2022 Fall

FINE 6310-Market Neutral Project

Group 1

Guangyu Wu, 214406763

Yujun Liu, 218374314

Vivek Solanki, 219101229

Yiyang Yao, 219309921

Abstract

Our objective is to create a market-neutral portfolio with three pairs of long-short pair trading from 3 different sectors to ensure the portfolio's value is not impacted by the market condition change. The three sectors have different extents of sensitivity to the economic condition in the investment universe. The first pair DVA and PHG are from the Healthcare industry, which is relatively defensive to the macro condition change. The second and third pairs are MEI and CAJ, SAH and F, which are from Technology and Consumer cyclical, respectively. The latter two-sector heavily relies on economic condition and are sensitive to market volatility. We used twenties years of monthly data from 2002 to 2021 of the respective companies and three Fama French model factors generated from WRDS to create a market-neutral portfolio. In the first stage, we compute each company's market beta; then, we use their Beta of market volatility to assign weight against its pair. Each pair was confirmed market-neutral, and the final portfolio was the average weight of the three pairs. The out-of-sample Beta is 1.3556, which indicates our portfolio will not remain neutral over time and requires reweighting.

Contents

Introduction

The focus of this research is to form a market-neutral (Beta = 0) equity-only portfolio by hedging two securities to develop a long-short pair trading to ensure the portfolio's value will not be impacted by the market condition. In our portfolio construction, we incorporated multiple strategies in an attempt to create a robust portfolio.

Literature Review

1. A note on market-neutral portfolio selection-Clarence C.Y. Kwan *
    a. Long-short equity strategies allow investors to benefit potentially from both undervalued and overvalued securities (p.g 1)
    b. Unlike the other long-short trading strategies, the market-neutral strategy does not involve derivative securities. This is where a portfolio response to market movements is eliminated by offsetting market exposures with long and short equity holdings. The spread in returns between long and short positions provides a residual return relative to an underlying market index (p.g 24)
2. Patton, A.J. "Are `Market Neutral' Hedge Funds Really Market Neutral?", The Review of Financial Studies, 22(7), 2009, 2495-2530
    a. The most commonly used risk-based definition of neutrality is based on correlation or "beta": a fund may be said to be market neutral if it generates returns that are uncorrelated with the returns on some market index(p.g 2)
    b. While correlation and Beta are simple to compute and easy to interpret, the dynamic strategies employed by hedge fund managers and the resulting nonlinear payoff profiles mean that these simple measures of neutrality may provide misleading information about the true diversification benefits offered by these funds(p.g 2)

c. The author proposes an alternative of five neutrality concepts: mean neutrality, Variance neutrality,value-at-risk neutrality, and tail neutrality. These four factors all relate to the neutrality of the risk of the hedge fund returns to market returns. And the final part is complete neutrality, which corresponds to the statistical independence of fund and market returns. (p.g 3)

3. Gatev, E., Goetzmann, W., Rouwenhorst, K.G. Pairs Trading: Performance of a Relative-Value Arbitrage Rule, The Review of Financial Studies, 19(3), 2006, 797–827.

   a. The concept of pairs trading is disarmingly simple. Find two stocks whose prices have moved together historically; when the spread between them widens, short the winner and buy the loser. If history repeats itself, prices will converge, and the arbitrageur will profit. It is hard to believe that such a simple strategy, based solely on past price dynamics and simple contrarian principles, could possibly make money. If the U.S. equity market were efficient at all times, risk-adjusted returns from pairs trading should not be positive(p.g 1)

   b. The author found that trading suitably formed pairs of stocks exhibits profits which are robust to conservative estimates of transaction costs. These profits are uncorrelated to the S&P 500;

      i. however, they do exhibit low sensitivity to the spreads between small and large stocks and between value and growth stocks in addition to the spread between high- and intermediate-grade corporate bonds and shifts in the yield curve.

   c. One view of the lower profitability of pairs trading in recent year is that returns are competed away by increased hedge fund activity. The alternative view, taken in this article, is that abnormal returns to pairs strategies are compensation to arbitrageurs for enforcing the "Law of One Price." (p.g 30)

Data Description

WRDS (Wharton research data services, 2022) was used to source all of the data tested in this report. We have collected 20 years of the HPR (holding period returns) from CRSP of six selected companies traded on the U.S. stock exchanges from December 31, 2001, to December 31, 2021, with a total of 241 observations per company. The first ten years are used as an estimation window, while the last ten years will be out of the sample window to examine our profile evolution over time.

Data procedures: The first ten years from 2002 -2011 are defined as the in-sample period, and the last ten years are defined as the out-of-sample period.

The following data is for the in-sample period data for each company (a complete summary can be found in Appendix: C)

	Defensive		Sensitive		Cyclical	
SECTOR	**Healthcare**		**Technology**		**Consumer cyclical**	
Ticker	DVA	PHG	MEI	CAJ	SAH	F
Mean	0.01426994	0.002891595	0.0106667	0.008130833	0.01839155	0.008876813
Standard Error	0.0061256	0.008750481	0.0130954	0.006242874	0.02345998	0.016534361
Median	0.01019679	-0.006320895	-0.0063991	0.010540459	0.00126945	-0.02203607
Standard Deviation	0.06710259	0.095856721	0.14345294	0.068387261	0.2569912	0.181124852
Sample Variance	0.00450276	0.009188511	0.02057875	0.004676817	0.06604448	0.032806212

The return data suggest that the characteristics of each security vary from period to period. Although not significant, we predict that these variations between the in-period and out-period will likely cause our market-neutral portfolio will fail to keep neutrality from the in-sample period to the out-of-sample period.

Methodology

(Complete steps see Appendix:B)

We have selected six securities to form 3 long-short trading pairs from Healthcare, Technology and Consumer cyclical to achieve such a goal. These three industries have different extents of sensitivity to the economic condition. For example, where Health Care is relatively defensive to the macro condition change, Consumer cyclical heavily relies on economic conditions, and Technology is sensitive to market volatility. Therefore, in each target industry, we selected a large-cap and a small stock with the intent of longing the smaller-cap companies and shorting the larger cap to increase Abnormal return.. The six companies are DaVita Inc. (DVA) and Koninklijke Philips N.V. (PHG), Methode Inc. (MEI) and Canon Inc. (CAJ), Sonic Automotive, Inc. (SAH) and Ford Motor Company (F).

