

Financial Time Series Analysis by Using MATLAB

Xinyuan Zheng

Correspondence: Xinyuan Zheng, 13 Fonti Street, Eastwood, NSW, Sydney. Tel: 61-41-661-8518.

Received: November 6, 2019

Accepted: December 6, 2019

Online Published: December 14, 2019

doi:10.20849/abr.v4i3.687

URL: <https://doi.org/10.20849/abr.v4i3.687>

Abstract

This report contains two parts. For part A, performing a Principle Components Analysis (PCA) and analyzing the drivers. Then, carrying out factor analyses and comparing them. For part B, employing 5 different quantitative models to forecast and generate moving origin horizon one forecasts of both return and volatility. Then, figuring out the optimal weights for the portfolio and assigning the optimal portfolio. Finally, comparing the returns and risk measure from all portfolio and models.

Keywords: MATLAB application, factor modelling, principal component analysis, dynamic portfolio optimization, data processing and exploratory data analysis, forecasting model building

Part A – Factor Modelling

1. Principal Component Analysis

1.1 Original Data Description and Processing

The data include the historical daily index prices for the S&P ASX200 Index and the 11 Global Industry Classification Standard (GICS) Industry sectors during the five years from 6th April 2014 to 6th April 2019. These 11 sectors consist of more than 43000 global companies from 24 industry groups (“GICS”, 2019). The whole data set contains 2 null rows, 27/12/15 and 04/09/17. To make the statistics complete, the 2 null rows could be deleted, which only occupy 0.15% of the whole data. This manipulation is reasonable because the nulls mean no transactions occurred on those days.

1.2 Daily Percentage Return of GICS

The daily percentage returns (Simple Returns) equal to $(Y_t - Y_{t-1}) / Y_{t-1} \times 100\%$, where Y_t represents the adjusted closing price of today, Y_{t-1} represents yesterday's adjusted closing price.

1.3 Principal Component Analysis

Principal Component Analysis (PCA), a statistical approach employed in a multivariate system to extract well-arranged and uncorrelated components of variation (Jolliffe, 1986). This means that each component does not have any correlation, as shown in the Appendix table ‘Principal Components (Correlation Coefficient)’, each pairwise correlation coefficient is zero like PC1 and the rest of PCs. However, each industry sectors should relate to each other. Otherwise, it might be difficult to find a PC that would drive them jointly under linear relationship. PCA aims to use the least dimensions to explain the most variations (Barber & Copper, 2012).

To perform a PCA, the first step is to estimate the correlation matrix of this finite sample. Table 1 shows all pairwise correlations. None of the pairwise correlations are exactly equal to one or zero which represents no clearly perfect multicollinearity or perfectly independent variables. The pairwise correlations are generally quite strong with the smallest one 0.2554 between XTJ and XMJ ranging up to 0.7055 between XDJ and XNJ. Therefore, the other statistics are necessary for further analyze, like p-value and confidence interval, to prove a more precise estimation.

Table 1. Correlation matrix

Industry	XPJ	XDJ	XSJ	XEJ	XXJ	XHJ	XNJ	XIJ	XMJ	XTJ	XUJ
XPJ	1.0000	0.5368	0.4544	0.2959	0.4978	0.4311	0.5847	0.3293	0.2991	0.3634	0.5324
XDJ	0.5368	1.0000	0.6277	0.4866	0.6645	0.5962	0.7055	0.6087	0.4860	0.4145	0.4889
XSJ	0.4544	0.6277	1.0000	0.4466	0.5892	0.4991	0.5750	0.4842	0.4538	0.3800	0.4092
XEJ	0.2959	0.4866	0.4466	1.0000	0.5276	0.3428	0.4540	0.4161	0.6818	0.2760	0.3397
XXJ	0.4978	0.6645	0.5892	0.5276	1.0000	0.5343	0.6215	0.5537	0.5447	0.3834	0.4288
XHJ	0.4311	0.5962	0.4991	0.3428	0.5343	1.0000	0.6224	0.5496	0.3389	0.2855	0.4027
XNJ	0.5847	0.7055	0.5750	0.4540	0.6215	0.6224	1.0000	0.5969	0.4848	0.3985	0.5288
XIJ	0.3293	0.6087	0.4842	0.4161	0.5537	0.5496	0.5969	1.0000	0.4384	0.3461	0.3053
XMJ	0.2991	0.4860	0.4538	0.6818	0.5447	0.3389	0.4848	0.4384	1.0000	0.2554	0.3241
XTJ	0.3634	0.4145	0.3800	0.2760	0.3834	0.2855	0.3985	0.3461	0.2554	1.0000	0.3344
XUJ	0.5324	0.4889	0.4092	0.3397	0.4288	0.4027	0.5288	0.3053	0.3241	0.3344	1.0000

The p-values as shown in the appendix table named ‘P-Value (Correlation Coefficients)’ of these correlation coefficients is less than 1%. Therefore, the correlation coefficient is statistically significant not zero at 95% confidence level, even at 99% confidence level to prove the linear relationship among these 11 industry sectors. Additionally, from the tables named ‘Upper Bound (Correlation Coefficients)’ and ‘Lower Bound (Correlation Coefficients)’ show all the confidence level are between zero and one. All the evidences mean these indices are related but would not cause multicollinearity. Thus, the data is suitable for PCA because all simple returns are positively linear associated but not colinear.

Table 2. PCA

Industry	1st PC Weight s	2nd PC Weight s	3rd PC Weight s	4th PC Weight s	5th PC Weight s	6th PC Weight s	7th PC Weight s	8th PC Weight s	9th PC Weight s	10th PC Weight s	11th PC Weight s
XPJ	0.2151	-0.2581	0.0760	0.4233	0.0560	-0.0718	-0.2456	-0.5907	-0.4401	-0.2440	-0.1771
XDJ	0.2862	-0.1651	-0.1067	0.0205	-0.0012	-0.1174	-0.1242	0.0358	-0.0107	0.8223	-0.4150
XSJ	0.2640	-0.1250	-0.0455	0.0353	0.0875	-0.5172	-0.2166	0.6554	-0.3140	-0.2445	0.0363
XEJ	0.4633	0.6476	0.1682	0.0491	-0.5713	-0.0301	-0.0526	-0.0486	-0.0488	-0.0141	0.0161
XXJ	0.3149	-0.0779	-0.0916	0.0158	0.1399	-0.3459	-0.2440	-0.2173	0.7784	-0.1774	0.0525
XHJ	0.2675	-0.2661	-0.3541	0.0030	-0.2475	-0.1470	0.7784	-0.0591	0.0028	-0.1564	-0.1243
XNJ	0.2512	-0.1667	-0.1037	0.0810	0.0233	0.0407	0.0201	-0.1001	-0.1268	0.3160	0.8736
XIJ	0.3250	-0.1776	-0.3787	-0.5640	-0.0806	0.4776	-0.3306	-0.0030	-0.1016	-0.2041	-0.0738
XMJ	0.3707	0.4270	-0.0103	-0.0243	0.7569	0.1412	0.2657	-0.0215	-0.1034	-0.0175	-0.0681
XTJ	0.2428	-0.3326	0.8048	-0.3910	0.0079	0.0050	0.1698	-0.0210	0.0170	-0.0106	-0.0020
XUJ	0.2271	-0.2047	0.1292	0.5817	-0.0448	0.5671	0.0052	0.3951	0.2489	-0.0795	-0.0582
Percentage % Variance	50.8914	12.8977	7.8729	6.6489	4.6697	3.9576	3.5145	3.1412	2.7926	2.0700	1.5435
Cumulative % Variance	50.8914	63.7891	71.6620	78.3109	82.9806	86.9382	90.4527	93.5939	96.3865	98.4565	100.0000

From the above PCA components Table 2, the first principal component can explain about 50.89% of the variation. The cumulatively explained percentage variance are 63.79%, 71.66%, 78.31%, 82.98%, 86.94%, 90.45% respectively for the second to seventh principal component. Generally, the principal components, which explain 70%-90% (cumulative) of the variation, are preferable to choose. Here the first six principal components are chosen because these six principal components have already explained 86.94% of the variation. Though the

seventh principal component increases the explained variation to 90.45%, it does not achieve the aim of dimensionality reduction significantly from the original 11 components. Due to the seventh principal component only have about 3.51% explanatory power. Thus, six principal components are enough to explain about 87% variation in the original series and is probably the best choice.

2. First Four PCs Analysis

2.1 Equations

The equations for the first four principal components are:

$$W_{1t} = 0.2151 \times XPJ_t + 0.2862 \times XDJ_t + 0.2640 \times XSJ_t + 0.4633 \times XEJ_t + 0.3149 \times XXJ_t + 0.2675 \times XHJ_t + 0.2512 \times XNJ_t + 0.3250 \times XIJ_t + 0.3707 \times XMJ_t + 0.2428 \times XTJ_t + 0.2271 \times XUJ_t$$

$$W_{2t} = -0.2581 \times XPJ_t - 0.1651 \times XDJ_t - 0.1250 \times XSJ_t + 0.6476 \times XEJ_t - 0.0779 \times XXJ_t - 0.2661 \times XHJ_t - 0.1667 \times XNJ_t - 0.1776 \times XIJ_t + 0.4270 \times XMJ_t - 0.3326 \times XTJ_t - 0.2047 \times XUJ_t$$

$$W_{3t} = 0.0760 \times XPJ_t - 0.1067 \times XDJ_t - 0.0455 \times XSJ_t + 0.1682 \times XEJ_t - 0.0916 \times XXJ_t - 0.3541 \times XHJ_t - 0.1037 \times XNJ_t - 0.3787 \times XIJ_t - 0.0103 \times XMJ_t + 0.8048 \times XTJ_t + 0.1292 \times XUJ_t$$

$$W_{4t} = 0.4233 \times XPJ_t + 0.0205 \times XDJ_t + 0.0353 \times XSJ_t + 0.0491 \times XEJ_t + 0.0158 \times XXJ_t + 0.0030 \times XHJ_t + 0.0810 \times XNJ_t - 0.5640 \times XIJ_t - 0.0243 \times XMJ_t - 0.3910 \times XTJ_t + 0.5817 \times XUJ_t$$

2.2 First Four PCs Analysis

2.2.1 The 1st Principal Component (PC1) – Market

PC1 is asserted as market-wide effect due to it is positively correlated with all of 11 industry sector's simple returns. The magnitude of the coefficients is quite similar except for XEJ, XMJ and XIJ with coefficients 0.4633, 0.3707 and 0.3250 respectively. The rest of indices is in the range between 0.2 to 0.3 from 0.2151 to 0.2862. PC1 distinguishes XEJ and XMJ from the others as shown in the Appendix Figure 1. The correlation between PC1 and the mean of all industry sectors is 0.9936. The correlation between PC1 and the market return is 0.9606. The visualize plots of two correlation are showed as followed. All the signs indicate that PC1 is market.

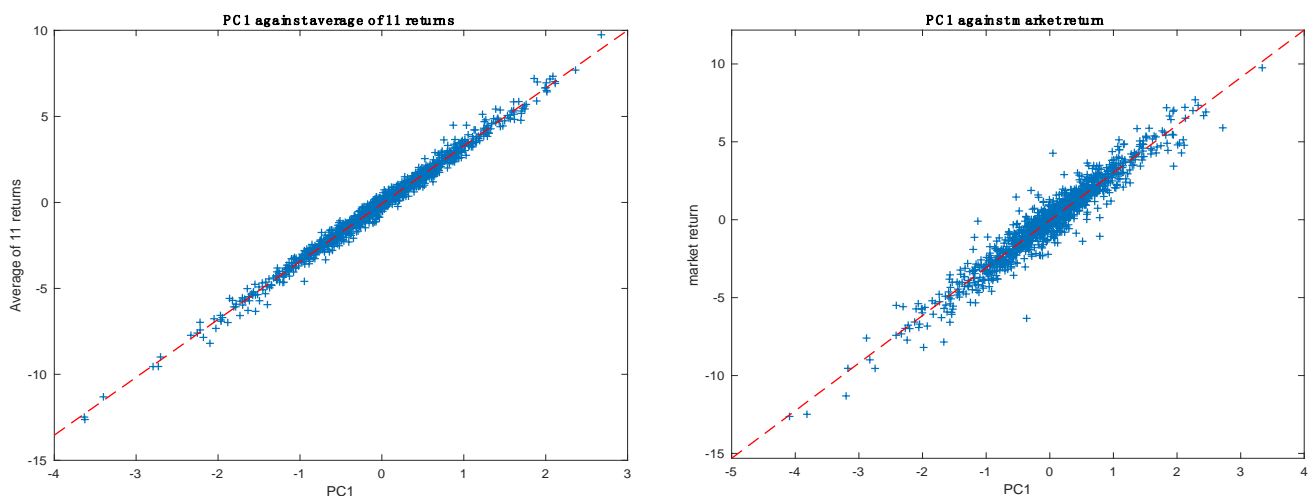


Figure 1. PCA 1 against average of 11 returns and market return

2.2.2 The 2nd Principal Component – Commodity Price

PC2 might be asserted as commodity price. It is positively linear associated with XEJ (0.6476) and XMJ (0.4270) but negatively linear associated with the rest from XTJ (-0.3326) to XXJ, which further distinguishes XEJ (-0.0779), XMJ from the others as shown in the appendix Figure 2. The commodity price especially crude oil

price, metal price and natural gas price, which may affect the mining and energy companies (Beattie, 2018). When prices increase, the production costs also increase for manufacture factories due to oil price, one of the most important manufacture materials for most daily products. However, for mining and energy companies, their production costs are relatively stable (Maverick, 2015). Hence, when the price is high, the share price increases as future cash flow increases (Resource and Mining Stocks, 2019). For XPJ (-0.2581) and XUJ (-0.2047) with the negative relation, as the commodity price increases, individuals decrease utility usage and rental expense with fixed income. XHJ (-0.2661), similar with XPJ and XUJ, normal families pay less attention on health care.

