Quantitative Finance MANG6299

ID: 32447248
Words: 2993

Question_1

Figure_1 demonstrates stock's risk premium against market risk premium. Market risk premium is the additional return expected on a portfolio of investments above the risk-free rate. Stock risk premium is the expected return of stock above the risk-free rate.

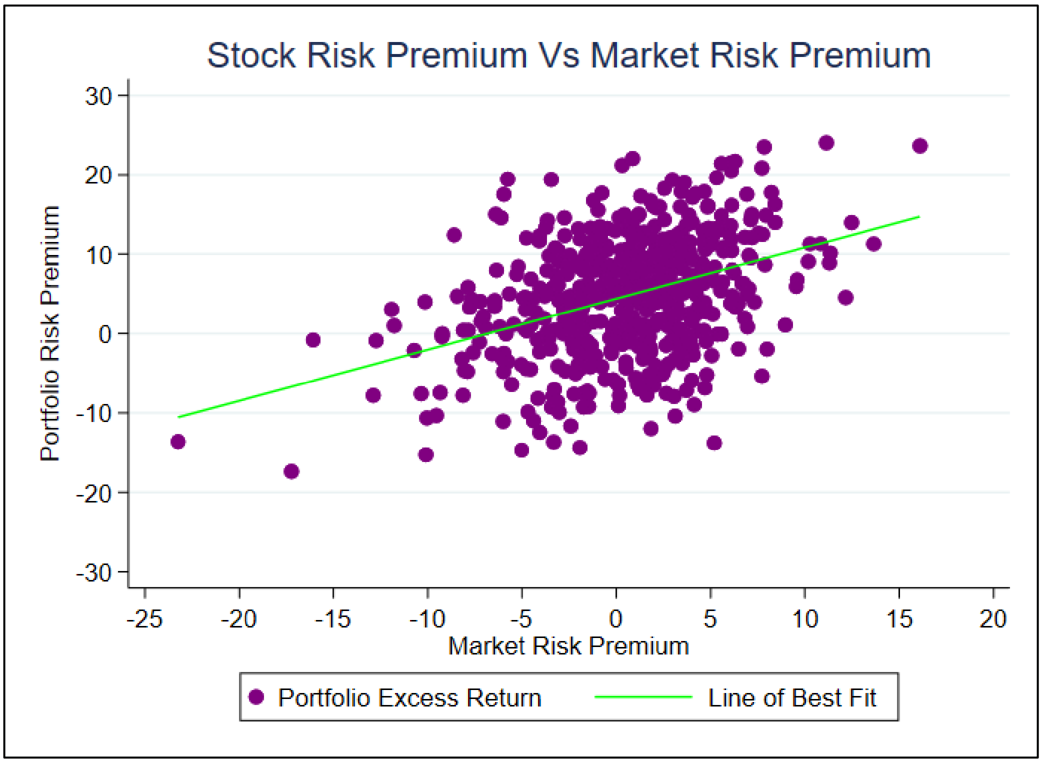


Figure_1:_Risk_Premium_Graph

There is a positive and direct correlation between the two excess returns. Hence, an increase in the excess return of the market is accompanied by increase in the excess return of portfolio.

Figure_2 shows the labelled output of the Capital Asset Pricing Model (CAPM) regression.

```
. regress e_portfolio e_mktport                               F-Test Value

      Source |       SS           df       MS      Number of obs   =       576
-------------+----------------------------------   F(1, 574)       =    105.26
       Model |  4766.46774         1  4766.46774   Prob > F        =    0.0000
    Residual |  25993.3155       574  45.2845217   R-squared       =    0.1550
-------------+----------------------------------   Adj R-squared   =    0.1535
       Total |  30759.7832       575  53.4952751   Root MSE        =    6.7294

------------------------------------------------------------------------------
 e_portfolio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   e_mktport |   .6428565     .06266    10.26   0.000     .5197856    .7659274
       _cons |   4.383935   .2826824    15.51   0.000     3.828717    4.939153
              Alpha (Constant) Standard Error T-test   Two-tail p-value test
            Beta (Coefficient)
```

Figure_2:_CAPM_Regression_Labelled_Diagram

The CAPM Equation of regression is written as:

$$Rp = ai + Bi(Rm - Rf)$$

By inserting the output and $(Rm - Rf)$ for xt, the equation is simplified:

$$\hat{y} = 4.384 + 0.6429\text{xt}$$
$$(0.2827)(0.06266)$$

Numbers in parenthesis are the standard errors, which provide a measurement of uncertainty in the estimated values report given by:

$$s = \sqrt{\frac{\sum(Uthat)^2}{T-2}}$$

Beta coefficient reports how much the dependant variable is expected to change when the independent variable changes. Figure_2 reports that portfolio excess return increases by 0.6429% when market excess return increases by 1%. In CAPM, stocks with high Beta are riskier, therefore, average returns for the risk taken are expected to be higher. The obtained Beta (0.6429) is positive but less than one, indicating the portfolio is defensive. A beta of 1 means the portfolio is in line with SP500 and negative beta indicates the portfolio is uncorrelated with SN500, yet it would be a great asset as it diversifies the portfolio.