We used the Fama Frech Asset Pricing Model to estimate the expected return, abnormal return α, and corresponding sensitivity factors βs for each stock by running three platforms (Python, SAS, and Excel).

1. We ran a regression of excessive returns on the three factors for each of the six securities using the in-period data. Following this, we recorded the Beta of each security. Finally, the weights of each pair were calculated based on the Beta(βi_{Mkt})
2. We regressed the three factors for each pair using the in-period factor data to confirm if we successfully created three market-neutral pairs.
3. Lastly, we average the weight of the three pairs to create a market-neutral portfolio.

Through dynamic pair trading, we achieve a market-neutral investment goal, with the correlation coefficient βi_{Mkt} determining the weight of each security within the pair. The factor exposure β_s could change over time, and the portfolio may need to be re-assessed, and the weights of the stock in each pair may be required to be reweighted in the long term.

Result

The results we obtained from Excel, Python, and SAS are consistent and share the same values. For the in-sample final portfolio, the Beta of the portfolio is close to 0, which is -4.18e-18, and the fitness of the data is 0.033, which means our portfolio is well diversified and market neutral hedged. The Alpha of estimates is 0.0176, which shares that our portfolio brings a positive return while market-neutral hedged. For the small minus big (SMB), we obtained a value of 0.829, and the P-value for SMB is 0.076 indicating the SMB is indeed statistically significant. On the other hand, we obtained a value of 0.099 for the high minus low (HML). The P-value of HML is 0.817. The P-values indicate that the result for HML is not statistically significant. Hence for the in-sample period, our primary(market-neutral) and secondary objective(small minus big) both hold, and it can generate an abnormal return.

After assigning the weight using the formula from Appendix B, we obtained data for the out-of-sample Market-Neutral Hedged portfolio. Again, the results we got from Excel, python and SAS are consistent since they share the same values. In the out of the sample, the Beta is worse than what we obtained from the in-sample. With a beta of 1.3557 and an R-square of 0.072, the Beta is statistically significant. The coefficient of return, Alpha, in the out-of-sample is -0.011, which means we now have a portfolio of slightly negative returns. The SMB and HML are now both negative and statistically insignificant. Overall, our out-of-sample portfolio is starting to fall apart in terms of our primary and secondary objectives; it has a negative return and is higher than market risk. Market neutral is not hedged; therefore, our market neutral hedge strategy did not meet our expectations.

Conclusion

Our work is consistent with our expectations and the literature we reviewed earlier in the report, such that our portfolio being a simple long-short market-neutral portfolio, cannot withstand the challenge of time. In addition, our report has not considered transaction costs, and it will be too costly to rebalance the portfolio constantly; it will futher decrease the already low excessive return of our market-neutral portfolio. We also believe a simple long-short trading pair that only considers Beta and SMB are insufficient and misleading, and more variables should be considered.

APPENDIX A: Companies Selections:

	Smaller-Cap		Larger-Cap	
SECTOR	**TICKER**	**Market Cap**	**TICKER**	**Market Cap**
Healthcare	DVA	6.70 billion	PHG	13.47 billion
Technology	MEI	1.84 billion	CAJ	30.16 billion
Consumer cyclical	SAH	1.90 billion	F	55.72 billion

APPENDIX B: Functional Form of Fama French Model

$Return\ = \alpha + \beta_1(r_m - r_f) + \beta_2(SMB) + \beta_3(HML) + \varepsilon$

- $r_m - r_f$: Market risk
- SMB: $size\ premium(small\ minus\ big)$
- HML: value premium (High minus low)

Weight calculation:

- $w_{MEI}\beta_{MEI} + w_{CAJ}\beta_{CAJ} = 0$
- $w_{SAH}\beta_{SAH} + w_F\beta_F = 0$
- $w_{DVA}\beta_{DVA} + w_{PHG}\beta_{PHG} = 0$

$$\text{Hedge Portfolio} = \frac{Weight_{MEI-CAJ} + Weight_{SAH-F} + Weight_{DVA-PHG}}{3}$$

Excel Summary Statistics

APPENDIX C: Summary Statistics of Raw Data

Summary Stats in Estimation Window	*MEI*	*CAJ*	*SAH*	*F*	*DVA*	*PHG*	*mktrf*	*SMB*	*HML*
12/31/01 - 11/30/11	0.010666697	0.008130833	0.018391546	0.008876813	0.014269937	0.002891595	0.002745833	0.004034167	0.000715
Mean	0.013095402	0.006242874	0.02345998	0.016534361	0.0061256	0.008750481	0.004277468	0.002274349	0.002378437
Standard Error	-0.006399149	0.010540459	0.001269452	-0.022036067	0.010196794	-0.006320895	0.00895	0.0013	0.0009
Median	0.143452944	0.068387261	0.256991203	0.181124852	0.067102591	0.095856721	0.046857314	0.024914245	0.02605447
Standard Deviation	0.020578747	0.004676817	0.066044478	0.032806212	0.004502758	0.009188511	0.002195608	0.00062072	0.000678835
Sample Variance	4.168665451	-0.267460712	46.291073	19.98343216	-0.144985332	1.201800027	1.117232505	-0.53796512	2.926355473
Kurtosis	1.176575854	0.008798469	5.528067869	2.910194204	0.137385071	-0.103436484	-0.552445644	0.166640952	-0.561719182
Skewness	1.005171643	0.318718842	2.714849667	1.85331029	0.336619777	0.557414386	0.2858	0.1132	0.1892
Range	-0.304154307	-0.157736067	-0.489949763	-0.579646157	-0.148274034	-0.32190092	-0.1723	-0.0528	-0.1129
Minimum	0.701017337	0.160982774	2.224899905	1.273664133	0.188345742	0.235513465	0.1135	0.0604	0.0763
Maximum	1.280003691	0.975700018	2.206985507	1.065217572	1.712392478	0.346991389	0.3295	0.4841	0.0858
Sum	120	120	120	120	120	120	120	120	120
Count									