2.2.3 The 3rd Principal Component – 5G and NBN

PC3 might be asserted as Internet like NBN and 5G. It is strongly positive on XTJ (0.8048) and negative on XHJ (-0.3541) and XIJ (-0.3787). So PC3 distinguishes XTJ, XHJ and XIJ from the rest industry sectors. The remaining have weak or near zero eigenvalues like XPJ (0.0760) and XSJ (-0.0455). The policy of carrying out NBN has boosted the communication services, which NBN is the high-speed Internet fiber network (Purvis, 2013). Recently, Telstra has announced that they will sell Australian's first 5G device (The West Australian, 2019). Baillieu analysts Nick Burgess also states that Telstra's 5G strategy might be a potential opportunity from the view of the Telstra investment case (Livewire, 2019). Therefore, the high speed of internet and the coming 5G mobile era have stimulated the communication industry sector index. However, the NBN and 5G do not stimulate XIJ, the information technology sector. It mainly because the increasing internet speed does not connect to the cloud computing, big data and mobile computing that the primary functions in the information technology sector (Compare Sector Characteristics, 2019).

2.2.4 The 4th Principal Component – Exchange Rate

PC4 might be asserted as exchange rate. It is relatively strongly positive with XUJ (0.5817) and XPJ (0.4233). XIJ, XMJ and XTJ with the negative loadings of -0.5640, -0.0243 and -0.3910. The rest are around zero. Thus, PC4 distinguishes XUJ and XPJ from the remaining sectors. Under the surge of studying abroad, Australia has become one of the most popular choice for international students (Koprowski, 2016). Koprowski stated that the number of international students rises sharply to 500,000 (2016). This directly make Australia become the fourth largest country that export education (2016). The national education department announced that in 2025, there will be more than one million international students (AIE2025, 2019). Furthermore, the exchange rates decrease means the depreciation of the AUD in terms of other currencies. It might be much less expensive for the parents. When the population increases, the rental return and gas, electricity and water consumption increase. The depreciation of the AUD which might also mean the government prefer to stimulate the economics. Therefore, it is slightly positive related to the rest of industries. However, XIJ and XTJ, the communication and information technology sector do not boost by the depreciation of AUD, mainly because they are sensitive to economic cycles (Compare Sector Characteristics, 2019). The weak exchange rate means the country is trying to boost the economy.

3. Factor Modelling

The main goal of factor modelling aims to find the least number of factors explaining most variation.

	m = 1	m = 2	m = 3	m = 4	m = 5
P-value	3.6E-132	2.74E-39	7.48E-07	0.0208	NA
Chisq	764.374	273.917	74.7211	30.8537	NA
Dfe	44	34	25	17	10
loglike	-0.6077	-0.2179	-0.0595	-0.0246	-0.0112

The overall R^2 is 45%. The model is moderate strength fit to the data. The single factor captures from about 24% (XTJ) to 72% (XDJ) of the variation in each industry sectors' returns (Adjusted R^2). Furthermore, the chi-squared test statistics is 764.374 with a p-value of 0. Thus, the null hypothesis that the number of factors equals to 1 is rejected. Additionally, the SER are between 0.4541 (XNJ) to 1.2209 (XEJ). They are significantly large for the daily return prediction error. Hence, one factor model is not good enough.

m=1					m=2					
Industry	Factor Loadings	Specific Variance	SER	R ² adjusted	Industry	1st Factor Loadings	2nd Factor Loadings	Specific Variance	SER	R ² adjusted
XPJ	0.5812	0.5048	0.7105	0.4009	XPJ	0.5987	0.1385	0.4649	0.6819	0.4482
XDJ	0.7700	0.2348	0.4846	0.7163	XDJ	0.6955	0.3310	0.2344	0.4842	0.7168
XSJ	0.6783	0.4183	0.6468	0.5238	XSJ	0.5800	0.3443	0.4235	0.6507	0.5179
XEJ	0.9255	1.4905	1.2209	0.3650	XEJ	0.4220	1.1562	0.8323	0.9123	0.6454
XXJ	0.7956	0.3868	0.6219	0.6207	XXJ	0.6395	0.4750	0.3851	0.6206	0.6224
XHJ	0.7122	0.5454	0.7385	0.4819	XHJ	0.7035	0.2122	0.5127	0.7160	0.5129
XNJ	0.6806	0.2062	0.4541	0.6919	XNJ	0.6379	0.2665	0.1916	0.4377	0.7139
XIJ	0.8080	0.7026	0.8382	0.4816	XIJ	0.6919	0.4085	0.7097	0.8424	0.4764
XMJ	0.7789	0.9981	0.9990	0.3781	XMJ	0.3359	1.0185	0.4545	0.6742	0.7168
XTJ	0.5730	1.0504	1.0249	0.2381	XTJ	0.5350	0.2167	1.0456	1.0226	0.2417
XUJ	0.5788	0.6374	0.7984	0.3445	XUJ	0.5558	0.2055	0.6212	0.7882	0.3611
Percentage (%)					Percentage (%)					
Variance 44.58770187					Variance 54.6259					

The overall R² for two factors is 10% over than one factor model. The adjusted R² for each industry sectors' return has increased especially XEJ and XMJ, XSJ and XIJ with small decrease. The chi-squared test statistics is 273.917 with a p-value of 0. Thus, the null hypothesis that the number of factors equal to 2 is strongly rejected. The SER has lightly decreased to the range of 0.4377 (XNJ) to 1.0226 (XTJ).

m=3							m=4							
Industry	1st Factor Loadings	2nd Factor Loadings	3rd Factor Loadings	Specific Variance	SER	R ² adjusted	Industry	1st Factor Loadings	2nd Factor Loadings	3rd Factor Loadings	4th Factor Loadings	Specific Variance	SER	R ² adjusted
XPJ	0.2247	0.6831	0.1163	0.3119	0.5585	0.6298	XPJ	0.2170	0.6690	0.1200	0.0935	0.3247	0.5698	0.6146
XDJ	0.5688	0.4263	0.2937	0.2361	0.4860	0.7147	XDJ	0.5458	0.4157	0.2833	0.2189	0.2288	0.4784	0.7235
XSJ	0.4552	0.3783	0.3165	0.4279	0.6542	0.5128	XSJ	0.4199	0.3626	0.3028	0.2906	0.3944	0.6281	0.5510
XEJ	0.3617	0.3055	1.1515	0.7969	0.8927	0.6605	XEJ	0.3705	0.3070	1.0633	0.2045	0.9431	0.9711	0.5982
XXJ	0.5195	0.4120	0.4400	0.3866	0.6218	0.6209	XXJ	0.4872	0.3956	0.4267	0.2811	0.3649	0.6040	0.6422
XHJ	0.6424	0.3642	0.1598	0.4817	0.6940	0.5424	XHJ	0.6427	0.3643	0.1624	0.0603	0.4768	0.6905	0.5470
XNJ	0.4801	0.4370	0.2344	0.1931	0.4394	0.7116	XNJ	0.4994	0.4542	0.2370	-0.0236	0.1571	0.3964	0.7653
XIJ	0.8285	0.1859	0.3364	0.5213	0.7220	0.6154	XIJ	0.8167	0.1842	0.3382	0.1200	0.5257	0.7251	0.6121
XMJ	0.3206	0.2322	0.9857	0.4765	0.6903	0.7031	XMJ	0.2904	0.2246	1.0751	0.0235	0.3136	0.5600	0.8046
XTJ	0.3487	0.4215	0.2000	1.0396	1.0196	0.2461	XTJ	0.3144	0.4098	0.1830	0.2547	1.0137	1.0068	0.2648
XUJ	0.2141	0.6075	0.2009	0.5171	0.7191	0.4682	XUJ	0.2099	0.6133	0.1956	0.0471	0.5118	0.7154	0.4737
Percentage (%)							Percentage (%)							
Variance 58.3850							Variance 59.4204							

The overall R² for three factor model and four factor model is about 58% and 59% respectively with only 3% and 4% increase from two factor model. The chi-squared for these two models are 74.7211 and 30.8527. Their p-values are 0 and 0.0208. Both are rejected at 95% confidence level. This means that three factor model and four factor model are not good enough. Though XPJ, XIJ and XMJ for their individually adjusted R² has increased, XTJ only have 26.48%, which is too low.

Industry	1st Factor Loadings	2nd Factor Loadings	3rd Factor Loadings	4th Factor Loadings	5th Factor Loadings	Specific Variance	SER	R ² adjusted
XPJ	0.2305	0.6503	0.1186	0.1330	0.0683	0.3302	0.5746	0.6081
XDJ	0.5610	0.3969	0.2832	0.1491	0.1549	0.2290	0.4785	0.7233
XSJ	0.4399	0.3415	0.3053	0.1590	0.2345	0.3948	0.6283	0.5506
XEJ	0.3836	0.2864	1.0754	0.1545	0.1159	0.9241	0.9613	0.6063
XXJ	0.5094	0.3750	0.4297	0.1536	0.2261	0.3604	0.6003	0.6466
XHJ	0.6477	0.3595	0.1570	0.0725	0.0252	0.4733	0.6880	0.5504
XNJ	0.4939	0.4421	0.2310	0.1116	-0.0744	0.1588	0.3985	0.7628
XIJ	0.8280	0.1458	0.3307	0.1821	-0.0094	0.5058	0.7112	0.6268
XMJ	0.3001	0.2167	1.0571	0.0897	-0.0298	0.3414	0.5843	0.7873
XTJ	0.2217	0.2665	0.1487	1.1084	0.0339	0.0069	0.0830	0.9950
XUJ	0.2140	0.5988	0.1942	0.1363	0.0074	0.5117	0.7153	0.4738
Percentage (%)								
Variance	67.2845							

m = 5

Industry	1st Factor Loadings	2nd Factor Loadings	3rd Factor Loadings	4th Factor Loadings	5th Factor Loadings	6th Factor Loadings	Specific Variance	SER	R ² adjusted
XPJ	0.1786	0.6359	0.1071	0.1523	0.1561	0.0926	0.3388	0.5820	0.5979
XDJ	0.4787	0.3967	0.2687	0.1980	0.2004	0.2178	0.2422	0.4921	0.7074
XSJ	0.3339	0.3089	0.2782	0.1830	0.2112	0.4568	0.3074	0.5545	0.6500
XEJ	0.3130	0.2805	1.0458	0.2326	0.2066	0.1727	0.9500	0.9747	0.5952
XXJ	0.3537	0.3161	0.3592	0.7824	0.1730	0.1364	0.0051	0.0714	0.9950
XHJ	0.6110	0.3785	0.1496	0.1737	0.0270	0.1721	0.4530	0.6731	0.5696
XNJ	0.4512	0.4537	0.2328	0.1228	0.1086	0.0883	0.1712	0.4138	0.7443
XIJ	0.8258	0.1413	0.3204	0.1803	0.2908	0.0566	0.4305	0.6561	0.6824
XMJ	0.2587	0.2157	1.0616	0.1640	0.0916	0.0852	0.3218	0.5673	0.7995
XTJ	0.2195	0.3494	0.1578	1.256	0.5586	0.1098	0.8437	0.9185	0.3881
XUJ	0.1718	0.6059	0.1893	0.0891	0.1516	0.0791	0.5028	0.7091	0.4830
Percentage (%)									
Variance	64.7345								

m = 6

For the 5-factor model, the overall R² has increased to 67.28% and the individually adjusted R² is from 55% to almost 100% like XTJ. The p-value on the hypothesis that m = 5 cannot be calculated, since one of the specific variances is too close to zero. For the 6-factor model, the overall R² decreases from about 67% to about 65% when m = 5 (See Appendix). For the 7-factor model, due to the M is too large for the number of the observed variables. Therefore, both the 6-factor model and the 7-factor model is not suitable for these 11 industry sectors' returns. Combine all the situation, the 5-factor model might be the best model. The equations are showed as followed.