The alpha of 4.384 suggests portfolio has systematically beaten the market and earns excess returns, therefore it performs well. The beta least square of $\bar{\alpha}$ and $\hat{B}$ are unbiased, thus the estimated value for coefficient is equal to their true value.

The purpose of t-values is to test the null hypothesis that independent variables have no impact on the dependant variable, therefore it has a coefficient of zero. To reject this at a 95% confidence level, at t-value greater than 1.96 is required, which in this case hypothesis is rejected as the t-values obtained are 10.26 and 15.51, suggesting independent variable does have an impact as they are statically significant. T-test is used for single hypotheses only, using the following equation:

$$t - stat = \frac{\hat{B} - B^*}{SE(\hat{B})} = \frac{0.6429 - B^*}{0.06266}$$

The confidence intervals reported are $0.5198 < B^* < 0.7659$. This suggests in many repeated samples, true value of beta will be within these intervals 95% of the times.

Fama-French-3 model is tested by adding two variables.

$$Rp = ai + Bi(Rm - Rf) + SMB + HML$$

$SMB = \text{Small Minus Big Factor}$

$HML = \text{High Minus low Factor}$

$$\hat{y} = 4.219 + 0.8629\text{x1} - 0.7781\text{x2} + 0.5511\text{x3}$$
$$(0.2617) \quad (0.06124) \quad (0.0887) \quad (0.0915)$$

Next, adding two new variables, Fama-French-5 is tested.

$$Rp = ai + Bi(Rm - Rf) + SMB + HML + RMW + CMA$$

$$\hat{y} = 4.025 + 0.9683x1 - 0.8604x2 + 0.0605x3 - 0.1774x4 + 1.177x5$$

$$(0.260) \quad (0.0630) \quad (0.0909) \quad (0.1189) \quad (0.1222) \quad (0.1865)$$

Additional macroeconomic factors are:
$CMA = Concervative\ minus\ aggressive\ factor$
$RMW = Robust\ minus\ weak\ factor$

F-test is required for joint significance since there is more than one coefficient to be tested simultaneously in FF models. Summary of these regressions are embedded in Table_1.

Table_1:_Regressions_Results

	e_mktport	SMB	HML	RMW	CMA	_cons	F-Stat	R^2	Adj R^2	Root MSE
CAPM	0.6429					4.384	105.3	0.155	0.1535	6.729
	(0.0627)					(0.283)				
FF3 FM	0.8629	-0.7721	0.5510			4.219	79.17	0.2934	0.2897	6.164
	(0.0611)	(0.0887)	(0.0914)			(0.262)				
FF5 FM	0.9683	-0.8604	*0.0605	*-0.1774	1.177	4.025	60.16	0.3454	0.3397	5.943
	(0.0630)	(0.0909)	(0.1189)	(0.1222)	(0.1856)	(0.260)				
FF5 Robust	0.9684	-0.8604	*0.0605	*-0.1774	1.1772	4.025	65.94	0.3454		5.943

Note: (*) Indicates the p-value is NOT statically significant at 5% level.

Two-tail p-values tests the hypothesis that each coefficient is different from zero. It represents the probability of obtaining a value as extreme as the one measured in a collection of random data in which the variable had no effect. P-value of 5% or less is the accepted point at which to reject the null hypothesis. This means there is only 5% probability that the results obtained would have come up in a random distribution. Therefore, assuming the model is correctly specified, it is concluded with a 95% confidence level, the variable is having some impact.

The coefficient value shows the size of the impact of that independent variable has on the dependant variable (holding all other independent variables constant) and the sign on the coefficient (positive or negative) provides the direction of the effect. For example, SMB (-77.8%) would have a larger but negative impact whereas HML (55%) has a smaller yet positive influence.

Conclusively, FF3 model shows a better fit than CAPM since the added variables are statistically significant. Therefore, SMB and HML do impact the portfolio. This is proven by increased R^2 value. Table_1 shows Root MSE of 6.727 in CAPM has reduced to 6,164 in FF3 model, indicating FF3 is a better model as the standard deviations has reduced.

In FF5 regression, the F-statistic has dropped to 60.16 from 79.17 of FF3, making it less useful. The p-values of HML and RMW are 61.1% and 14.7%, respectively, making them statistically insignificant. These values are outside the 95% confidence level, making them useless for the interpretation. The F-test statistic for all three models are large and significant, suggesting a statistically significant relationship between the variable and independent variables.

Table_1 shows the standard errors have become larger in FF5 relative to the previous tests. For example, SE of HML has increased from 9.411% (in FF3) to 11.89% (FF5). Results *may* have obtained large standard errors relative to the values they should take.

R^2 is the goodness of fit statistics test to see how precisely the sample regression function (SRF) fits the data. By the addition of regressors, the R^2 has increased from 15.5% in CAPM, to 29.34% in FF3 and to 34.54% in FF5 model. Hence, CAPM explains fewer variability of the response data around its mean, whereas FF5 explains more. Figure_3 illustrates the theory.