Summary Stats 12/30/11 - 12/31/21	*MEI*	*CAJ*	*SAH*	*F*	*DVA*	*PHG*	*mktrf*	*SMB*	*HML*
Mean	0.020890793	-0.002195332	0.017988617	0.011890818	0.011600716	0.009290232	0.013116529	-6.8595E-05	-0.001908264
Standard Error	0.009768154	0.004901214	0.011259718	0.007611603	0.007125278	0.006218458	0.003571071	0.002231842	0.002751592
Median	0.023251176	-0.003659736	0.00440081	0.007673689	0.004468352	0.011276238	0.0154	0.0023	-0.0042
Standard Deviation	0.107449689	0.053913351	0.123856894	0.083727633	0.078378061	0.068403037	0.03928178	0.024550261	0.030267512
Sample Variance	0.011545436	0.002906649	0.01534053	0.007010317	0.00614312	0.004678975	0.001543058	0.000602715	0.000916122
Kurtosis	2.078758465	1.488606103	6.672092605	1.995008986	2.931044207	1.266205172	2.246400792	0.002398913	3.576264313
Skewness	-0.363838961	-0.327949959	0.663094457	-0.062527758	0.520457663	-0.431068227	-0.46158946	0.341750695	-0.28618832
Range	0.777872145	0.328873783	1.137147747	0.566433274	0.516927477	0.426889047	0.2704	0.1222	0.2216
Minimum	-0.415086299	-0.1903854	-0.523442887	-0.307334505	-0.223011192	-0.255449742	-0.1339	-0.0488	-0.1397
Maximum	0.362785846	0.138488384	0.61370486	0.259098768	0.293916286	0.171439305	0.1365	0.0734	0.0819
Sum	2.527785985	-0.265635218	2.176622674	1.438788946	1.403686632	1.124118063	1.5871	-0.0083	-0.2309
Count	121	121	121	121	121	121	121	121	121

APPENDIX D: Summary Statistics of Portfolio

Summary Stats in Estimation Window				
12/31/01 - 11/30/11	*Final Hedge Portfolio*	*Hedge Portfolio CAJ-MEI*	*Hedge Portfolio SAH-F*	*Hedge Portfolio DVA-PHG*
Mean	0.022322811	-0.00207086	0.007731693	0.057277906
Standard Error	0.010855429	0.013342106	0.016172671	0.024853571
Median	0.029705278	0.028870079	0.014462012	0.046689294
Standard Deviation	0.118915267	0.146155452	0.17716273	0.27225723
Sample Variance	0.014140841	0.021361416	0.031386633	0.074123999
Kurtosis	0.684320555	3.540818226	5.537872631	-0.339328264
Skewness	0.191964884	-1.249778142	0.79598791	0.179801811
Range	0.702356123	0.982877498	1.364259866	1.326741866
Minimum	-0.242094028	-0.658150863	-0.564505133	-0.560503986
Maximum	0.460262094	0.324726635	0.799754733	0.766237881
Sum	2.678737322	-0.248503195	0.927803172	6.87334873
Count	120	120	120	120

Appendix E: Hedge Portfolio CAJ-MEI Reg 01-11

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.29363377
R Square	0.086220791
Adjusted R Square	0.06258857
Standard Error	0.141507723
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.219173985	0.073057995	3.648442141	0.014734639
Residual	116	2.322834545	0.020024436		
Total	119	2.54200853			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.004255893	0.013087583	0.325185574	0.745626404	-0.021665715	0.030177501	-0.021665715	0.030177501
mktrf	-3.258E-16	0.30506849	-1.06796E-15	1	-0.604226583	0.604226583	-0.604226583	0.604226583
SMB	-1.451621787	0.561875094	-2.58353111	0.01102127	-2.564486209	-0.338757364	-2.564486209	-0.338757364
HML	-0.658278434	0.520956254	-1.263596373	0.208908512	-1.690097937	0.37354107	-1.690097937	0.37354107

Appendix F: Hedge Portfolio SAH-F Reg 01-11

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.281061602
R Square	0.078995624
Adjusted R Square	0.055176545
Standard Error	0.172205769
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.295049393	0.098349798	3.316485291	0.02242962
Residual	116	3.439959933	0.029654827		
Total	119	3.735009325			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.002846386	0.015926745	0.178717351	0.858471407	-0.028698539	0.03439131	-0.028698539	0.03439131
mktrf	-5.42999E-17	0.37124867	-1.46263E-16	1	-0.735304769	0.735304769	-0.735304769	0.735304769
SMB	0.941429741	0.683765739	1.376830817	0.171216007	-0.412854461	2.295713943	-0.412854461	2.295713943
HML	1.520871317	0.633970151	2.398963603	0.01803487	0.265213552	2.776529082	0.265213552	2.776529082

APPENDIX G: Hedge Portfolio DVA-PHG Reg 01-11

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.26899257
R Square	0.072357003
Adjusted R Square	0.048366235
Standard Error	0.265591605
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.638243459	0.21274782	3.016035382	0.032809944
Residual	116	8.18251246	0.070538901		
Total	119	8.820755919			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.045595311	0.024563693	1.856207492	0.065961652	-0.003056178	0.094246801	-0.003056178	0.094246801
mktrf	-3.258E-16	0.572573906	-5.69009E-16	1	-1.134054766	1.134054766	-1.134054766	1.134054766
SMB	2.995901291	1.054566526	2.840884116	0.005315406	0.907199409	5.084603173	0.907199409	5.084603173
HML	-0.564154398	0.977767181	-0.576982342	0.565069448	-2.500745502	1.372436705	-2.500745502	1.372436705

APPENDIX H: Hedge Portfolio Reg 2012-2021

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.459128035
R Square	0.210798552
Adjusted R Square	0.190562618
Standard Error	0.100904006
Observations	121

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.318186995	0.106062332	10.41704061	3.97467E-06
Residual	117	1.191249347	0.010181618		
Total	120	1.509436342			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.011399932	0.009766787	-1.167214193	0.245497215	-0.030742542	0.007942677	-0.030742542	0.007942677
mktrf	1.355756479	0.246311052	5.504245409	2.22219E-07	0.867950347	1.843562611	0.867950347	1.843562611
SMB	-0.223133497	0.394545626	-0.565545484	0.572785764	-1.004510434	0.55824344	-1.004510434	0.55824344
HML	-0.274596449	0.309801997	-0.886361134	0.377241182	-0.888143069	0.338950171	-0.888143069	0.338950171