$$\begin{aligned}
 R_{XPJ,t} &= 0.0361 + 0.2305F_{1t} + 0.6503F_{2t} + 0.1186F_{3t} + 0.1330F_{4t} + 0.0683F_{5t} + \varepsilon_{1,t} \\
 R_{XDJ,t} &= 0.0244 + 0.5610F_{1t} + 0.3969F_{2t} + 0.2832F_{3t} + 0.1491F_{4t} + 0.1549F_{5t} + \varepsilon_{2,t} \\
 R_{XSJ,t} &= 0.0116 + 0.4399F_{1t} + 0.3415F_{2t} + 0.3053F_{3t} + 0.1590F_{4t} + 0.2345F_{5t} + \varepsilon_{3,t} \\
 R_{XEJ,t} &= -0.0062 + 0.3836F_{1t} + 0.2864F_{2t} + 1.0754F_{3t} + 0.1545F_{4t} + 0.1159F_{5t} + \varepsilon_{4,t} \\
 R_{XXJ,t} &= -0.0026 + 0.5094F_{1t} + 0.3750F_{2t} + 0.4297F_{3t} + 0.1536F_{4t} + 0.2261F_{5t} + \varepsilon_{5,t} \\
 R_{XHJ,t} &= 0.0640 + 0.6477F_{1t} + 0.3595F_{2t} + 0.1570F_{3t} + 0.0725F_{4t} + 0.0252F_{5t} + \varepsilon_{6,t} \\
 R_{XNJ,t} &= 0.0377 + 0.4939F_{1t} + 0.4421F_{2t} + 0.2310F_{3t} + 0.1116F_{4t} - 0.0744F_{5t} + \varepsilon_{7,t} \\
 R_{XIJ,t} &= 0.0470 + 0.8280F_{1t} + 0.1458F_{2t} + 0.3307F_{3t} + 0.1821F_{4t} - 0.0094F_{5t} + \varepsilon_{8,t} \\
 R_{XMJ,t} &= 0.0289 + 0.3001F_{1t} + 0.2167F_{2t} + 1.0571F_{3t} + 0.0897F_{4t} - 0.0298F_{5t} + \varepsilon_{9,t} \\
 R_{XTJ,t} &= -0.0255 + 0.2217F_{1t} + 0.2665F_{2t} + 0.1487F_{3t} + 1.1084F_{4t} + 0.0339F_{5t} + \varepsilon_{10,t} \\
 R_{XUJ,t} &= 0.0333 + 0.2140F_{1t} + 0.5988F_{2t} + 0.1942F_{3t} + 0.1363F_{4t} + 0.0074F_{5t} + \varepsilon_{11,t}
 \end{aligned}$$

Industry	FM = 5 (Method 1)		FM = 5 (Method 2)	
	R ² adjusted	SER	R ² adjusted	SER
XPJ	60.8099	0.5746	68.5077	0.2653
XDJ	72.3318	0.4785	72.0624	0.2312
XSJ	55.0564	0.6283	55.9861	0.3866
XEJ	60.6273	0.9613	99.8438	0.0037
XXJ	64.6634	0.6003	66.9484	0.3371
XHJ	55.0360	0.6880	71.5832	0.2991
XNJ	76.2815	0.3985	71.4734	0.1910
XIJ	62.6813	0.7112	86.4715	0.1834
XMJ	78.7274	0.5843	97.0349	0.0476
XTJ	99.5000	0.0830	99.0209	0.0135
XUJ	47.3785	0.7153	73.8663	0.2541
Overall R ²	67.2845		77.4362	

4. Comparing Method 1 and Method 2 of Factor Modelling

Through re-estimating a factor model when m = 5 by applying the second method, exploiting a PCA analysis decomposition. The first five principal components are from question 2. By comparing the method 1 and method 2, the above figure shows that method 2 is a better one.

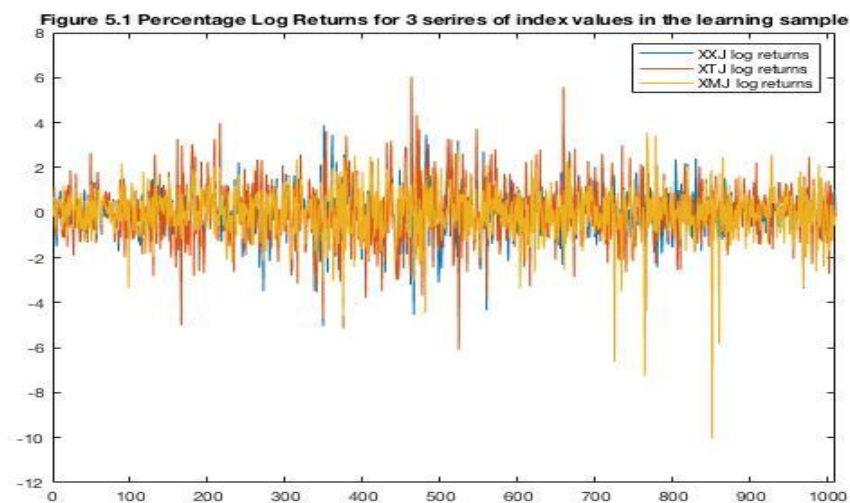
The adjusted R^2 for XPJ has increased from 60.81% to 68.51%. XDJ's adjusted R^2 has decreased slightly from 72.33% to 72.06%. For XSJ, which almost stay at the same level, between 55% to 56%. XEJ's adjusted R^2 has increased significantly from 60.62% to 99.84%, which can be explained totally. XXJ has increased slightly, about 2.3% to 66.95%. XHJ has increased over 16% to 71.58%, which is quite dramatically. However, XNJ has dropped a little to 71.47%. Both XIJ and XMJ has improved the adjusted R^2 considerably to 86.47% and 97.03% respectively. Particularly, the XMJ, whose variance has already could be explained over 90%. XTJ has declined slightly, while its adjusted R^2 still stay at 99%. It could be asserted as completely explained. The last one, XUJ, has raised from 47.38% to 73.87%, from the poor level to good one.

Due to the increasing of the adjusted R^2 , all the SERs have decreased. Especially for those adjusted R^2 have increased significantly. In addition, the overall R^2 has increased from 67.28% to 77.44%, which means the overall variation could be explained increased about 10%.

Part B – Dynamic Portfolio Optimisation

5. Data Processing and Exploratory Data Analysis (EDA)

To build the suitable time-series forecasting models, analyzing the percentage log return of XXJ, XMJ and XTJ Index is indispensable. Then converting the data to percentage log returns ($100 \times \log(\frac{P_t}{P_{t-1}})$) and splitting the data sets into two parts: the in-sample and the forecast sample. Since the forecasting model build by using on the time series data from 7/4/2014 to the 6/4/2018, this section statistical analysis focuses on data in this learning period. The plot below provides the visualized movements of the log returns. All of three series percentage log returns stay around 0, which may imply these three assets follow the mean stationary. However, the variances seem not stationary based on some significant fluctuations existing during the period.



To further statistical analysis, the summary statistic table of the three indices is showed below:

Table 2. The summary statistic table

	Mean	Median	Std	Min	Max	Skewness	Kurtosis
XXJ	-0.0050	0.0534	1.0332	-5.0415	4.2102	-0.2886	5.0831
XMJ	0.0078	-0.0341	1.3226	-6.0939	6.0415	0.0070	4.4867
XTJ	-0.0440	0.0051	1.1201	-10.0600	3.5724	-1.4395	12.7851

The means of the three indices are close to 0 and satisfy the assumption by the figure 5.1 and the standard deviation of the three indices are larger than 1. Moreover, the XXJ index and XTJ index have the negative

skewness and the skewness of the XMJ is positive. Meanwhile, all the kurtosis is larger than 3, which indicates they have fat tails. According to the histograms of the three indices below, the outliers exists in the three indices. These signs imply they are not followed the normal distribution and can be assumed as follow student-t distribution.

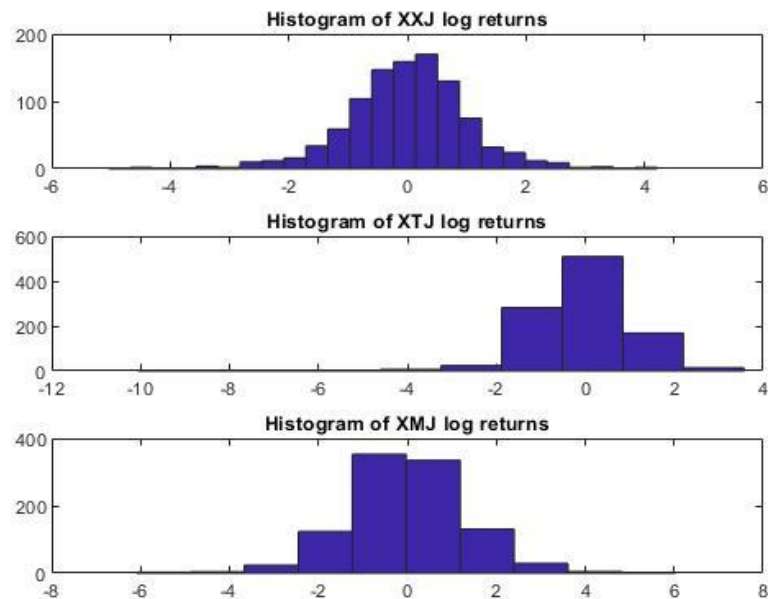


Figure 2. Histograms of the three indices

6. Forecasting Model Building

6.1 Model Selection and Motivation

According to the statistics analysis in the previous section, the three indices percentage log returns follow the mean stationary. Furthermore, the autocorrelation effects of the three indices do not exist. These imply that the constant mean equations can be applied to these indices. In terms of the variance, the ARCH, GARCH, GJR-ARCH, GJR-GARCH models are chosen to represent the assets' volatilities.

- ARCH model can describe the lag residual error to predict the variance of innovation change.
- GARCH model import the number of lag variance errors to solve the time-dependent variance issue.
- GJR-ARCH and GJR-GARCH model can be applied to the asymmetric statistical data by employing the leverage term.
- Ad-hoc model extracts a certain number of previous time series data and calculates their expected value as the following forecast return and variance.

Also, Ad-hoc model can be set as the benchmark model for comparison. Considering the asset distributions, the model with the student-t distribution need to be utilized.

6.2 Model Analysis

6.2.1 Model Assumption Test

To choose the suitable forecasting models with the appropriate mean and volatility equation, the distribution of indices log returns requires further verification, which is corresponding to the previous model assumption. The autocorrelation effects of the three indices are tested by the Autoregression Function (ACF) plot and Ljung-Box (LB) test.

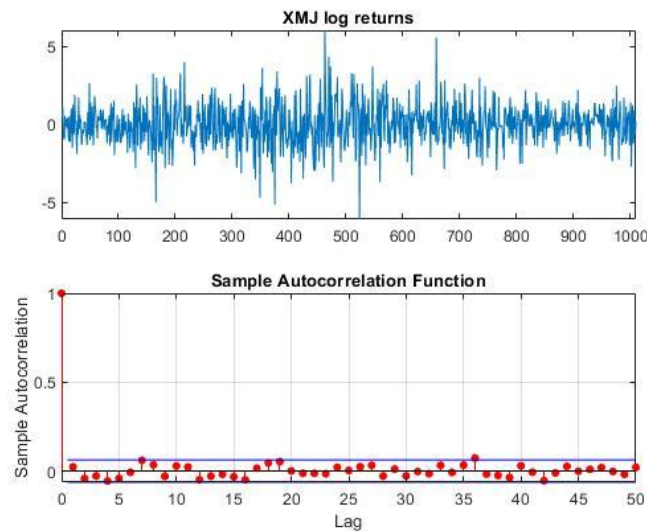


Figure 3. XMJ log return and ACF plot

Firstly, the XMJ log return at the in-sample data seems follow the mean stationary. The ACF plot above shows that the correlation dies down to 0 quickly at lag 1, and all correlations presented within the confidence interval indicates that the marginally significant does not exist in the first 50 lags. The LB test can verify the existence of the autocorrelation and the null and alternative hypothesis are showed below:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_m = 0$$

$$H_1: \text{at least } \rho_i \neq 0; i = 1, \dots, m$$

The LB test statistic equation is:

$$Q_m = T(t + 2) \sum_{i=1}^m \left(\frac{\hat{\rho}_i^2}{T - i} \right)$$

For the LB test of the XMJ index, the p-value of lag 5 and lag 10 are 0.1439 and 0.1346, which indicate that no autocorrelation existing and H_0 cannot be rejected at the 5% significant level. Therefore, the constant mean equation is suitable for the XMJ index.

Secondly, the Jarque and Bera (JB) test can diagnose whether time series are normally distributed or not by checking their sample skewness and kurtosis. The null hypothesis and alternative hypothesis of this joint test are shown as below:

$$H_0: \text{skewness} = 0 \ \& \ \text{kurtosis} = 3$$

$$H_1: \text{either skewness} \neq 0 \ \text{or kurtosis} \neq 3$$

The test statistic is:

$$JB = \frac{n}{6} \left(S^2 + \frac{(\hat{k} - 3)^2}{4} \right)$$

The sample skewness and kurtosis of XMJ are: 0.0070 and 4.48667. The p-value of the JB test is close to zero, so the null distribution is strongly rejected at the 5% significant level, even at 1% significant level. This concludes that either the sample skewness does not equal to zero or kurtosis does not equal to three. Therefore, XMJ does not follow the normal distribution.

Furthermore, the Engle’s ARCH test can prove whether the ARCH effect exists in the mean-corrected returns of the series. The null hypothesis of this test is that no ARCH effects in the first 5 and 10 lags. The p-values for lag 5 and lag 10 are 0.0002 and 2.3125×10^{-8} respectively. Therefore, the null hypothesis is strongly rejected at 5% significant level. This indicates that the ARCH effects exist in both lags.