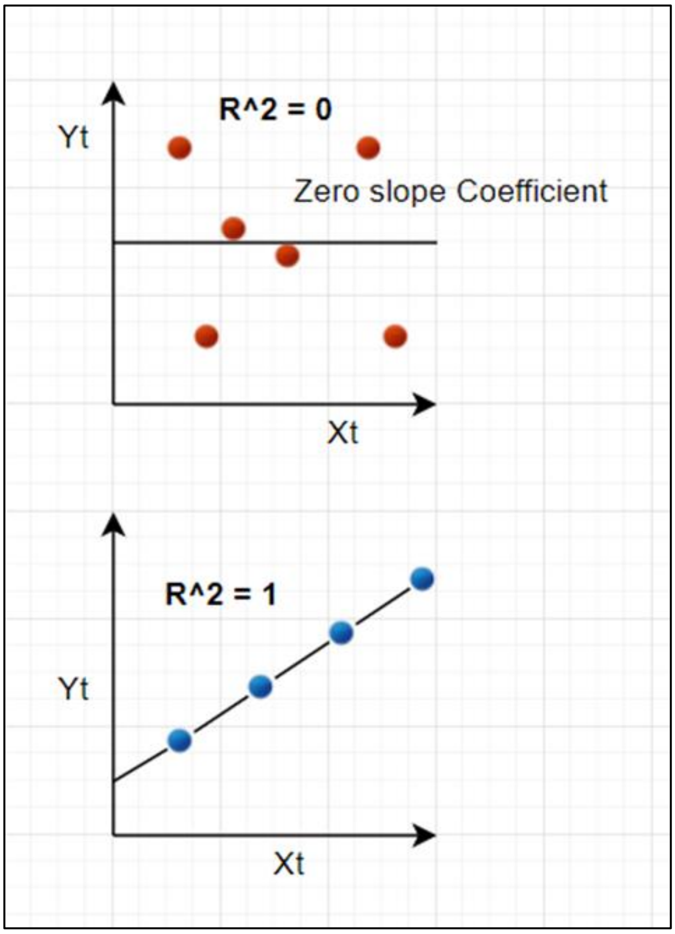


Figure_3:_R-Squared_Illustriation

Root MSE shows the standard deviation of the regression, therefore a lower number is more ideal. FF5 has proven to indicate a better fit for the model with RMSE of 5.943. Adjusted R^2 displays the same as R^2, however, adjusted by number of cases and number of variables. When the number of variables is small, and the number of cases is large then Adjusted R^2 is closer to R^2. The Adjusted R^2 increased from 0.15 in CAPM to 0.34 in FF5, therefore, FF5 is a better model.

Test for heteroskedasticity is completed by visual checking of residuals in Figure_4. Residuals are the error terms or the differences between the observed value of the dependent variable and the predicted value.

