The FF5-SSAEPD-GARCH (1,1) is simulated as follows.

$$R_t-R_{ft}=\beta_0+\beta_1(R_{mt}-R_{ft})+\beta_2SMB_t+\beta_3HMLO_t \tag{19}$$

$$+\beta_4 RMW_t + \beta_5 CMA_t + u_t, t = 1, 2, \cdots, T, \quad (20)$$

$$u_t = \sigma_t z_t, z_t \sim SSAEPD(\alpha, p_1, p_2),$$

$$\sigma_t^2 = a_0 + a_1 u_{t-1}^2 + b_1 \sigma_{t-1}^2.$$

The data generation process is as follows:

1) Given $\alpha = 0.5, p_1 = p_2 = 2$, generate SSAEPD random numbers $\{z_t\}_{t=1}^{T}$ [11].

2) Set $\sigma_0^2 = 1, \varepsilon_0 = 0, a = 0.3, b = 0.5, c = 0.4$, generate $\{\sigma_t^2\}_{t=1}^{T}$ and $\{u_t\}_{t=1}^{T}$ with following formula:

Table 10. Simulation results.

	β_0	β_1	β_2	β_3	β_4	β_5	α	p_1	p_2	a	b	c
T	0.2	1	0.5	0.5	0.5	0.5	0.5	2	2	0.3	0.5	0.4
E	0.2011	0.9788	0.5150	0.4780	0.5152	0.5367	0.4600	1.8670	2.1095	0.3149	0.5132	0.3841
P	5.50%	2.12%	3.00%	4.40%	3.04%	7.33%	8.00%	6.65%	5.47%	4.98%	2.65%	3.98%
T	0.2	1	0.5	0.5	0.5	0.5	0.5	2	2	0.3	0.5	0.4
E	0.2303	0.9460	0.5036	0.4594	0.5346	0.4939	0.5331	2.0996	1.8691	0.2849	0.4873	0.4224
P	15.13%	5.40%	0.71%	8.13%	6.92%	1.22%	6.62%	4.98%	6.55%	5.02%	2.55%	5.60%
T	0.1	1	0.5	0.5	0.5	0.5	0.5	2	2	0.3	0.5	0.4
E	0.0858	1.0218	0.5014	0.5054	0.5353	0.4450	0.4866	1.9291	2.0242	0.2912	0.5016	0.3990
P	14.19%	2.18%	0.29%	1.09%	7.07%	11.01%	2.69%	3.54%	1.21%	2.93%	0.32%	0.26%
T	0.2	0.8	0.5	0.5	0.5	0.5	0.5	2	2	0.3	0.5	0.4
E	0.2355	0.7871	0.4910	0.4421	0.5272	0.4759	0.4559	1.8879	2.1762	0.3213	0.4969	0.3931
P	17.74%	1.61%	1.80%	11.58%	5.44%	4.81%	8.82%	5.61%	8.81%	7.09%	0.62%	1.72%
T	0.2	1	0.8	0.5	0.5	0.5	0.5	2	2	0.3	0.5	0.4
E	0.1526	0.9918	0.8290	0.5065	0.4985	0.5511	0.4545	1.8716	2.1702	0.3161	0.5161	0.3774
P	23.68%	0.82%	3.63%	1.31%	0.30%	10.22%	9.09%	6.42%	8.51%	5.36%	3.22%	5.64%
T	0.2	1	0.5	0.8	0.5	0.5	0.5	2	2	0.3	0.5	0.4
E	0.2026	1.0015	0.4462	0.7692	0.5195	0.5291	0.5219	2.0682	1.8839	0.2874	0.5153	0.3918
P	1.29%	0.15%	10.77%	3.85%	3.89%	5.82%	4.37%	3.41%	5.81%	4.22%	3.06%	2.05%
T	0.2	1	0.5	0.5	0.8	0.5	0.5	2	2	0.3	0.5	0.4
E	0.1841	0.9958	0.4931	0.5213	0.7973	0.5330	0.5464	2.1344	1.8229	0.3198	0.5043	0.3763
P	7.93%	0.42%	1.38%	4.26%	0.33%	6.60%	9.27%	6.72%	8.85%	6.60%	0.86%	5.91%
T	0.2	1	0.5	0.5	0.5	0.8	0.5	2	2	0.3	0.5	0.4
E	0.2003	0.9743	0.5274	0.5228	0.4866	0.7987	0.5221	2.1624	2.0246	0.2943	0.4840	0.4055
P	0.13%	2.57%	5.48%	4.57%	2.69%	0.16%	4.43%	8.12%	1.23%	1.91%	3.20%	1.39%
T	0.2	1	0.5	0.5	0.5	0.5	0.5	2	2	0.3	0.6	0.3
E	0.2620	0.9962	0.4686	0.4520	0.4569	0.5049	0.5349	2.1566	1.8604	0.3024	0.6210	0.2880
P	31.01%	0.38%	6.28%	9.61%	8.63%	0.97%	6.99%	7.83%	6.98%	0.80%	3.5%	4.00%

Notes: T means the true value of parameters. E means the estimates. *P* means the error in percentage.

*note

$$\sigma_1^2 = a_0 + a_1\sigma_0^2 z_t^2 + b_1\sigma_0^2,$$

$$u_1 = z_1\sigma_1.$$

3) Generate $\{X_{1t}\}_{t=1}^T$, $\{X_{2t}\}_{t=1}^T$, $\{X_{3t}\}_{t=1}^T$, $\{X_{4t}\}_{t=1}^T$, $\{X_{5t}\}_{t=1}^T$ from Uniform (0,1)[12].

4) Set $\beta_0 = 0.2, \beta_1 = 1, \beta_2 = 0.5, \beta_3 = 0.5, \beta_4 = 0.5, \beta_5 = 0.5$, and we can get $\{Y_t\}_{t=1}^T$.

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + u_t, t = 1, 2, \cdots, T.$$

After getting the simulated data $\{X_{1t}, X_{2t}, X_{3t}, X_{4t}, X_{5t} Y_t\}_{t=1}^T$, we can use them to estimate the parameters in the FF5-SSAEPD-GARCH model. The simulation results are reported in **Table 10**, almost all the estimates are close to the true values of the parameters. Hence, we can draw the conclusion that this MatLab program is valid from empirical analysis.

[12]For simplicity, we use Xs to represent the 5 factors in simulation.