In conclusion, all the model assumptions are proved based on the several tests above. Those models can be applied to the forecasting process.

6.2.2 Benchmark Model for 3 Indices – Ad-hoc (25) Model

Ad-hoc model is set as the benchmark model to compare with these GARCH models above. Using Ad-hoc model to generate the return and volatility forecasts of three indexes, the forecast based on the previous month data (25 days), due to the assumption of 5 transaction days per week and 5 weeks per month. The equation of the Ad-hoc model is showed below.

$$r_{t+1} = \frac{\sum_{t-24}^t r_n}{25}$$

$$\sigma_t^2 = \frac{\sum_{t-24}^t (r_n - r_{t+1})}{25}$$

6.2.3 Model Application and Analysis – XMJ

- ARCH(8)-t model

The AIC and SIC functions can be used to identify an appropriate ARCH model order. The XMJ’s AIC and SIC’s plot are shown as the following. The lowest point of the AIC is 13, while the lowest one for SIC is 8. Combing the suggestion by AIC and SIC, the model order is chosen as 8.

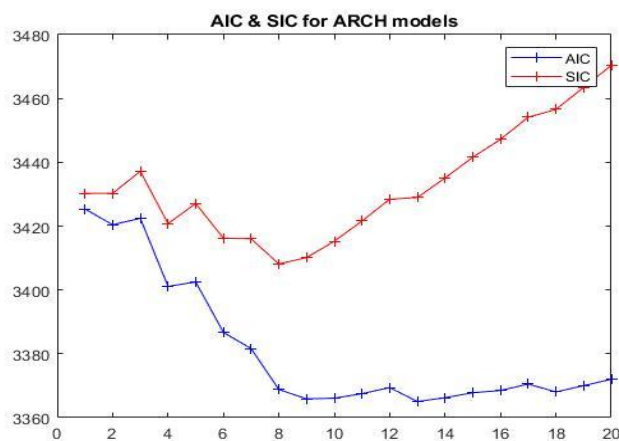


Figure 4. AIC and SIC for ARCH (8)-t model

ARCH (8) Conditional Variance Model with Offset (t Distribution):

Parameter	Value	Standard Error	T-Statistic	P-Value
Constant	0.8033	0.1294	6.2089	5.34e-10
ARCH{1}	0.0636	0.0403	1.5778	0.1146
ARCH{2}	0.0113	0.0411	0.2744	0.7838
ARCH{3}	0.0149	0.0293	0.5091	0.6107
ARCH{4}	0.1253	0.0484	2.5920	0.0095
ARCH{5}	0.0229	0.0302	0.7599	0.4473
ARCH{6}	0.1477	0.0524	2.8182	0.0048
ARCH{7}	0.0567	0.0392	1.4462	0.1481
ARCH{8}	0.1058	0.0424	2.4945	0.0126
DoF	10.2340	2.8550	3.5844	0.0003
Offset	0.0148	0.0375	0.3949	0.6930

The estimated model is:

$$r_t = a_t + 0.0148; a_t | F_{t-1} \sim t_{10.2}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.8033 + 0.0636a_{t-1}^2 + 0.0113a_{t-2}^2 + 0.0149a_{t-3}^2 + 0.1253a_{t-4}^2 + 0.0229a_{t-5}^2$$

$$+ 0.1477a_{t-6}^2 + 0.0567a_{t-7}^2 + 0.1058a_{t-8}^2$$

The t-statistics vary from 0.27 to 6.21 in a big range. Not all of them are greater than 2. Hence, some of these parameter estimates are significantly different to zero at 5% significant level, such as Constant, ARCH{4}, ARCH{6} and ARCH{8}. Others are not significantly different to zero. The p-values of these parameters support this conclusion.

The following plots show that the standard deviation process performs nice and smooth and locates nicely on the shoulder of the returns data.

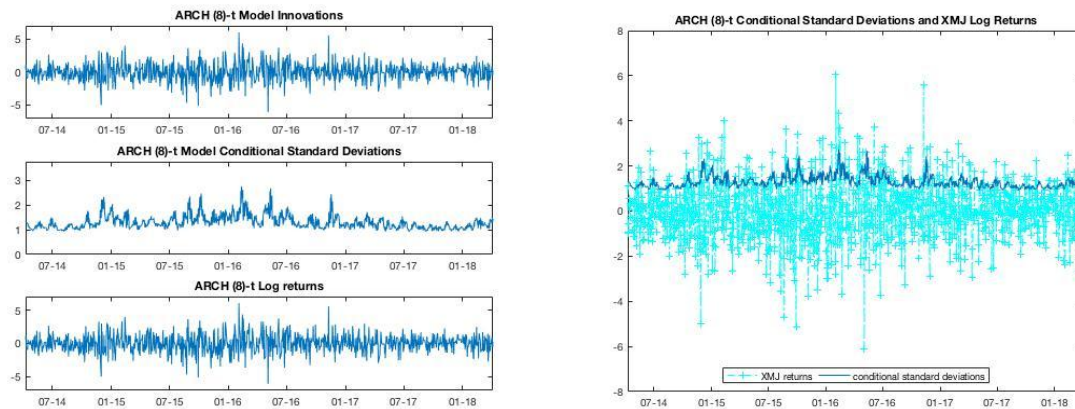


Figure 5. The ARCH (8) –t model standard deviation and XMJ log return

Firstly, the standardised residual $(\frac{\hat{a}_t}{\hat{\sigma}_t})$ need to be transformed to a normal distribution from a Student-t distribution with degree of freedom (df) of 10.23. So, the transformed standardised residuals, as presented in the below figure left part, there is no apparent autocorrelation or heteroskedasticity exists in both ACF plots. Yet, the p-values from the LB test, when $m = 13$ and $m = 18$, with df of 5 (13-8) and 10 (18-8) are 0.0300 and 0.0584, respectively. Thus, the constant cannot model the mean equation perfectly. The p-values of the transformed squared standardised residuals through a LB test are 0.0910 and 0.0831. Hence, ARCH effects do not exist at 5% significant level.

Secondly, the right graph in the below, the histogram of transformed standardised residuals could be asserted as a normal distribution though there might be two to three outliers. While, through the qq plot, all the transformed standardised residuals lie into the range between -4 and 4 prove the normality of distribution. Additionally, JB test could be applied, from which the p-value is equal to 0.5, so the null hypothesis is strongly not rejected. Furthermore, the sample skewness and kurtosis are: 0.0030 and 3.0118, which seem quite close to a normal distribution. Hence, the distribution seems like a standard Gaussian at 5% significant level.

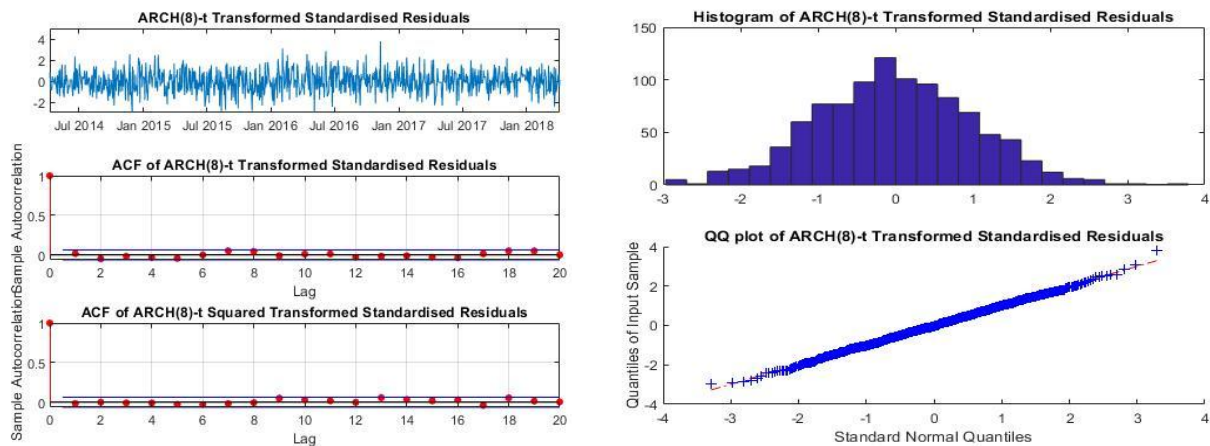


Figure 6. Assess the ARCH (8) –t model

• **GARCH(1,1)-t model**

GARCH(1,1) Conditional Variance Model with Offset (t Distribution):

Parameter	Value	Standard Error	T-Statistic	P-Value
Constant	0.0114	0.0088	1.2982	0.1942
GARCH{1}	0.9580	0.0136	70.5000	0.0000
ARCH{1}	0.0360	0.0117	3.0856	0.0020
DoF	10.2010	2.8001	3.6431	0.0003
Offset	0.0176	0.0372	0.4722	0.6368

Thus, the estimated model is:

$$r_t = a_t + 0.0176; a_t | F_{t-1} \sim t_{10.20}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.0114 + 0.0360a_{t-1}^2 + 0.9580\sigma_{t-1}^2$$

The t-statistics GARCH{1} and ARCH{1} are above 2, especially GARCH{1} which is about 70.50. Thus these two parameter estimates are significantly different to 0, especially GARCH{1} at a 5% significance level. Their p-values also show this conclusion.

The standard deviation process is smoother than the ARCH(8) model, and its location is on the shoulder of the returns data as shown in the below:

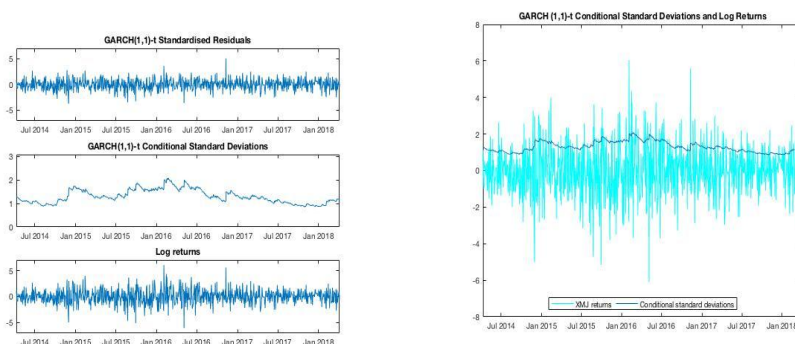


Figure 7. The GARCH (1,1) –t model standard deviation and XMJ log return

Here, the standardised residual $\frac{\hat{a}_t}{\hat{\sigma}_t}$, also need to be transformed to a normal distribution that following $\sim N(0,1)$ from Student-t distribution with df equals to 10.2. Therefore, the transformed standardised residuals, as presented in the below figure left part with its ACF plots, indicate that no apparent autocorrelation or heteroskedasticity exist in both ACF plots. While, the p-values from the LB test, when $m = 8$ and $m = 13$, with df of 5 (8-3) and 10 (13-3) are 0.0829 and 0.3269, respectively. Thus, the constant cannot model the mean equation perfectly. The p-values of the transformed squared standardised residuals (the bottom left one of Figure 6) through a LB test is 0.0534 and 0.2573. So ARCH effects statistically significantly not exist at 5% significant level.

Then, the right part in the below figure, the histogram of GARCH(1,1) transformed standardised residuals could be stated as a normal distribution though there might be two to three outliers. While, through the q-q plot, almost all the transformed standardised residuals lie close to -3 and 3, though there are some outliers. This still can prove the normality of distribution. In addition, after applying a JB test to test normality, which the p-value is exactly 0.5, so the null hypothesis is strongly not rejected. Moreover, the sample skewness and kurtosis are: -0.0393 and 3.0022, which are close to a Gaussian distribution. So, the distribution might be asserted as a standard Gaussian at 5% significant level.