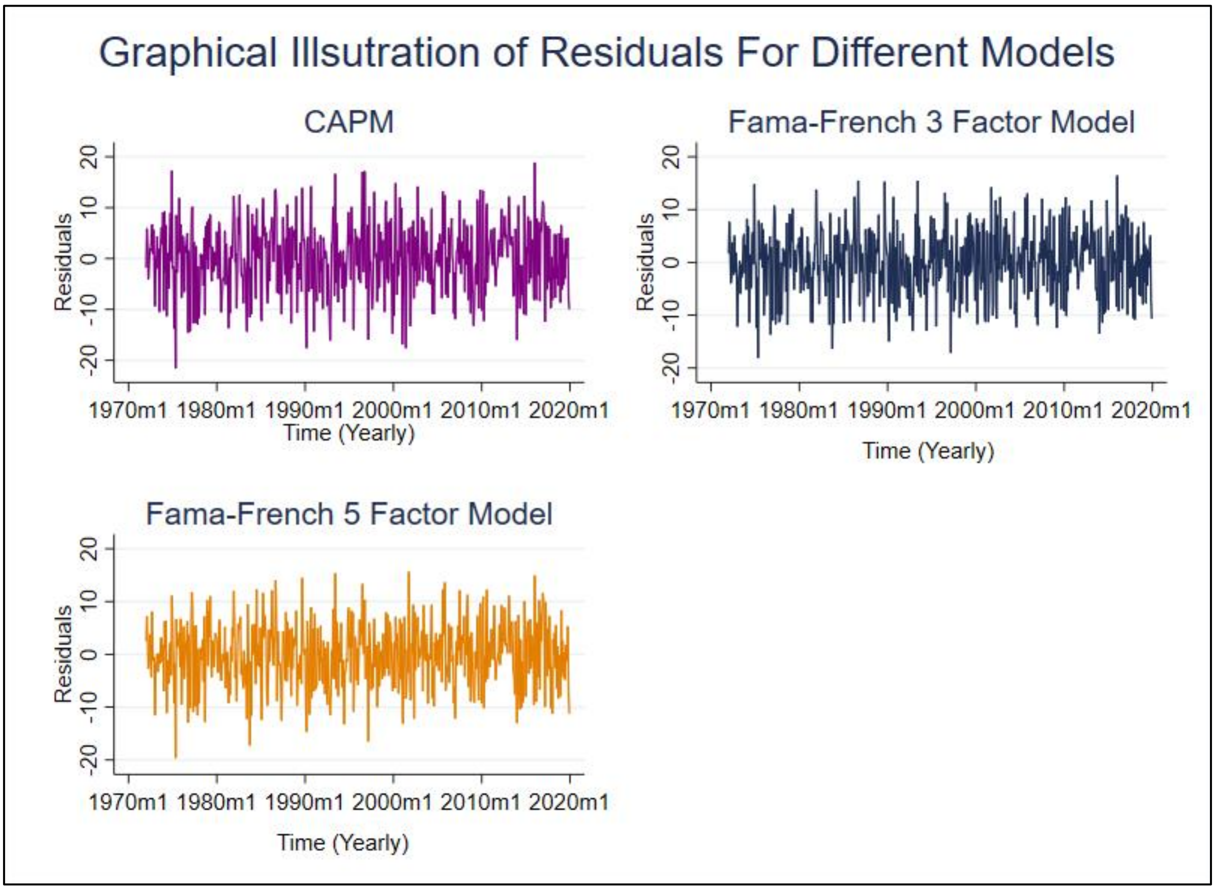


Figure_4:_Visual_Checking_For_Heteroskedasticity

Interpretation is hard from Figure_4. Thus, Breusch-Pagan test (H(0): constant variance) is implemented. Consequently, all the p-values generated are above 5%, advising to accept the null, hence there is no heteroskedasticity in the data. *Assuming* there is heteroskedasticity, the issue can be fixed by implementing a robustness command. (Example in Table_1: FF5_Robust).

The Portmantua's (H(0): white noise), and Breusch-Godfrey Lm (H(0): no-serial correlation), are used to test for autocorrelation. The report showed significant values of above 5%, indicating there is no autocorrelation. 12 lags are used since it is the optimum number of lags. A complementary test of Durbin-Watson carried out also resulted in all values close to 2, meaning the null (H(0): no serial correlation), is accepted.

Wald-test is applied to test for structural break (H(0): No Structural break), with middle date of 1995m12. P-values reported are significantly greater than 5%, which is an indication of no structural break. Chi2 value demonstrates how strongly the variables are linked, and p-value shows the probability of it begin true.

Ramsey's test (H(0): No omitted variables) is used, and based on the report, p-values are significantly larger than 5%, hence there are no omitted variables.

There is no multicollinearity in regressions, however, on FF5, there is collinearity of 0.6917 between HML and CMA. A Variance-Inflation-Factors is calculated to be 1.528. Since this number is less than 5, it is not severe, and hence it is concluded there is no multicollinearity.

$$VIF = \left(\frac{1}{1-R^2}\right) = \left(\frac{1}{1-0.3454}\right) = 1.5277. \qquad VIF \leq 5, \qquad Not\ severe.$$

Jarque-Bera tests for normality, and from the obtained data, it can be argued the residuals follow a relatively normal distribution. The histograms in Figure_5 also show a bell-shaped graph, and since the theoretical line (Navy) is very similar to the actual data points (Orange), it can be concluded there are no violations.