APPENDIX I: Hedge Portfolio Reg 2012-2017

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.339517311
R Square	0.115272004
Adjusted R Square	0.067875862
Standard Error	0.084981228
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.052692415	0.017564138	2.432096749	0.074532647
Residual	56	0.40442131	0.007221809		
Total	59	0.457113726			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.011640565	0.011881967	-0.979683343	0.331456012	-0.035443005	0.012161875	-0.035443005	0.012161875
X Variable 1	0.881121193	0.37022614	2.37995403	0.020744697	0.139469115	1.622773271	0.139469115	1.622773271
X Variable 2	0.269454876	0.513672329	0.524565682	0.601955582	-0.75955445	1.298464201	-0.75955445	1.298464201
X Variable 3	0.074877173	0.514448937	0.145548308	0.884800611	-0.955687886	1.105442231	-0.955687886	1.105442231

APPENDIX J: MEI

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.541702837
R Square	0.293441964
Adjusted R Square	0.275168911
Standard Error	0.122131551
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.718601492	0.239533831	16.05872696	8.50951E-09
Residual	116	1.730269429	0.014916116		
Total	119	2.448870921			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.001307353	0.011295545	0.115740605	0.908058249	-0.021064897	0.023679604	-0.021064897	0.023679604
mktrf	0.987169768	0.263296497	3.749270416	0.000278179	0.465677879	1.508661656	0.465677879	1.508661656
SMB	1.510964734	0.484939445	3.115780226	0.00231246	0.550481059	2.471448409	0.550481059	2.471448409
HML	0.77378598	0.449623483	1.720964338	0.087923415	-0.116749982	1.664321942	-0.116749982	1.664321942

APPENDIX K: CAJ

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.66036694
R Square	0.436084496
Adjusted R Square	0.421500474
Standard Error	0.052014774
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.242699023	0.080899674	29.90152551	2.14706E-14
Residual	116	0.313842256	0.002705537		
Total	119	0.556541279			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.005262295	0.004810675	1.09387865	0.276274505	-0.004265853	0.014790443	-0.004265853	0.014790443
mktrf	0.933767424	0.112135706	8.327119483	1.873E-13	0.711668528	1.155866319	0.711668528	1.155866319
SMB	0.056132707	0.206531525	0.271787597	0.786268471	-0.352929002	0.465194416	-0.352929002	0.465194416
HML	0.109259003	0.191490762	0.570570619	0.569394339	-0.270012579	0.488530584	-0.270012579	0.488530584

APPENDIX L: SAH

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.588483188
R Square	0.346312462
Adjusted R Square	0.32940675
Standard Error	0.210449474
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	2.721771077	0.907257026	20.48493777	1.01148E-10
Residual	116	5.137521829	0.044288981		
Total	119	7.859292906			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.004100345	0.019463779	0.210665431	0.833517965	-0.034450122	0.042650812	-0.034450122	0.042650812
mktrf	2.231466657	0.453696108	4.918417008	2.90386E-06	1.332864347	3.130068966	1.332864347	3.130068966
SMB	1.647710584	0.835617418	1.971848059	0.05100717	-0.007335002	3.302756171	-0.007335002	3.302756171
HML	2.121434729	0.774763154	2.738171941	0.007152928	0.586918681	3.655950778	0.586918681	3.655950778

APPENDIX M: F

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.572007598
R Square	0.327192692
Adjusted R Square	0.309792503
Standard Error	0.150476222
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	1.277340386	0.425780129	18.80397349	5.24089E-10
Residual	116	2.626598839	0.022643093		
Total	119	3.903939225			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.000856436	0.01391705	0.061538608	0.951036172	-0.026708036	0.028420908	-0.026708036	0.028420908
mktrf	2.015128474	0.324403169	6.211802677	8.4394E-09	1.372607122	2.657649826	1.372607122	2.657649826
SMB	0.546537155	0.597485703	0.914728422	0.362231738	-0.636858565	1.729932875	-0.636858565	1.729932875
HML	0.394892701	0.553973503	0.712836804	0.477378242	-0.702321623	1.492107025	-0.702321623	1.492107025

APPENDIX O: DVA

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.45029216
R Square	0.202763029
Adjusted R Square	0.182144832
Standard Error	0.06068446
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.108646142	0.036215381	9.834178234	7.87664E-06
Residual	116	0.427182023	0.003682604		
Total	119	0.535828164			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.010116593	0.005612506	1.802509098	0.074061923	-0.000999682	0.021232868	-0.000999682	0.021232868
mktrf	0.378279618	0.130826191	2.891467028	0.004579325	0.119161862	0.637397375	0.119161862	0.637397375
SMB	0.802385196	0.240955658	3.330011855	0.001164851	0.325142155	1.279628237	0.325142155	1.279628237
HML	-0.171054586	0.223407939	-0.765660284	0.445433579	-0.613542159	0.271432987	-0.613542159	0.271432987

APPENDIX P: PHG

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.810442192
R Square	0.656816547
Adjusted R Square	0.647941113
Standard Error	0.056876155
Observations	120

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.718184763	0.239394921	74.00387838	8.09322E-27
Residual	116	0.375248047	0.003234897		
Total	119	1.093432809			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.002938477	0.005260288	-0.55861524	0.577501069	-0.013357141	0.007480187	-0.013357141	0.007480187
mktrf	1.595024237	0.122616082	13.00827928	2.08061E-24	1.352167637	1.837880837	1.352167637	1.837880837
SMB	0.387373337	0.225834279	1.715299102	0.088960402	-0.059919909	0.834666583	-0.059919909	0.834666583
HML	-0.157100984	0.209387783	-0.750287249	0.454600863	-0.571819871	0.257617904	-0.571819871	0.257617904