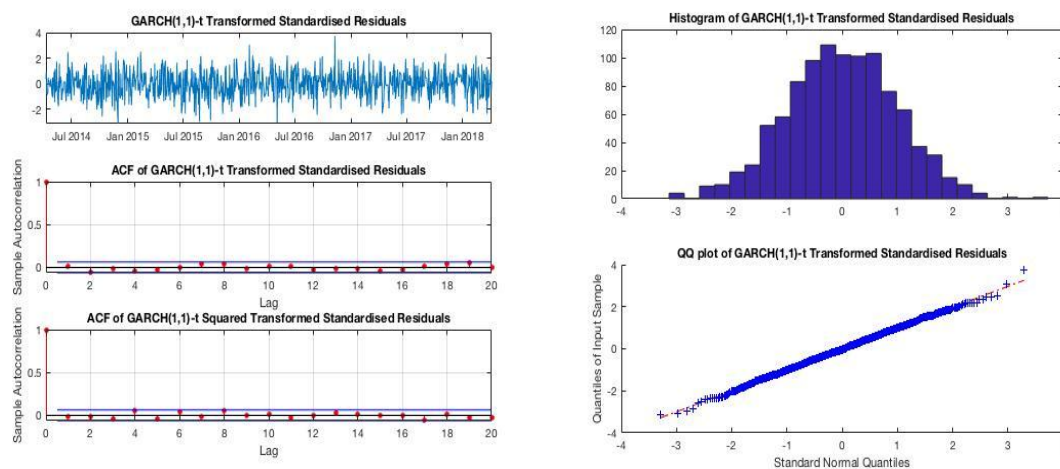


Figure 8. Assess the GARCH (1,1) –t model

• **GJR-ARCH (8) -t model**

By using the AIC and SIC criteria, the suitable ARCH model order for XMJ index is 14 for AIC and 8 for SIC. Comprehensively apply these criteria, the GJR-ARCH (8) model is suitable in this section. Then the GJR-ARCH (8) -t model is fitted below:

Table 3. MATLAB output for GJR-ARCH (8) -t Model

Parameter	Value	Standard Error	T-Statistic	P-Value
Constant	0.7350	0.1255	5.8546	4.7824e-09
ARCH{1}	0.0721	0.0576	1.2513	0.2108
ARCH{3}	0.0424	0.0457	0.9273	0.3538
ARCH{4}	0.0651	0.0495	1.3159	0.1882
ARCH{5}	0.0228	0.0361	0.6302	0.5285
ARCH{6}	0.1142	0.0626	1.8262	0.0678

ARCH{7}	0.0983	0.0570	1.7254	0.0845
ARCH{8}	0.0824	0.0587	1.4063	0.1600
Leverage{1}	-0.0119	0.0702	-0.1670	0.8653
Leverage{2}	0.1075	0.0755	1.4236	0.1546
Leverage{3}	-0.0424	0.0523	-0.8105	0.4177
Leverage{4}	0.1191	0.0861	1.3828	0.1667
Leverage{5}	0.0115	0.0582	0.1978	0.8432
Leverage{6}	0.0557	0.0885	0.6291	0.5293
Leverage{7}	-0.0717	0.0706	-1.0159	0.3097
Leverage{8}	0.0299	0.0750	0.3990	0.6899
DoF	10.7560	3.0586	3.5167	0.0004
Offset	-0.0043	0.0378	-0.1142	0.9091

Meanwhile, the estimated model is:

$$r_t = a_t - 0.0043; a_t | F_{t-1} \sim t_{10.76}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.7350 + (0.0721 - 0.0119I_{t-1})a_{t-1}^2 + 0.1075I_{t-2}a_{t-2}^2 + (0.0424 - 0.0424I_{t-3})a_{t-3}^2 + (0.0651 + 0.1191I_{t-4})a_{t-4}^2 + (0.0228 + 0.0115I_{t-5})a_{t-5}^2 + (0.1142 + 0.0557I_{t-6})a_{t-6}^2 + (0.0983 - 0.0717I_{t-7})a_{t-7}^2 + (0.0825 + 0.0299I_{t-8})a_{t-8}^2$$

However, not all the t-statistics are above 2, except for the constant term have p-values closes to 0, while the other parameters are not significantly different to 0 at a 5% significance level.

The below plots show that the standard deviation is slightly smoother than the ARCH model and locates nicely on the shoulder of the XMJ log returns data below.

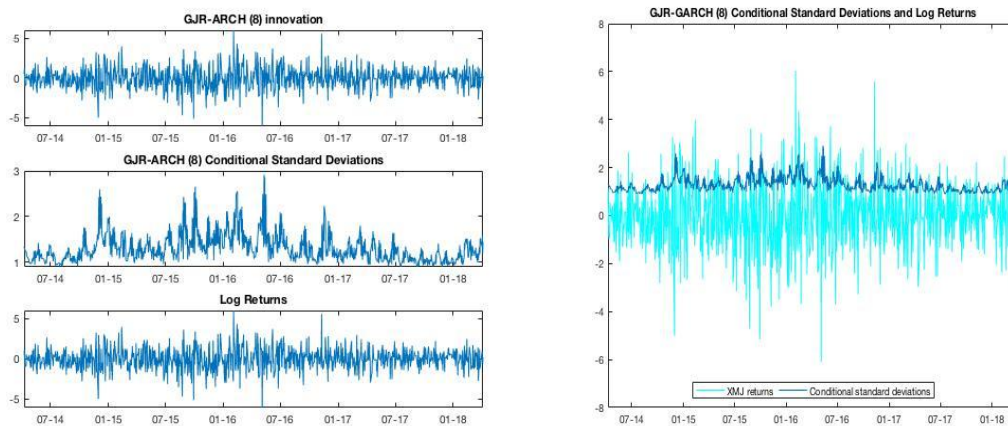


Figure 9. The GJR-ARCH (8) –t model standard deviation and XMJ log return

To assess the GJR-ARCH (8) model fitting, transforming the standard residual from a Student-t distribution with 10.76 degree of freedom to the normal distribution. The transformed standard residual plot and the ACF plot of the transformed standard residual and transformed squared standard residuals shows below. There are no clear outliers presented in the ACF plot of the standard residual. To confirm this observation, the p-values from a LB test for standard residual, with degree of freedom equal to 5 and 10 are 0.0006 and 0.0128 separately. It seems like the constant mean equation is not suitable enough. Meanwhile, the squared transformed standardised residuals at the ninth lag seems to exist the autocorrelation. Since only 1 lag autocorrelations are significant among the 20 lagged autocorrelations, the observation may be incorrect. Then the LB test applied to the squared transformed standard residual, with df=5 and df=10, the p-values of LB test are 0.0018 and 0.0277 respectively. This result may imply that some remaining ARCH effects are still exist in the data.

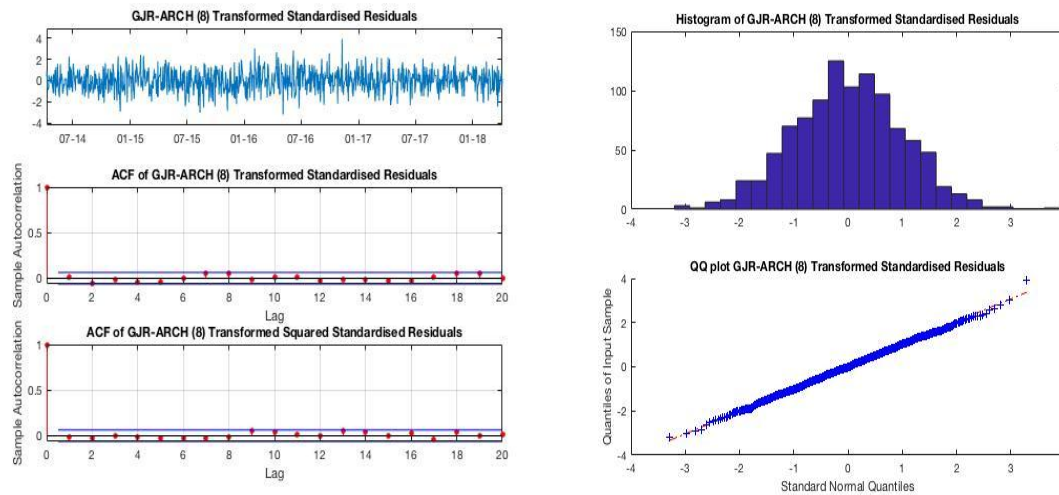


Figure 10. Assess the GJR-ARCH (8) –t model

Furthermore, according to the histogram of GJR-ARCH (8) -t transformed standardised residuals above, its distribution can be assumed to follow the normal distribution. Moreover, the outliers are hard to observed in this histogram and all transformed standardised residuals lie between -4 to 4. Also, the qq plot proves the normality of this distribution. The JB test is applied to further confirm, the p-value of the JB test is 0.5, so the null hypothesis cannot be rejected. Also, the skewness is -0.0177 and kurtosis is 3.0227, which is close to the 0 and 3 like the normal distribution. Thus, the residuals could come from the Student-t (10.76) cannot be reject.

• **GJR-GARCH (1,1) – t model**

By applying the AIC and SIC criteria, the AIC choose GJR-GARCH (2,2) -t for constant mean models with the student t distribution, while the SIC chooses GJR-GARCH (1,1) -t. Overall, the GJR-GARCH (1,1) -t should be selected since the SIC criteria choose the more parsimonious one. Then the relative parameters of the student t distribution GJR- GARCH (1,1) conditional variance model showed below:

Table 4. MATLAB output for GJR-GARCH (1,1) -t Model

Parameter	Value	Standard Error	T-Statistic	P-Value
Constant	0.0086	0.0069	1.2408	0.2147
GARCH{1}	0.9657	0.0121	79.9590	0.0000
ARCH{1}	0.0104	0.0111	0.9331	0.3508
Leverage{1}	0.0390	0.0156	2.4953	0.0126
DoF	10.9660	3.1572	3.4732	0.0005
Offset	0.0043	0.0374	0.1151	0.9084

Thus, the estimated model is:

$$r_t = a_t + 0.0043; a_t | F_{t-1} \sim t_{10.97}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.0086 + (0.0104 + 0.0390I_{t-1})a_{t-1}^2 + 0.9657\sigma_{t-1}^2$$

Except for ARCH (1), the p-value of the remaining parameters is close to 0 and the t-statistic of these parameters are larger than 2, which means these parameters are significantly different to zero at a 95% confidence level. Since the ARCH (1) has the smaller t-statistic, this parameter may not have the significant effect on this model.

The below plots show that the standard deviation is slightly smoother than the previous GARCH model and locates on the shoulder of the XMJ log returns data.

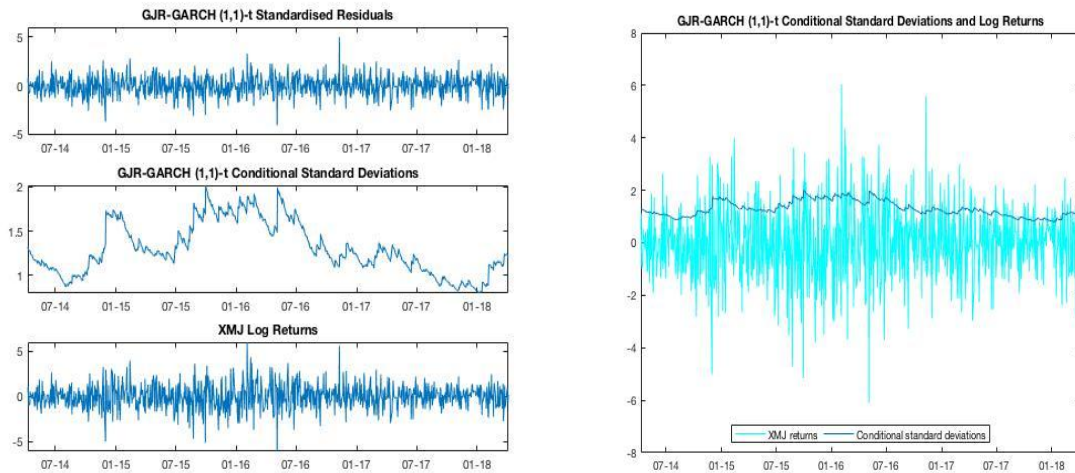


Figure 11. The GJR-GARCH (1,1) –t model standard deviation and XMJ log return

To assess the fitting effect of the GJR-GARCH (1,1)-t model, the standard residual from a Student-t distribution with 10.96 degree of freedom also need to be transformed to the normal distribution. According to the ACF plot of the transformed standard residual below, there are no clear outliers presented in this ACF plot. To confirm whether the observation is accurate, the LB test for the standard residual is applied. The p-values from a LB test, with degree of freedom equal to 5 is 0.0577 and with df=10 is 0.2887, which imply the constant mean equation is suitable for this time series data. Moreover, according to the ACF plot of the squared transformed standard residual, the squared transformed residuals at the fourth lag and ninth lag seems to display significant autocorrelation, but no others show any significant autocorrelation. Since only 2 lag autocorrelations are significant among the 20 lagged autocorrelations, the observation is probably spurious under this situation. Then the LB test applied to the squared transformed standard residual, with m=13 and df=5 and m=17 and df=10, the p-values of LB test are 0.0922 and 0.274 respectively. This result may imply that there are no remaining ARCH effects existing.

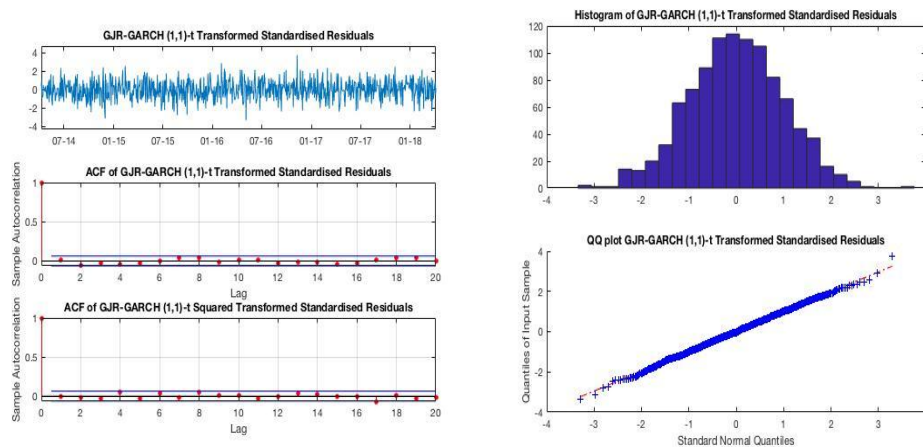


Figure 12. Assess the GJR-GARCH (1,1) –t model

Moreover, based on the histogram of GJR-GARCH (1,1) -t transformed standardised residuals above, the distribution of the transformed standardised residual can be assumed to follow the normal distribution. Moreover, all transformed standardised residuals lie between -4 to 4 and no significant outliers in this histogram. Also, the qq plot proves the normality of this distribution. The JB test is applied to further confirm, the p-value of the JB test is 0.5, so the null hypothesis cannot be rejected at 5% significant level. Also, the skewness is -0.0432 and

kurtosis is 3.0046, which is close to the 0 and 3 like the normal distribution. Thus, the residuals could come from the Student-t (10.96) cannot be reject.