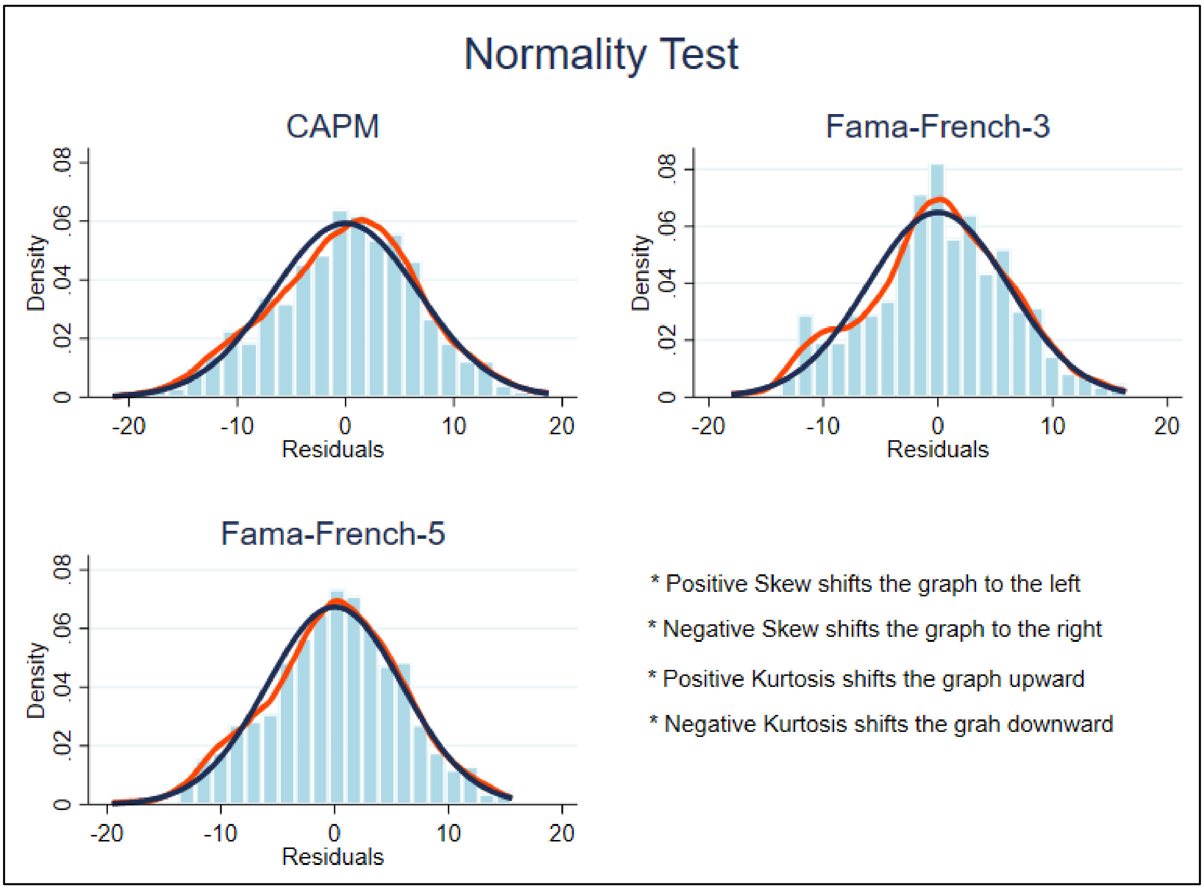


Figure_5:_Normality_Test_Histogram

Question_2

Initially, the variable can be plotted to check for stationarity. From Figure_6, the Time Series (TS) illustrates a mean-reverting and stable variance over time. Conclusively, it is white noise, thus no unit root is present. Data points are small, hence it is assumed the data is already in log from.

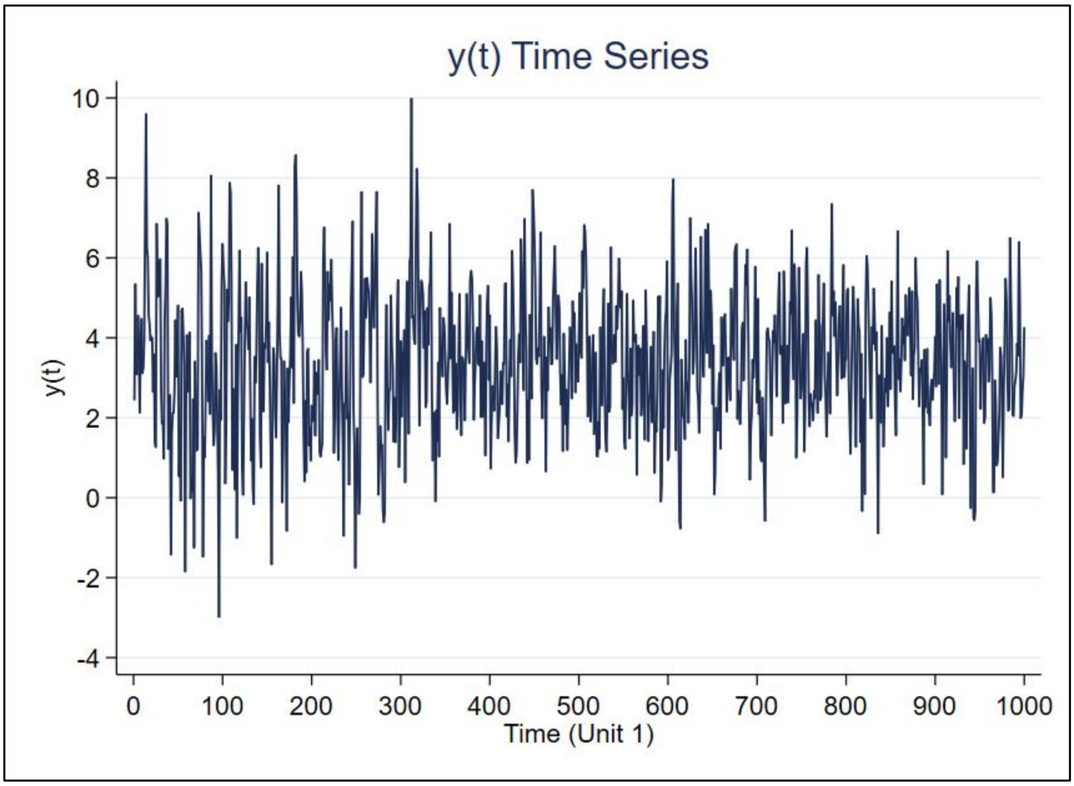


Figure_6:_Unit_Root_Visual_Check

Three different stationarity tests of ADF, KPSS and PP are executed. T-stats of ADF and PP (-19.71 and -19.34, respectively) are greater than all Critical Values (CVs), and the p-values are zero, making them statically significant, consequently, the null that the data has unit root can be dismissed. KPSS tests for stationarity rather than unit root. Table_2 shows the t-stat is lower than the CVs, concluding the series is stationary.

Table_2:_Stationarity_Tests

			Critical Values				
	Null(H0)	Test Statistic	1%	5%	10%	P-Value	Stationary
ADF	Unit Root	-19.714	-3.43	-2.86	-2.57	0	Yes
PP	Unit Root	-19.341	-3.43	-2.86	-2.57	0	Yes
KPSS	Stationary	0.0926	0.216	0.146	0.119		Yes

The ACF and PACF are used as complementary and further evidence of stationarity. ACF of a stationary process shows a weak persistence of series by geometrically decaying, which holds true in Figure_7. ACF is very close to zero apart from the first single peak, and the same is true for PACF, hence the data is stationary. The ACF and PACF can suggest the number of q(MA) and p(AR) lags for a suitable model to fit the data.