Python Summary Statistics

APPENDIX Q: Python Code:

```
import pandas as pd
import numpy as np
import statsmodels.api as sm

indata = r'/Users/wewake/Desktop/Fall 2022/FINE6310/Assignment 3/Data.xlsx'
t = pd.read_excel(indata)
tic_pair = [['MEI', 'CAJ'], ['SAH', 'F'], ['DVA', 'PHG']]
longshort = []
longshort_in = []

for x in tic_pair:
    Intercept = []
    Beta = []
    ls = []
    for tic in x:
        s1 = t[t['TICKER'].str.contains(tic)]
        s1.reset_index(drop=True, inplace=True)
        X = s1[['mktrf', 'SMB', 'HML']].iloc[:120]
        print(x)
        X = sm.add_constant(X)
        ret = s1.RET - s1.rf
        est = sm.OLS(ret.iloc[:120], X).fit()
        print("The model is for {}".format(tic))
        print(est.summary())
        df = pd.read_html(est.summary().tables[1].as_html(), header=0, index_col=0)[0]
        i = df['coef'].values[0]
        b = df['coef'].values[1]
        Intercept.append(i)
        Beta.append(b)
        ls.append(ret)
        ret.reset_index(drop=True, inplace=True)
        X.reset_index(drop=True, inplace=True)
        print()
        print()
    long = np.argmax(Intercept)
    short = np.argmin(Intercept)
    b_long = Beta[short] / Beta[long]
    temp = b_long * ls[long] - ls[short]
    temp1 = temp.iloc[:120]
    temp = temp.iloc[120:]
    longshort.append(temp)
    longshort_in.append(temp1)


average_ls = np.mean(np.array(longshort), axis=0)
average_ls_in = np.mean(np.array(longshort_in), axis=0)
vwretd = s1[['mktrf', 'SMB', 'HML']].iloc[120:]
vwretd.reset_index(drop=True, inplace=True)
print(vwretd.describe())
vwretd = sm.add_constant(vwretd)
est = sm.OLS(average_ls, vwretd).fit()
print("Final model taking average of longshort")
print(est.summary())
print("\n" * 5)
```

APPENDIX R: Python Output - In-sample Final Portfolio

```
Final model taking average of longshort In-Sample
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.033
Model:                            OLS   Adj. R-squared:                  0.008
Method:                 Least Squares   F-statistic:                     1.324
Date:                Fri, 02 Dec 2022   Prob (F-statistic):              0.270
Time:                        14:53:01   Log-Likelihood:                 89.701
No. Observations:                 120   AIC:                            -171.4
Df Residuals:                     116   BIC:                            -160.3
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0176      0.011      1.630      0.106      -0.004       0.039
mktrf      -5.765e-05      0.251     -0.000      1.000      -0.498       0.498
SMB            0.8285      0.463      1.790      0.076      -0.088       1.745
HML            0.0995      0.429      0.232      0.817      -0.750       0.949
==============================================================================
Omnibus:                        3.491   Durbin-Watson:                   2.131
Prob(Omnibus):                  0.175   Jarque-Bera (JB):                3.224
Skew:                          -0.220   Prob(JB):                        0.199
Kurtosis:                       3.671   Cond. No.                         44.9
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly spec
ified.
```

APPENDIX S: Python Output – Out-of-sample Final Portfolio

```
Model:                            OLS   Adj. R-squared:                  0.191
Method:                 Least Squares   F-statistic:                     10.42
Date:                Fri, 02 Dec 2022   Prob (F-statistic):           3.98e-06
Time:                        14:53:01   Log-Likelihood:                 107.87
No. Observations:                 121   AIC:                            -207.7
Df Residuals:                     117   BIC:                            -196.6
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0114      0.010     -1.167      0.245      -0.031       0.008
mktrf          1.3556      0.246      5.504      0.000       0.868       1.843
SMB           -0.2231      0.395     -0.566      0.573      -1.004       0.558
HML           -0.2746      0.310     -0.886      0.377      -0.888       0.339
==============================================================================
Omnibus:                       11.304   Durbin-Watson:                   1.850
Prob(Omnibus):                  0.004   Jarque-Bera (JB):               30.347
Skew:                          -0.080   Prob(JB):                     2.57e-07
Kurtosis:                       5.448   Cond. No.                         44.2
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly spec
ified.
```

SAS Summary Statistics

APPENDIX T: SAS Code

```
data CAJ; set week2.tech2; if ticker="CAJ";
CAJ=ret;
proc sort; by date;
data MEI; set week2.tech2; if permno="53197";
MEI=ret;
proc sort; by date;
data SAH; set week2.tech2; if ticker="SAH";
SAH=ret;
proc sort; by date;
data F; set week2.tech2; if ticker="F";
F=ret;
proc sort; by date;
data DVA; set week2.tech2; if ticker="DVA";
DVA=ret;
proc sort; by date;
data PHG; set week2.tech2; if ticker="PHG";
PHG=ret;
proc sort; by date;
data allt; merge CAJ MEI SAH F DVA PHG ; by date;
data ff; set week2.ff1;
date=dateff;
HML=HML;
SMB=SMB;
UMD=UMD;
mktrf=mktrf;
proc sort; by date;
data all; merge allt (in=a) ff; by date;
if a;
year=year(date);
month=month(date);
CAJ=CAJ-rf;
MEI=MEI-rf;
SAH=SAH-rf;
F=F-rf;
DVA=DVA-rf;
PHG=PHG-rf;
proc means;
proc model data=all;
where date<mdy(12,1,2011);
CAJ= a1 +b_mktrf1*mktrf +b_SMB1*SMB +b_HML1*HML;
MEI= a2 +b_mktrf2*mktrf +b_SMB2*SMB +b_HML2*HML;
SAH= a3 +b_mktrf3*mktrf +b_SMB3*SMB +b_HML3*HML;
F= a4 +b_mktrf4*mktrf +b_SMB4*SMB +b_HML4*HML;
DVA= a5 +b_mktrf5*mktrf +b_SMB5*SMB +b_HML5*HML;
PHG= a6 +b_mktrf6*mktrf +b_SMB6*SMB +b_HML6*HML;
fit CAJ MEI SAH F DVA PHG/outest=neutral1 gmm kernel=(bart,2,0);
test a1, a2, a3 ,a4 ,a5 ,a6, /out=p;
data mktneutral; set all; if _n_=1 then set neutral1;
mhedge1=(CAJ*b_mktrf2/b_mktrf1)-MEI;
mhedge2=(SAH*b_mktrf4/b_mktrf3)-F;
mhedge3=(DVA*b_mktrf6/b_mktrf5)-PHG;
hedge=(mhedge1+mhedge2+mhedge3)/3;
proc means;
var hedge mhedge1 mhedge2 mhedge3 CAJ MEI SAH F DVA PHG;
where date<mdy(12,1,2011);
proc corr data=mktneutral;
var hedge CAJ MEI SAH F DVA PHG ;
proc model data=mktneutral;
where date<mdy(12,1,2011);
hedge= alpha +b_mktrf*mktrf +b_SMB*SMB +b_HML*HML;
fit hedge/gmm kernel=(bart,2,0);
title "Simple Market-Neutral Hedge, In-Sample"; run;
proc model data=mktneutral;
where date>mdy(12,1,2011);
hedge= alpha +b_mktrf*mktrf +b_SMB*SMB +b_HML*HML;
fit hedge/gmm kernel=(bart,2,0);
title "Simple Market-Neutral Hedge, Out-of-Sample"; run;
proc model data=mktneutral;
where date>mdy(12,1,2011);
hedge= alpha +b_mktrf*mktrf +b_SMB*SMB +b_HML*HML + b_UMD*UMD;
fit hedge/gmm kernel=(bart,2,0);
title "Simple Market-Neutral Hedge, Out-of-Sample"; run;
ods graphics / reset width=6.4in height=4.8in imagemap;
proc sgplot data=WORK.MKTNEUTRAL; where date>mdy(12,1,2011);
reg x=MEI y=hedge /CURVELABEL="Average Hedge Portfolio" LINEATTRS=( pattern=solid)
LEGENDLABEL="Regression Line Average Hedge";
reg x=MEI y=mhedge1 /CURVELABEL="Hedge Using Just MEI and CAJ" LINEATTRS=
(pattern=dot ) LEGENDLABEL="Regression Line CAJ/MEI Hedge";
xaxis grid;
yaxis grid;
run;
```