6.2.4 Model Application and Analysis – XTJ

For XTJ, follow the same process above, the relative estimated models and diagnostic test result are showing below:

ARCH (2) model

$$r_t = a_t - 0.0033; a_t | F_{t-1} \sim t_{5.87}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.9418 + 0.1009a_{t-1}^2 + 0.0845a_{t-2}^2$$

GARCH (1,1)-t model

$$r_t = a_t + 0.0043; a_t | F_{t-1} \sim t_{6.08}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.0913 + 0.0593a_{t-1}^2 + 0.8610\sigma_{t-1}^2$$

GJR-ARCH (1) model

$$r_t = a_t - 0.0006; a_t | F_{t-1} \sim t_{5.64}(0, \sigma_t^2)$$

$$\sigma_t^2 = 1.0242 + (0.1001 - 0.0364I_{t-1})a_{t-1}^2$$

GJR-GARCH (1,1) model

$$r_t = a_t + 0.0035; a_t | F_{t-1} \sim t_{6.08}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.0927 + (0.0466 + 0.0167I_{t-1})a_{t-1}^2 + 0.8627\sigma_{t-1}^2$$

Table 5. Diagnostic test results for XTJ index

XTJ	ARCH(2)-t model	GARCH(1,1)-t model	GJR-ARCH (1)-t model	GJR-GARCH(1,1)-t model
P-Value of LB test for transformed standardized residual	0.1422(m=7,df=5)	0.1802(m=8,df=5)	0.0943(m=7,df=5)	0.1760(m=9,df=5)
	0.4717(m=12,df=10)	0.5902(m=13,df=10)	0.3423(m=12, df=10)	0.6008(m=14,df=10)
P-Value of LB test for squared transformed standardized residual	0.4106(m=7,df=5)	0.3377(m=8,df=5)	0.0102(m=7,df=5)	0.0279(m=9,df=5)
	0.0043(m=12,df=10)	0.1412(m=13,df=10)	0.0002(m=12,df=10)	0.0938(m=14,df=10)
P-value of JB test for Gaussian residuals	0.0061	0.0056	0.0058	0.0054

Considering the result of model assessing for XTJ, the p-values of LB test for the residual are all larger than 0.05, which implies the constant mean equation is suitable for XTJ index. Although the GJR-ARCH model still has the significant ARCH effect on the squared transformed standardized residual since the p-value of it are lower than 0.5, and the GARCH (1,1) model is the best model to fit the volatility for the XTJ index. However, the result of JB test cannot prove that the residuals are following the normal distribution at 5% significant level since the p-value is smaller than 0.05.

6.2.5 Model Application and Analysis – XXJ

For XXJ, follow the same process and results are showing below:

ARCH (5)-t model

$$r_t = a_t + 0.0294; a_t | F_{t-1} \sim t_{9.74}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.3997 + 0.2123a_{t-1}^2 + 0.0179a_{t-2}^2 + 0.0778a_{t-3}^2 + 0.1976a_{t-4}^2 + 0.1526a_{t-5}^2$$

GARCH (1,1)-t model

$$r_t = a_t + 0.0144; a_t | F_{t-1} \sim t_{9.17}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.0134 + 0.0732a_{t-1}^2 + 0.9160\sigma_{t-1}^2$$

GJR-ARCH (5)-t model

$$r_t = a_t + 0.0115; a_t | F_{t-1} \sim t_{11.66}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.3888 + (0.1395 + 0.1595I_{t-1})a_{t-1}^2 + 0.0554I_{t-2}a_{t-2}^2 + (0.059 + 0.0114)a_{t-3}^2$$

$$+ (0.1195 + 0.1681I_{t-4})a_{t-4}^2 + (0.1079 + 0.077I_{t-5})a_{t-5}^2$$

GJR-GARCH (1,1)-t model

$$r_t = a_t + 0.001; a_t | F_{t-1} \sim t_{10.5}(0, \sigma_t^2)$$

$$\sigma_t^2 = 0.0179 + (0.0145 + 0.1083I_{t-1})a_{t-1}^2 + 0.9132\sigma_{t-1}^2$$

Table 6. Diagnostic test results for XXJ index

XXJ	ARCH(5)-t model	GARCH((1,1)-t model)	GJR-ARCH (5)-t model	GJR-GARCH(1,1)-t model
P-Value of LB test for transformed standardized residual	0.0231(m=10,df=5)	0.0332(m=8,df=5)	0.0005(m=15,df=5)	0.0323(m=9,df=5)
	0.0270(m=20,df=10)	0.1068(m=13,df=10)	0.0065(m=20,df=10)	0.0457(m=14,df=10)
P-Value of LB test for squared transformed standardized residual	0.4617(m=10,df=5)	0.0104(m=8,df=5)	0.0018(m=15,df=5)	0.0163(m=9,df=5)
	0.0899(m=20,df=10)	0.0464(m=13,df=10)	0.0142(m=20,df=10)	0.0835(m=14,df=10)
P-value of JB test to Gaussian residuals	0.0568	0.0599	0.0345	0.0450

Based on model assessing result above, except to the GARCH (1,1), the other models' p-values of LB test for residual are lower than 0.05, which implies the constant mean equation may not quite suitable for XXJ index. Although the GJR-ARCH and GARCH model still has the significant ARCH effect on the squared transformed standardized residual since the p-value of it are lower than 0.5, and the ARCH (5) model seems to be the best model to fit the volatility for the XTJ index. Moreover, the result of JB test shows that the residuals are following the normal distribution for ARCH and GARCH model since their p-value is larger than 0.05.

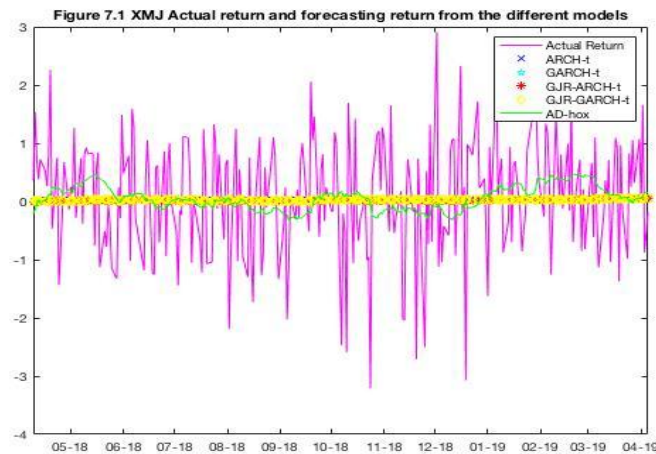
7. Different Model Forecasts Results and Accuracy Assessment

By applying the forecasting models above, using the in-sample data to train the ARCH, GARCH, GJR-ARCH, GJR-ARCH, Ad-hoc model to predict the future log return of XMJ, XTJ, XXJ. In this section, the forecast horizon of 1 day and a fixed size moving window of T= 254 is used for the forecasting.

Then to assess the accuracy of the forecasting result, focusing on the difference between the forecasting results of different models and the true values of the forecasting sample (the log return of the three series of index from 7th April 2018 to 6th April 2019). Considering the features of the index log return, root mean square error (RMSE) and mean absolute deviation (MAD) could set as the evaluation criteria and have the same unit as the percentage log return in this process. While the mean absolute percentage error (MAPE) cannot be used as a suitable measure to obtain the bias between the true values and predicted results due to the index log return could be 0. Therefore, choosing the optimal model based on the model with the minimize RMSE and MAD value to guarantee the accuracy of forecasting return and volatility.

7.1 Assess Return Forecasting Result

Firstly, we set the ad-hoc model as the benchmark model to assess the accuracy of the remaining advanced model. In ad-hoc model, the average of the last month index values (assuming 5 trading day per week and 5 weeks one month) is used to predict the future percentage log return. Then we obtain the return forecasts of the different models and visualize the comparing plot to describe the forecast result and the true values of forecasting sample for XMJ index as followed.



According to the plot above, except for the ad-hoc model, the rest of advanced models look like predicted value around to 0 merely and did not follow the trend and magnitudes of the actual log returns for sample data. Therefore, based on the following pattern, we cannot assume that the rest models fit the true data pattern as better as the benchmark model. This pattern is also applied for the XTJ and XXJ index.

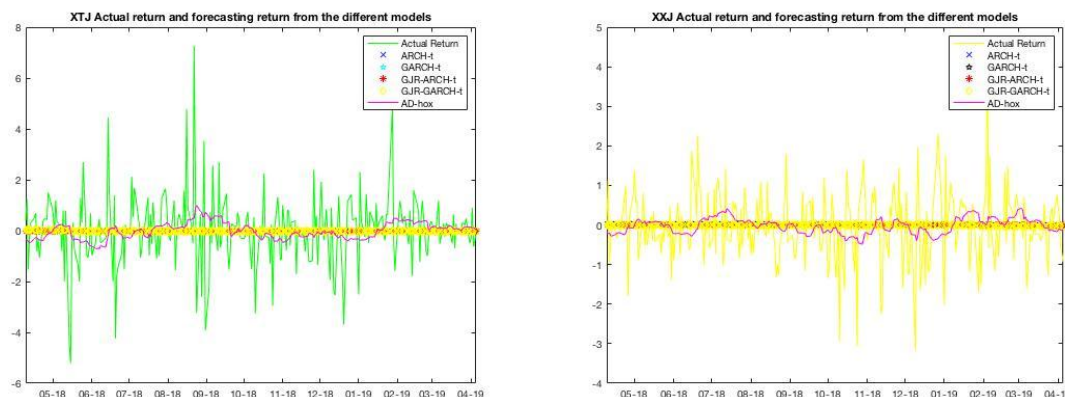


Figure 13. XTJ and XXJ Actual return and forecasting return from the different models

However, assessing accuracy based on the plots above merely is hard to distinguish the accuracy of the different advanced forecasting models. Therefore, to provide the enough evidence to have the confident conclusion, the RMSE and MAD of different models for 3 indices can be calculated to further assessing as followed.

		Ad-hoc	ARCH	GARCH	GJR-ARCH	GJR-GARCH
XMJ	RMSE	1.0316	1.0136	1.0134	1.0141	1.0139
	MAD	0.8337	0.8180	0.8181	0.8182	0.8182
XTJ	RMSE	1.4253	1.3937	1.3937	1.3937	1.3937
	MAD	0.9857	0.9551	0.9548	0.9550	0.9548
XXJ	RMSE	0.9464	0.9186	0.9184	0.9182	0.9183
	MAD	0.7165	0.6888	0.6884	0.6883	0.6881

According to the table above, all the highlight terms are the optimal models based on the corresponding criteria. For XMJ index, the GARCH model has the lowest RMSE value and the ARCH model has the lowest GARCH

model. Although different criteria imply the different optimal models, based on the slightly different, assuming the GARCH model performs better for XMJ index. For XTJ, the model with the minimize RMSE and MAD is the GJR-GARCH and GARCH model. Finally, for XXJ index, the RMSE criteria imply the GJR-ARCH model is the optimal while the MAD also select the GJR-GARCH model like XTJ. Therefore, considering there are 50 percent of probability the GJR-GARCH model perform better than the other models, the GJR-GARCH model can be assumed as the appropriate forecasting model for these indices.

7.2 Assess Volatility Forecasting Result

In this section, firstly, using the volatility proxies to measure the forecasting sample accuracy and focusing on XMJ index with proxy 1 as sample in this process. The plot below shows proxies 1 along with all forecasts – from both the ad hoc and GARCH-type models. Focusing on the proxy 1 in green pentagram, accuracy measures of proxy 1 are usually the highest since the returns are often close to 0 but volatility forecasts never are. According to the figure below, the GJR-GARCH (1,1) -t model performs optimally under the proxy 1.

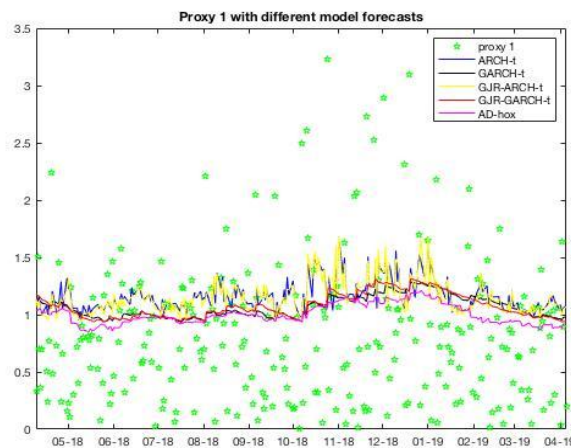


Figure 14. Proxy 1 with the different model forecasts

Then using the RMSE and MAD to further assess the volatility forecasting results for all three stocks individually.