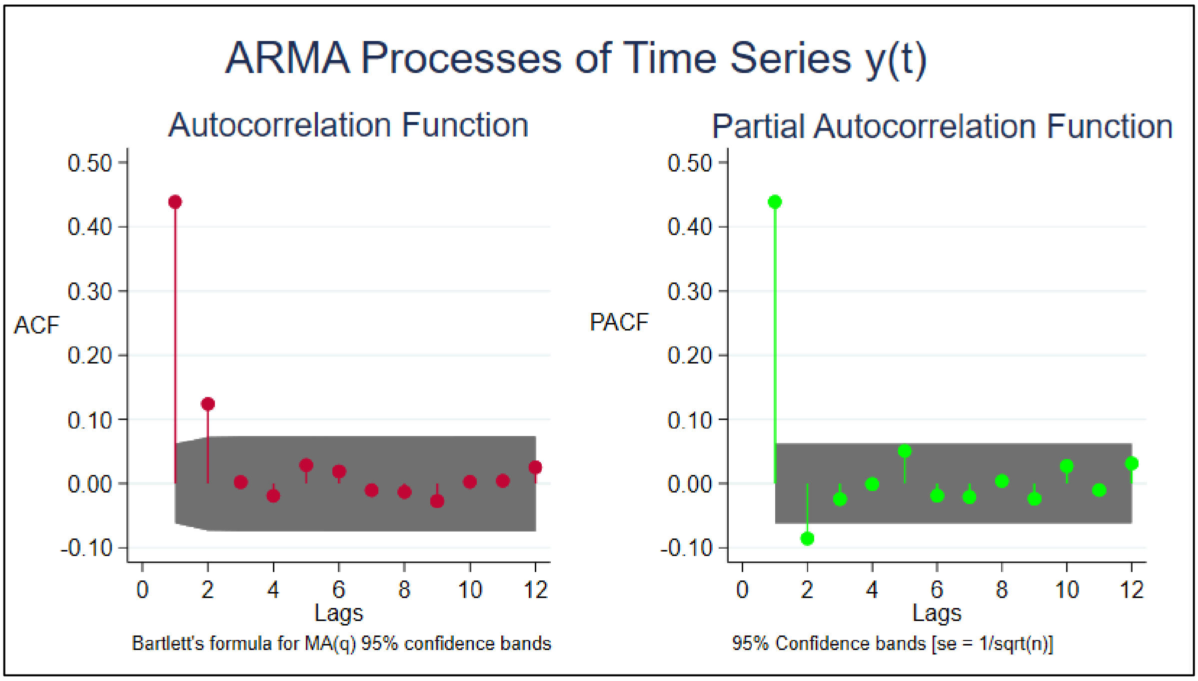


Figure_7:_ARMA_Processes_of_TS

Twelve different ARMA models are tested with different lags for AR and MA (See Table_3). Models 1-8 are the primary and 9-12 are the alternative tests.

SBIC and AIC are used for model selection. ICs consist of functions of RSS and penalty due to losses in degrees of freedom for extra added parameters. This is because adding a lag/variable increases the penalty term but reduces the RSS. AIC is related to the RSS, and SBIC to number of parameters k. AIC is efficient but inconsistent, whereas SBIC is very consistent yet inefficient.

ARMA models are also tested on Log-likelihood to measure the goodness of fit, number of significant coefficients and Sigma-squared for volatility. R^2 can also be utilized, however, it usually chooses the largest model, thus it is not desired. Ljung-Box with four lags (first four autocorrelations) is implemented to test for white noise processes.

Table_3:_ARMA_Models

Arma Models	1 (1,0)	2 (2,0)	3 (3,0)	4 (4,0)	5 (1,1)	6 (2,1)	7 (1,2)	8 (2,2)	9 (0,1)	10 (0,2)	11 (3,2)	12 (3,4)
y hat	3.383	3.383	3.383	3.383	3.383	3.383	3.383	3.383	3.383	3.383	3.383	3.383
std.err	0.0886	0.0814	0.0796	0.0795	0.0834	0.0797	0.0822	0.0796	0.0712	0.0798	0.0844	0.0807
u	1.899	2.060	2.110	2.111	2.355	1.683	2.850	2.187			1.741	
yt-1	0.4388	0.4764	0.4743	0.4743	0.3038	0.6764	0.1576	0.5015			1.230	-1.551
std.err	0.0271	0.0311	0.0312	0.0312	0.0663	0.3496	0.2149	0.6838			0.2805	0.2437
yt-2		-0.0854	-0.0739	-0.0740		-0.1739		-0.1480			-1.048	-0.8315
std.err		0.0327	0.0363	0.0363		0.1496		0.2144			0.2380	03091.
yt-3			-0.0240	-0.0238							0.3032	-0.2054
std.err			0.0313	0.0341							0.0831	0.1967
yt-4				-6E-04								
std.err				0.0321								
ut-1					0.1693	-0.2013	0.3159	-0.0279	0.4213	0.4677	-0.7609	2.029
std.err					0.0712	0.3539	0.2166	0.6838	0.0289	0.0312	-2.720	0.2410
ut-2							0.0785	0.0647		0.1402	0.6328	1.727
std.err							0.0959	0.1363		0.0302	3.500	0.3880
ut-3												0.8637
std.err												0.2801
ut-4												0.2204
sts.err												0.0748
Obs.	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
AIC	3750	3745	3746	3748	3746	3746	3747	3748	3762	3746	3748	3749
SBIC	3765	3764	3771	3778	3765	3771	3771	3777	3777	3765	3782	3793
Qk=4	9.542	1.502	0.6419	0.6183	2.467	0.6549	1.626	0.7535	19.46	159.3	1.016	1.011
p-value	0.0489	0.8263	0.9583	0.9610	0.6504	0.9568	0.8041	0.9446	0.0006	0.000	0.9074	1.000
Log Likelihood	-1872	-1868	-1868	-1868	-1869	-1868	-1868	-1868	-1878	-1869	-1867	-1866
Sigma^2	1.573	1.567	1.567	1.567	1.568	1.567	1.567	1.567	1.582	1.736	1.565	1.563
# of Sig Coeff	1	2	2	2	2	0	0	0	1	2	5	6