APPENDIX U: In-sample

Simple Market-Neutral Hedge, In-Sample

The MODEL Procedure

Nonlinear GMM Summary of Residual Errors							
Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq
hedge	4	116	1.5757	0.0136	0.1165	0.0331	0.0081

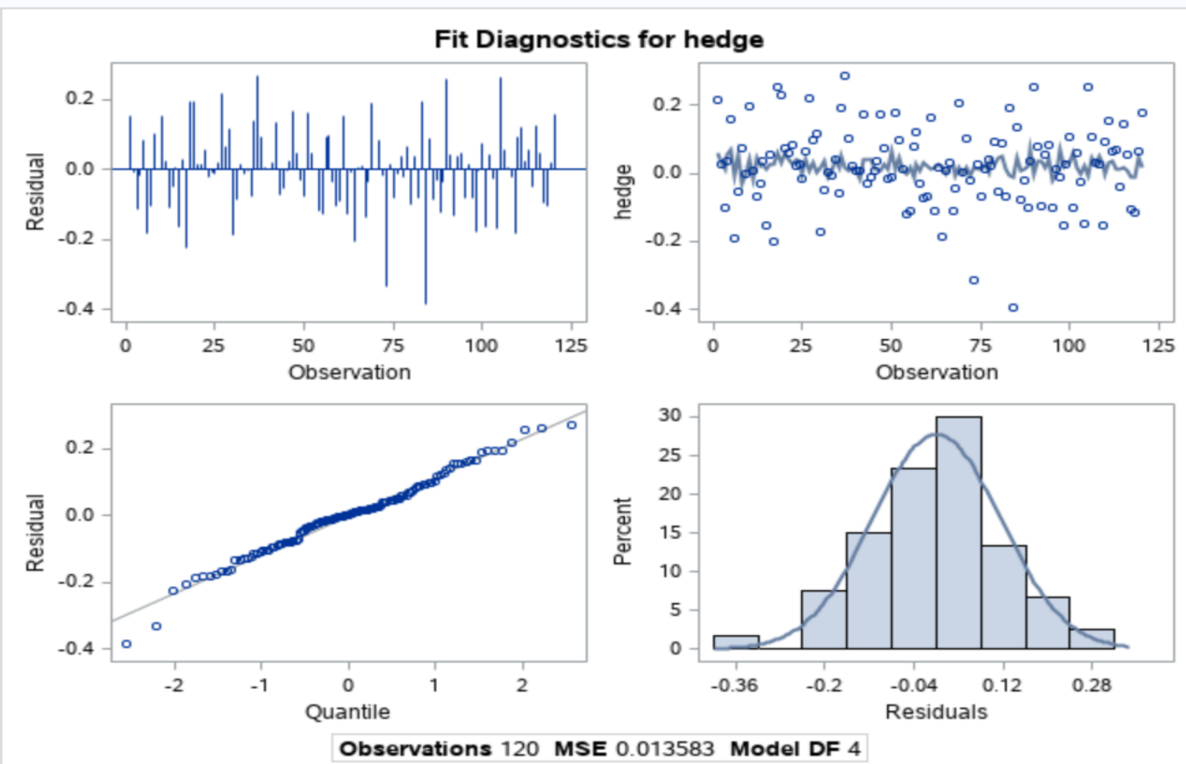
Nonlinear GMM Parameter Estimates				
Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
alpha	0.017566	0.0106	1.66	0.0993
b_mktrf	-418E-18	0.2251	-0.00	1.0000
b_SMB	0.82857	0.4121	2.01	0.0467
b_HML	0.099479	0.3855	0.26	0.7968