According to the table showed above, for index XMJ, although the GARCH (1,1) performs better in the four

	RMSE	Proxy1	Proxy2	Proxy3	Proxy4	MAD	Proxy1	Proxy2	Proxy3	Proxy4
XMJ	ARCH	0.6997	0.5206	0.4999	0.4645	XMJ	0.5740	0.4719	0.4474	0.3988
	GARCH	0.6540	0.4405	0.4269	0.4063		0.5304	0.3884	0.3711	0.3365
	GJR-ARCH	0.7002	0.5181	0.4978	0.4635		0.5727	0.4671	0.4429	0.3967
	GJR-GARCH	0.6545	0.4441	0.4295	0.4065		0.5322	0.3915	0.3738	0.3378
	Ad hoc	0.6295	0.3950	0.3863	0.3786		0.5018	0.3416	0.3297	0.3058
XTJ	ARCH	1.0094	0.5858	0.5986	0.6526	XTJ	0.7516	0.4441	0.4336	0.4275
	GARCH	1.0257	0.5965	0.6074	0.6570		0.7658	0.4500	0.4410	0.4376
	GJR-ARCH	1.0136	0.5912	0.6043	0.6588		0.7562	0.4480	0.4381	0.4352
	GJR-GARCH	1.0297	0.6017	0.6121	0.6606		0.7694	0.4549	0.4445	0.4399
	Ad hoc	1.0220	0.5983	0.6203	0.6931		0.7398	0.4270	0.4244	0.4366
XXJ	ARCH	0.6982	0.4387	0.4298	0.4260	XXJ	0.5538	0.3592	0.3451	0.3222
	GARCH	0.6605	0.4048	0.3977	0.3986		0.5273	0.3336	0.3194	0.2993
	GJR-ARCH	0.6995	0.4372	0.4283	0.4247		0.5516	0.3559	0.3425	0.3207
	GJR-GARCH	0.6485	0.3855	0.3776	0.3785		0.5172	0.3185	0.3043	0.2850
	Ad hoc	0.6559	0.4029	0.3967	0.3993		0.5314	0.3372	0.3240	0.3048

GARCH models, the Ad hoc model has the minimize RMSE and MAD value for all of 4 proxies among the all of five forecasting models. For XTJ, the ARCH model is selected as the optimal forecasting model under the minimize RMSE criteria and the Ad-hoc model performs better at the MAD criteria. Meanwhile, for the XXJ index, the GJR-GARCH model performs better at the RMSE and MAD criteria. In conclude, the Ad-hoc model performs better in volatility forecasting for these three indices.

8. Asset Portfolio Weights Allocation

The purpose for this section is providing different rules to choose the optimal portfolio weights for the three assets, which including XTJ that is telecommunication services, XXJ that is financials excluding A-REITs and XMJ that is materials. There is no doubt that choosing the optimal portfolio weights for the assets is very significant for the financial investment, it can help the investments with maximize return of the portfolio and minimize the risk of the investment. It chooses the equally weighted rule, return rule and risk rule to allocate the weights. The return rule means the higher return of the asset the higher weight for this asset is, and we choose the Sharpe ratio for the risk rule to match the risk management as the question required, it follow the standard that the higher Sharpe ratio of the asset, the higher weight for this asset is (Ritter& Chopra, 1989).

8.1 Analyzing Weights Allocation Rules

The equally weighted rule is providing the same proportion for the assets, for example, in this case it is 1/3 for XTJ, 1/3 for XXJ and 1/3 for XMJ. It is always recognized as the benchmark, and easier to use than other allocation rules. Normally, it can achieve the higher returns than the average returns. Therefore, the investors who without the investing knowledge and financial skills usually use this way to invest. Additionally, the equally weighted rule is a passive method for obtaining most variations of global opportunities, it can help the analysts to analyze the investment environment ((Ritter& Chopra, 1989).

Comparing to equally weighted rule, the return rule is more aggressive, it may lose sight of the unforeseen market volatilities that they may meet in the investment horizon. The standard of this rule is the weights allocated to each asset are depended on the maximum forecasted return, which means higher forecasted return asset, the higher percentage allocated is (Willenbrock, 2011).

The Sharpe ratio also known as the Sharp Index that is a standardized indicator of fund performance evaluation. The Sharpe ratio in modern investment theory research shows that the size of the risk plays a fundamental role in determining the performance of the combination. The risk-adjusted rate of return is a comprehensive indicator that considers both benefits and risks, with a view to eliminating the negative impact of risk factors on performance evaluation.

The Sharpe ratio formula is $Sharpe\ ratio = \frac{r_p - r_f}{\sigma_p}$

The Sharpe ratio is one of the three classic indicators that can simultaneously consider the benefits and risks. There is a regular feature in investment, that is, the higher the expected return of the investment target, the higher the risk of volatility that the investor can bear; on the contrary, the lower the expected return, the lower the risk of volatility, so the rational investor chooses the investment target and investment. The main purpose of the combination is to pursue maximum compensation at a fixed risk, or to pursue the lowest risk at a fixed expected return (Zakamouline & Koekebakker, 2009).

8.2 Formulas for Calculating & Forecasting Weights Allocation for Index Portfolio

To calculate and forecast the weights allocation for the index portfolio, we choose the all observations up to and including 6th April 2018 as the in-sample period, and the remaining 254 data will be the forecasting period. The allocated percentage for each index is related to the mean of performances that is forecasted. The more particular information for the process is shown below:

Weights Allocation Rules			
	Equally Weighted Rule	Return Rule	Sharpe ratio Rule
Formula	1/3	$W_{r,i} = \frac{\bar{r}_i}{\sum_{i=1}^3 \bar{r}_i}$	$W_{Sharpe\ ratio,i} = \frac{Sharpe\ ratio_i}{\sum_{i=1}^3 Sharpe\ ratio_i}$

$W_{r,i}$ denotes the percentage allocated for index I when using the return rule, \bar{r}_i is the mean of forecasted return for index i (i= 1, 2, 3 is the asset that our group analyzed, which is telecommunication services, financials excluding A-REITs and materials) in the forecasting period.

$W_{Sharpe\ ratio_i}$ means the percentage allocated for index i when using the Sharpe ratio method, $Sharpe\ ratio_i$ is the Sharpe ratio for index i in the forecasting period, and the $\sum_{i=1}^3 Sharpe\ ratio_i$ is the total of the Sharpe ratios for the three series, which are XTJ, XXJ and XMJ.

9. Rebalance & Re-estimated for Asset Portfolio Weights Allocation Using Different Frequencies

9.1 Analysis the Frequency for Re-estimated

In this section, it will provide the rebalance and re-estimated for the weight every period and every 5th period for asset portfolio weights allocation, every period presents the daily variation, and the every 5th period shows the weekly changing. The number of re-adjust frequencies have a relationship with the cost of the transaction, with the number of frequencies increase in the trading, the total cost such as transaction cost may increase; on the opposite, with the number of re-adjust frequencies decrease, the transaction cost may decrease. Also, when the frequency of trading in market is decreased, it can make the investors capture the opportunities difficultly and make more unforeseen and uncertainties result for the investment (Ghysels et al, 2006). Moreover, if investors cannot change their weights allocation for their asset portfolio based on the market change immediately when the market situation have huge volatility trend, the investors may have more unforeseen loss. Consequently, adopt more frequency to re-estimate the asset portfolio weights allocation is more accuracy than the infrequency, but it has more discourse cost, such the time use, manpower, attention and some other resources (Leibowitz& Bova, 2011). Because the frequency of in-sample data that given is daily, to ensure the result accuracy, reliable and confirm the market dynamics, we choose the daily and weekly frequency as our asset portfolio weight allocation re-estimation periodicity.

9.2 Analysis the Methods for Two Frequency of Re-estimated

9.2.1 Every Period

For every period re-estimated method for the asset portfolio weights allocation, the weight will be re-calculated every day, we use the similar formulas for asset portfolio weights allocation in 8th section, which are:

Every Period Weights Allocation Rules			
	Equally Weighted Rule	Return Rule	Sharpe ratio Rule
Formula	1/3	$W_{r,i,t} = \frac{\bar{r}_{i,t+t}}{\sum_{i=1}^3 \bar{r}_{i,t+t}}$	$W_{Sharpe\ ratio,i,t} = \frac{Sharpe\ ratio_{i,t+t}}{\sum_{i=1}^3 Sharpe\ ratio_{i,t+t}}$

$W_{r,i,t}$ means the percentage allocated for index I in day t when using the return rule, $\bar{r}_{i,t+t}$ is the mean of forecasted return in every period for index i in the forecasting period.

$W_{Sharpe\ ratio,i,t}$ means the percentage allocated for index i in day t when using the Sharpe ratio method, $Sharpe\ ratio_i$ is the Sharpe ratio for index i in every period of the forecasting period.

9.2.2 Every 5th Period

For every 5th period re-estimated method for the asset portfolio weights allocation, the weight will be re-calculated every five days, we use the similar formulas for asset portfolio weights allocation in 8th section, which are:

Every 5 th Period Weights Allocation Rules			
	Equally Weighted Rule	Return Rule	Sharpe ratio Rule
Formula	1/3	$W_{r,i,t} = \frac{\bar{r}_{i,t+5t}}{\sum_{i=1}^3 \bar{r}_{i,t+5t}}$	$W_{Sharpe\ ratio,i,t} = \frac{Sharpe\ ratio_{i,t+5t}}{\sum_{i=1}^3 Sharpe\ ratio_{i,t+5t}}$

$W_{r,i,t}$ means the percentage allocated for index I in day t when using the return rule, $\bar{r}_{i,t+5t}$ is the mean of forecasted return in every 5th period for index i in the forecasting period. $W_{Sharpe\ ratio,i,t}$ means the percentage allocated for index i in day t when using the Sharpe ratio method, $Sharpe\ ratio_i$ is the Sharpe ratio for index i in every five days of the forecasting period.

9.3 Analysis the Re-estimated Result

For the equally weighted rule, the mean is **0.0226%** and standard deviation is **0.7839%** for both **every period** and **every 5th period** is same.

9.3.1 Every Period Analysis

Every Period Portfolio Performance				
Model	Return Rule		Sharpe ratio Rule	
	Mean	Standard deviation	Return Mean	Standard deviation
AD-HOL	0.1979%	4.0773%	0.7636%	10.6582%
ARCH	0.0880%	3.8510%	-27.1378%	425.7994%
GARCH	-0.0008%	2.3382%	-0.0245%	1.8175%
GJR-ARCH	-0.4531%	14.6858%	16.9969%	253.1612%
GJR-GARCH	3.5790%	41.9566%	3.2184%	26.9087%

From calculated the mean and standard deviation for re-estimated the asset portfolio weights allocation using every period frequency with three different methods that are the equally weighted, return rule and Sharpe ratio rule, the Sharpe ratio rule weighted portfolio GJRARCH model has the highest return 16.9969% and the lowest return -27.1378% for ARCH model with the highest standard deviation 425.7994%. In the return rule, the GJRGARCH model has the highest return 3.5790% and highest standard deviation 41.9566%. Compare the standard deviation under return rule and Sharpe ratio rule, the lowest standard deviation is 1.8175% under the Sharpe ratio rule but with the negative mean of return. However, the real lowest standard deviation is 0.7839% under equally weighted rule and with the positive mean of return.

9.3.2 Every 5th Period Analysis

Every 5 th Period Portfolio Performance				
Model	Return Rule		Sharpe ratio Rule	
	Mean	Standard deviation	Return Mean	Standard deviation
AD-HOL	0.0506%	2.7570%	-0.1392%	2.2630%
ARCH	-0.0205%	5.0315%	-0.0122%	3.0289%
GARCH	-0.0408%	2.5842%	-0.1101%	1.9381%
GJR-ARCH	-1.1651%	15.3880%	-3.8858%	44.8532%
GJR-GARCH	3.2839%	59.5706%	-1.1277%	38.7938%

From calculated the mean and standard deviation for re-estimated the asset portfolio weights allocation using every 5th period frequency the equally weighted rule, return rule and Sharpe ratio rule, the return rule weighted portfolio GJRGARCH model has the highest return 3.2839%, but with the highest standard deviation 59.5706% under the return rule. And under the Sharpe ratio rule, the lowest return is -3.8858% under GJRARCH model with the second highest standard deviation 44.8532%. Comparing the return rule and Sharpe ratio rule, the highest mean of return is occurred in the return rule and the lowest mean of return is presented in the Sharpe ratio rule, and the lowest standard deviation is 1.9381% but with the negative mean of return that is -0.1101% under the Sharpe ratio rule. For the equally weighted method, the standard deviation 0.7839% is lower than the 1.9381%, it is the real lowest standard deviation in these three rules, compared to the lowest value under Sharpe ratio rule, it has the positive mean of return that is 0.0226%.