Reject	Yellow	Optimum choice	Green
1% significant	Blue		
5% significant	Red		

Due to many numbers of models, colour coding is used instead of (*) to fit the values. other models were also tested which were not statistically significant, hence they were eliminated.

AIC and BIC of 3745 and 3764 suggest ARMA(2,0) is the optimum model as it has the lowest errors, whereas Sigma-Squared and Log-likelihood suggest ARMA(3,4) is ideal. Sigma-Squared of 1.563 suggests ARMA(3,4) is less volatile, this model also has six significant coefficients to describe the data and consists of MA lags, making it invertible.

Ljung-Box tests whether any autocorrelation of TS is different from zero. It is a test for overall randomness based on lags with the H(0):white noise. Models 1, 9 and 10 are not statistically significant thus TS is not white noise. However, ARMA(3,4) shows 100% randomness of TS.

AC comparison in Figure_8 shows both models 2 and 12 act almost identically with similar numbers of data points outside the 6% region. Hence, increasing the number of lags does not necessarily improve the forecasting, meaning ARMA(2,0) is the better model.

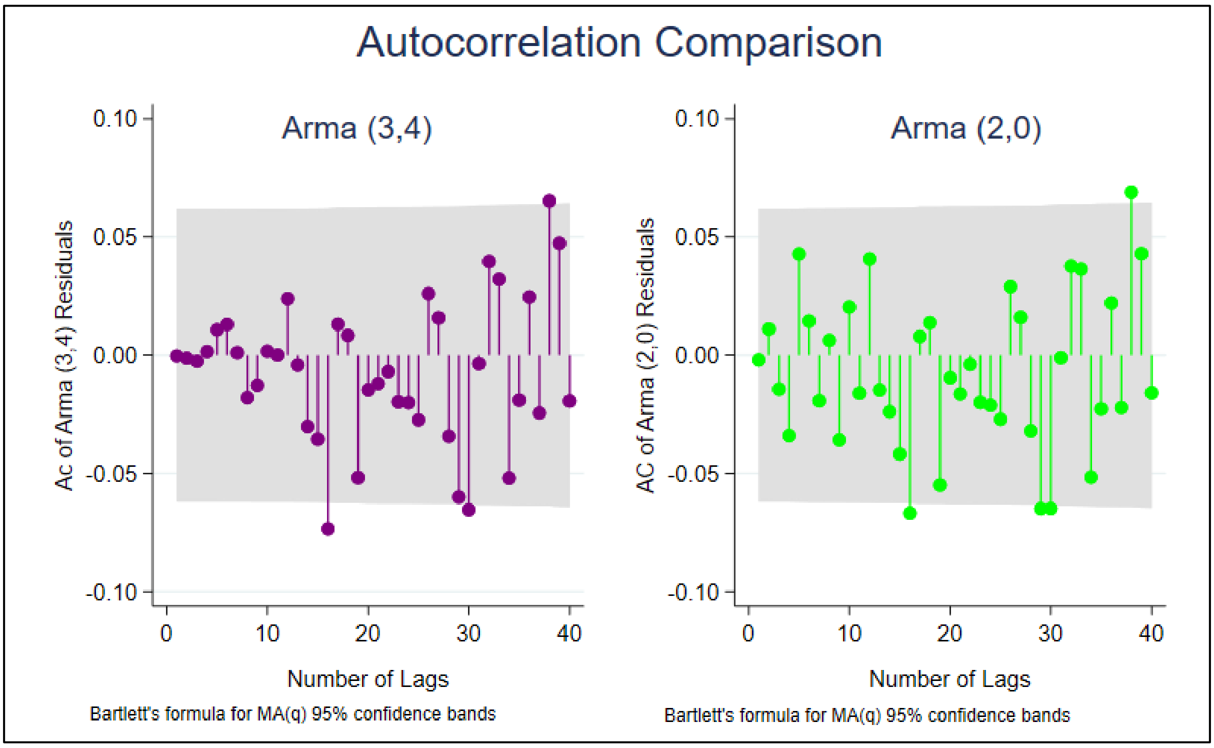


Figure_8:_Autocorrelation_Comparison

Question_3

To estimate the ARCH/GARCH models, a diagnostic test is implemented to check for the presence of ARCH effects. Based on the results obtained in Appendix_2 the *H(0):No arch effect* is rejected resulting in presence of ARCH effect. Seven lags are used to remove daily effects on the residuals. Figure_9 illustrates the presence of ARCH, known as volatility pooling.