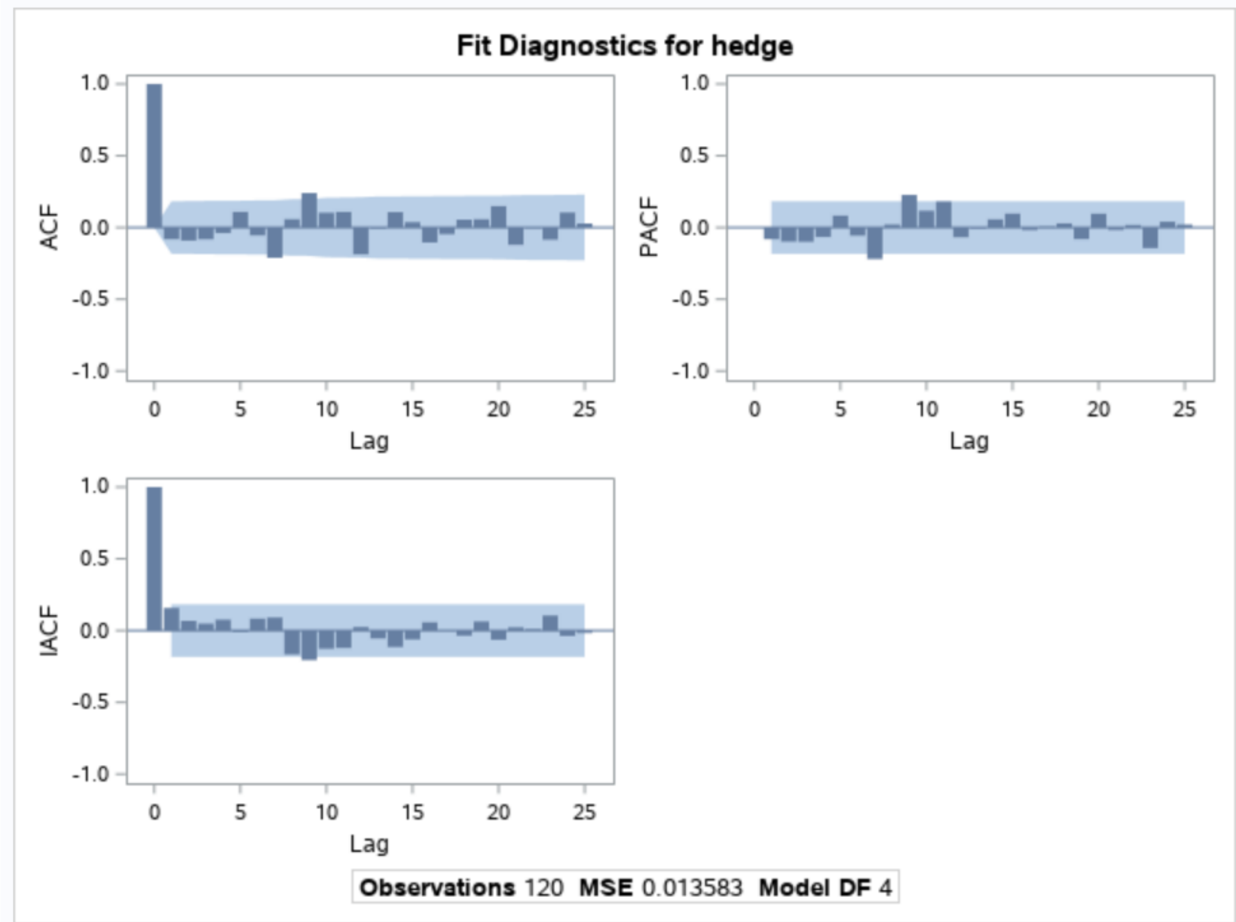
Number of Observations		Statistics for System	
Used	120	Objective	8.534E-33
Missing	0	Objective*N	1.024E-30

GMM Test Statistics			
Test	DF	Statistic	Prob
Overidentifying Restrictions	0	0.00	.

Simple Market-Neutral Hedge, In-Sample

The MODEL Procedure




Fit Diagnostics for hedge
ACF
PACF
IACF
Lag
Observations 120 MSE 0.013583 Model DF 4

APPENDIX V: Out of sample

Simple Market-Neutral Hedge, Out-of-Sample

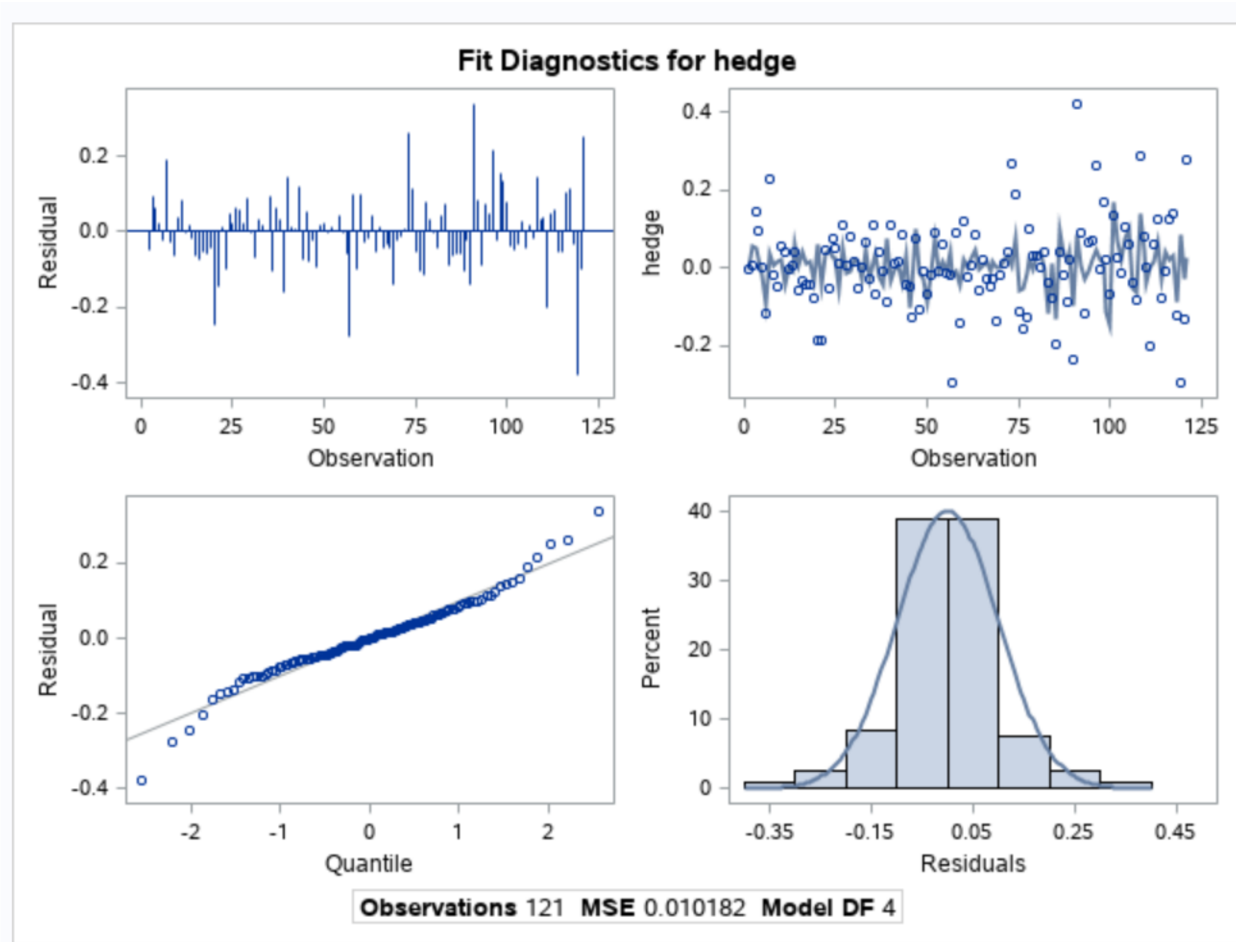
The MODEL Procedure

Nonlinear GMM Summary of Residual Errors							
Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq
hedge	4	117	1.1912	0.0102	0.1009	0.2108	0.1906