10. Use Sharpe Ratio to Compare Portfolio Performance

The Sharpe ratio indicates the risk premium of the portfolio per unit of the total risk in the portfolio, and the formula is $Sharpe\ ratio = \frac{r_p - r_f}{\sigma_p}$, according to Australia government the r_f in Australia is 1.55% that is 10 years bond yield (Bloomberg, 2019), therefore, the result for the portfolio performance by use the Sharpe ratio is shown below:

	Return rule	Sharpe ratio rule	Equally Weighted rule
AD-HOL-1	-0.3316	-0.0738	
ARCH-1	-0.3796	-0.0637	
GARCH-1	-0.6632	-0.0135	-1.9485
GJR-ARCH-1	-0.1364	0.0671	
GJR-GARCH-1	0.0484	0.1196	
ADHOL-5	-0.5439	-0.0615	

ARCH-5	-0.3121	-0.0040	
GARCH-5	-0.6156	-0.0568	-1.9485
GJR-ARCH-5	-0.1764	-0.0866	
GJR-GARCH-5	0.0291	-0.0291	

Consequently, calculating the Sharpe ratio for the rebalance result and re-estimating the asset portfolio weights allocation using different frequencies during every first and 5th period. And using the equally weighted rule, return rule and Sharpe ratio rule, the result by Sharpe ratio shows the highest value is 0.1196 under the GJR-GARCH model with every period frequency, which means the GJR-GARCH model with every period frequency under the Sharpe ratio has the best performance. The lowest value is -1.9485 under the equally weighted rule, which has the worst performance. Under the return rule, the best performance from the GJR-GARCH model with every period frequency, is 0.0484. For every 5th period frequency, the best performance is 0.0291 with the GIR-GARCH model under the return rule, and the performances for the portfolios under the Sharpe ratio rule are all negative.

11. Conclusion

In conclusion, for factor modeling, the results indicate the two-factor model performs better than the single-factor model. For dynamic portfolio optimization, since no autocorrelations proved by ACF plots, constant mean equation is adequate. LB test and Engle test exhibit the existence of ARCH effect. By the tests and statistics analysis, selecting the models as follows: ad-hoc, ARCH, GARCH, GJR-ARCH and GJR-GARCH. All models except ad-hoc help to explain the volatility dynamics but the remain ARCH effect in the volatility equation could be a problem. Considering the forecast accuracy, the Ad-hoc model perform better in volatility forecasting for these three indices. Then, to construct the portfolio, setting three methods: equally weights, high return and low risk. Sharpe ratio is the best method for finding the optimal portfolio allocation. It is worth to build the GJR-GARCH model because it has the best performance for the rebalance result and re-estimating the asset portfolio.

References

- AIE2025. (2019). Retrieved April 22, 2019, from <https://www.austrade.gov.au/australian/education/services/australian-international-education-2025/aie2025>
- Barber, J.R., & Copper, M.L. P(2012). Principal component analysis of yield curve movements. *Journal of Economics and Finance*, 36, 750. <https://doi.org/10.1007/s12197-010-9142-y>
- Beattie, A. (2018). *A Beginner's Guide to Mining Stocks*, Retrieved 20 April 2018, from <https://www.investopedia.com/articles/basics/12/beginners-guide-mining-stocks.asp>
- Bloomberg. (2019). Retrieved 27 May 2019, from <https://www.bloomberg.com/markets/rates-bonds/government-bonds/australia>
- Compare Sector Characteristics. (2019). Retrieved 19 April 2019, from <https://www.fidelity.com/sector-investing/compare-sectors>
- Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1), 59-95.
- GICS. (2019). Retrieved 14 April 2019, from <https://www.asx.com.au/products/gics.htm>
- Jolliffe, I. T. (1986). *Principal Component Analysis*. New York: Springer.
- Koprowski, E. (2016). *Australia Wants One Million International Students by 2025*. Retrieved 14 April 2019, from <https://www.masterstudies.com/news/Australia-Wants-One-Million-International-Students-by-2025-705/>
- Leibowitz, M.L., & Bova, A. (2011) Policy portfolios and rebalancing behavior. *Journal of Porfolio Management*, 37(2), 60-71.
- Livewire. (2019). *What is Telstra's 5G strategy?*. Retrieved from 23 April 2019, from <https://www.livewiremarkets.com/wires/what-is-telstra-s-5g-strategy>
- Maverick, J. B. (2015). *What main factors affect share prices in the metals and mining sector?*. Investopedia.
- Resource and Mining Stocks. (2019). Retrieved 17 April 2019, from <https://www.money Morning.com.au/commodities/resources-and-mining/resources-and-mining-stocks>

Ritter, J.R., & Chopra, N. (1989). Portfolio Rebalancing and the Turn-of-the-Year Effect. *The Journal of Finance*, 44(1), 149-166.

The West Australian. (2019). Retrieved 22 April 2019, from <https://thewest.com.au/business/retail/buckle-up-5g-is-finally-here-with-telstra-rolling-out-hotspot-devices-and-smartphones-ng-b881207571z>

Willenbrock, S. (2011). Diversification Return, Portfolio Rebalancing, and the Commodity Return Puzzle. *Financial Analysts Journal*, 67(4), 42-49.

Yang, L., Rea, W., & Rea, A. (2017). Financial insights from the last few components of a stock market PCA. *International Journal of Financial Studies*, 5(3), 15. <https://doi.org/10.3390/ijfs5030015>

Zakamouline, V., & Koekebakker, S. (2009) Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *Journal of Banking & Finance*, 33(7), 1242-1254.

Appendixes

	XPJ	XDJ	XSJ	XEJ	XXJ	XHJ	XNJ	XIJ	XMJ	XTJ	XUJ
XPJ	1	2.94E-95	2.07E-65	5.87E-27	4.11E-80	2.30E-58	8.27E-117	2.35E-33	1.56E-27	9.30E-41	1.99E-93
XDJ	2.94E-95	1	1.99E-139	4.10E-76	7.71E-162	1.68E-122	6.38E-191	5.34E-129	6.68E-76	1.16E-53	6.20E-77
XSJ	2.07E-65	1.99E-139	1	5.34E-63	5.26E-119	1.39E-80	3.71E-112	2.91E-75	3.17E-65	1.08E-44	3.38E-52
XEJ	5.87E-27	4.10E-76	5.34E-63	1	1.62E-91	3.61E-36	2.84E-65	4.19E-54	1.56E-173	1.58E-23	1.67E-35
XXJ	4.11E-80	7.71E-162	5.26E-119	1.62E-91	1	3.10E-94	5.77E-136	1.88E-102	1.51E-98	1.54E-45	1.09E-57
XHJ	2.30E-58	1.68E-122	1.39E-80	3.61E-36	3.10E-94	1	1.86E-136	1.18E-100	2.45E-35	3.89E-25	1.75E-50
XNJ	8.27E-117	6.38E-191	3.71E-112	2.84E-65	5.77E-136	1.86E-136	1	7.05E-123	1.72E-75	2.32E-49	5.51E-92
XIJ	2.35E-33	5.34E-129	2.91E-75	4.19E-54	1.88E-102	1.18E-100	7.05E-123	1	1.63E-60	6.93E-37	1.12E-28
XMJ	1.56E-27	6.68E-76	3.17E-65	1.56E-173	1.51E-98	2.45E-35	1.72E-75	1.63E-60	1	2.87E-20	2.73E-32
XTJ	9.30E-41	1.16E-53	1.08E-44	1.58E-23	1.54E-45	3.89E-25	2.32E-49	6.93E-37	2.87E-20	1	2.10E-34
XUJ	1.99E-93	6.20E-77	3.38E-52	1.67E-35	1.09E-57	1.75E-50	5.51E-92	1.12E-28	2.73E-32	2.10E-34	1

Industry	XPJ	XDJ	XSJ	XEJ	XXJ	XHJ	XNJ	XIJ	XMJ	XTJ	XUJ
XPJ	1.0000	0.5368	0.4544	0.2959	0.4978	0.4311	0.5847	0.3293	0.2991	0.3634	0.5324
XDJ	0.5368	1.0000	0.6277	0.4866	0.6645	0.5962	0.7055	0.6087	0.4860	0.4145	0.4889
XSJ	0.4544	0.6277	1.0000	0.4466	0.5892	0.4991	0.5750	0.4842	0.4538	0.3800	0.4092
XEJ	0.2959	0.4866	0.4466	1.0000	0.5276	0.3428	0.4540	0.4161	0.6818	0.2760	0.3397
XXJ	0.4978	0.6645	0.5892	0.5276	1.0000	0.5343	0.6215	0.5537	0.5447	0.3834	0.4288
XHJ	0.4311	0.5962	0.4991	0.3428	0.5343	1.0000	0.6224	0.5496	0.3389	0.2855	0.4027
XNJ	0.5847	0.7055	0.5750	0.4540	0.6215	0.6224	1.0000	0.5969	0.4848	0.3985	0.5288
XIJ	0.3293	0.6087	0.4842	0.4161	0.5537	0.5496	0.5969	1.0000	0.4384	0.3461	0.3053
XMJ	0.2991	0.4860	0.4538	0.6818	0.5447	0.3389	0.4848	0.4384	1.0000	0.2554	0.3241
XTJ	0.3634	0.4145	0.3800	0.2760	0.3834	0.2855	0.3985	0.3461	0.2554	1.0000	0.3344
XUJ	0.5324	0.4889	0.4092	0.3397	0.4288	0.4027	0.5288	0.3053	0.3241	0.3344	1.0000

(p-value correlation coefficients)

Upper Bound (Correlation Coefficients)

	XPJ	XDJ	XSJ	XEJ	XXJ	XHJ	XNJ	XIJ	XMJ	XTJ	XUJ
XPJ	1.0000	0.4964	0.4095	0.2447	0.4552	0.3851	0.5472	0.2793	0.2480	0.3146	0.4917
XDJ	0.4964	1.0000	0.5931	0.4434	0.6325	0.5594	0.6767	0.5728	0.4427	0.3678	0.4458
XSJ	0.4095	0.5931	1.0000	0.4014	0.5520	0.4565	0.5369	0.4408	0.4089	0.3318	0.3622
XEJ	0.2447	0.4434	0.4014	1.0000	0.4867	0.2932	0.4091	0.3695	0.6511	0.2242	0.2900
XXJ	0.4552	0.6325	0.5520	0.4867	1.0000	0.4937	0.5865	0.5143	0.5047	0.3354	0.3827
XHJ	0.3851	0.5594	0.4565	0.2932	0.4937	1.0000	0.5874	0.5099	0.2891	0.2341	0.3555
XNJ	0.5472	0.6767	0.5369	0.4091	0.5865	0.5874	1.0000	0.5602	0.4415	0.3511	0.4879
XIJ	0.2793	0.5728	0.4408	0.3695	0.5143	0.5099	0.5602	1.0000	0.3928	0.2966	0.2544
XMJ	0.2480	0.4427	0.4089	0.6511	0.5047	0.2891	0.4415	0.3928	1.0000	0.2031	0.2738
XTJ	0.3146	0.3678	0.3318	0.2242	0.3354	0.2341	0.3511	0.2966	0.2031	1.0000	0.2845
XUJ	0.4917	0.4458	0.3622	0.2900	0.3827	0.3555	0.4879	0.2544	0.2738	0.2845	1.0000

Lower Bound (Correlation Coefficient)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
PC1	1	-8.01E-17	-1.13E-16	-1.44E-16	1.41E-18	7.08E-17	-6.05E-17	-1.84E-16	2.43E-16	1.34E-16	-5.51E-18
PC2	-8.01E-17	1	2.69E-18	6.33E-17	-8.24E-17	2.63E-16	-1.77E-17	-1.57E-16	3.43E-16	2.53E-16	7.30E-17
PC3	-1.13E-16	2.69E-18	1	2.82E-16	-5.55E-17	5.64E-17	3.52E-16	-2.46E-18	-2.32E-18	-3.09E-17	-9.72E-17
PC4	-1.44E-16	6.33E-17	2.82E-16	1	2.46E-16	1.06E-17	-1.58E-16	3.92E-17	-2.08E-17	-2.78E-17	2.36E-16
PC5	1.41E-18	-8.24E-17	-5.55E-17	2.46E-16	1	1.26E-17	1.58E-17	-6.03E-18	9.17E-17	-2.43E-16	1.21E-16
PC6	7.08E-17	2.63E-16	5.64E-17	1.06E-17	1.26E-17	1	4.05E-16	-5.84E-16	4.73E-16	-1.39E-16	1.10E-16
PC7	-6.05E-17	-1.77E-17	3.52E-16	-1.58E-16	1.58E-17	4.05E-16	1	3.06E-16	-6.54E-17	1.11E-17	2.11E-16
PC8	-1.84E-16	-1.57E-16	-2.46E-18	3.92E-17	-6.03E-18	-5.84E-16	3.06E-16	1	3.09E-16	-2.69E-16	3.36E-16
PC9	2.43E-16	3.43E-16	-2.32E-18	-2.08E-17	9.17E-17	4.73E-16	-6.54E-17	3.09E-16	1	6.78E-17	6.87E-17
PC10	1.34E-16	2.53E-16	-3.09E-17	-2.78E-17	-2.43E-16	-1.39E-16	1.11E-17	-2.69E-16	6.78E-17	1	4.25E-17
PC11	-5.51E-18	7.30E-17	-9.72E-17	2.36E-16	1.21E-16	1.10E-16	2.11E-16	3.36E-16	6.87E-17	4.25E-17	1

Principal Components (Correlation Coefficient)

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).