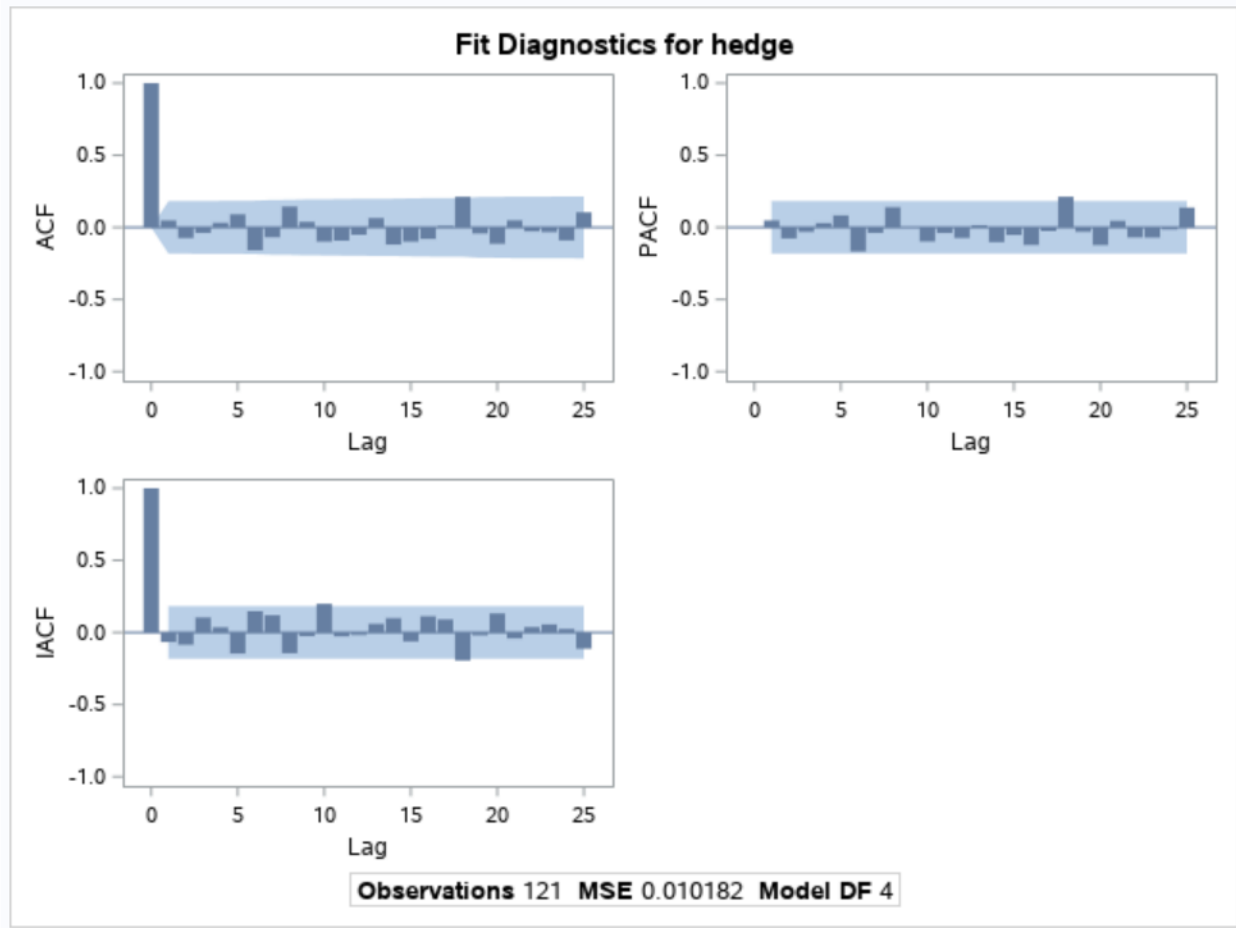
Nonlinear GMM Parameter Estimates				
Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
alpha	-0.0114	0.00881	-1.29	0.1981
b_mktrf	1.355756	0.2958	4.58	<.0001
b_SMB	-0.22313	0.4044	-0.55	0.5821
b_HML	-0.2746	0.2783	-0.99	0.3258

Number of Observations		Statistics for System	
Used	121	Objective	2.542E-32
Missing	0	Objective*N	3.075E-30

GMM Test Statistics			
Test	DF	Statistic	Prob
Overidentifying Restrictions	0	0.00	.



Fit Diagnostics for hedge
Residual
Observation
hedge
Observation
Residual
Quantile
Percent
Residuals
Observations 121 MSE 0.010182 Model DF 4


Fit Diagnostics for hedge
ACF
PACF
IACF
Lag
Observations 121 MSE 0.010182 Model DF 4

Reference

. 1) Borkovec, M., Domowitz, I., Kiernan, B., & Serbin, V. Portfolio optimization and the cost of trading. Journal of Investing, 19(2), 2010, 63-76.

. 2) Gatev, E., Goetzmann, W., Rouwenhorst, K.G. Pairs Trading: Performance of a Relative-Value Arbitrage Rule, The Review of Financial Studies, 19(3), 2006, 797–827.

. 3) Kwan, C.C.Y. "A note on market-neutral portfolio selection", Journal of Banking & Finance 23, 1999, 773-799.

. 4) Patton, A.J. "Are `Market Neutral' Hedge Funds Really Market Neutral?", The Review of Financial Studies, 22(7), 2009, 2495-2530.