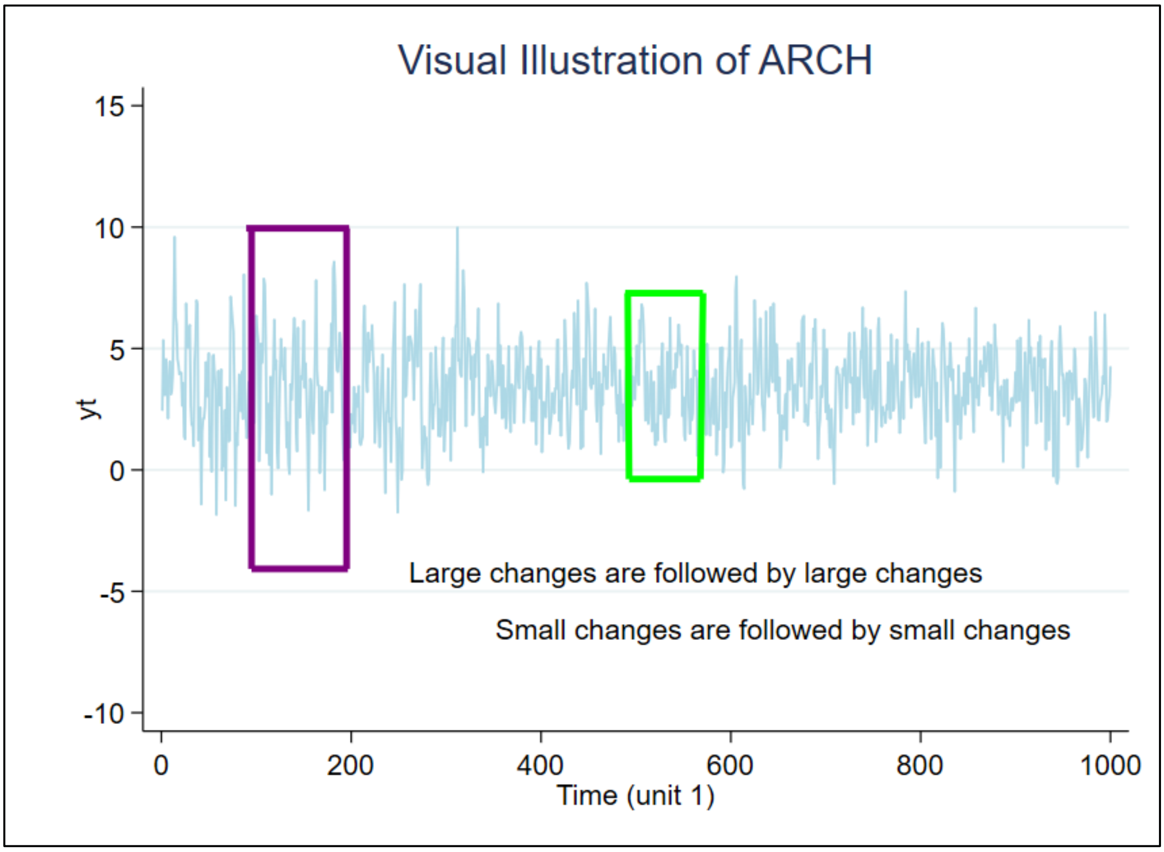


Figure_9:_Illustration_of_ARCH

An ARCH(q) model is required to satisfy the positivity conditions, that is the conditional variance should only take positive numbers. The H(0):$a0 = 0$ is rejected, meaning the constant is positive, yet the null $a1 = 0$ is accepted, meaning $a1 \geq 0$. For example, from Table_5, ARCH(1), $a0 > 0, and\ a1 \geq 0$, hence the positivity condition is satisfied. Whereas, ARCH(3), $a0 > 0\ and\ a1 \geq 0$, yet, $a2 < 0$, therefore, the positivity condition has not been satisfied.

ARCH(q) also requires a stationarity condition, that is the sum of the coefficients are less than one. For example, ARCH $a1 + a2 + a3 = 0.5139 < 1$, therefore, the stationarity condition is satisfied. Stationarity test has a great importance. If this requirement has not been met, the conditional variance will increase and "explode".

Graphical representations of standardised residuals in Figure_10 show that there is little volatility clustering, suggesting that ARCH(1) model is correctly specified. The predicted conditional variance is also plotted in Figure_10. The conditional variance is high from time 0 to 300, and low from 800 to 1000. Var of different models can be calculated using the following formula:

$$ARCH(1)\ Examples: \qquad Var(et) = \frac{a0}{1 - (a1 + \cdots an)} = \frac{2.337}{1 - 0.234} = 3.087$$

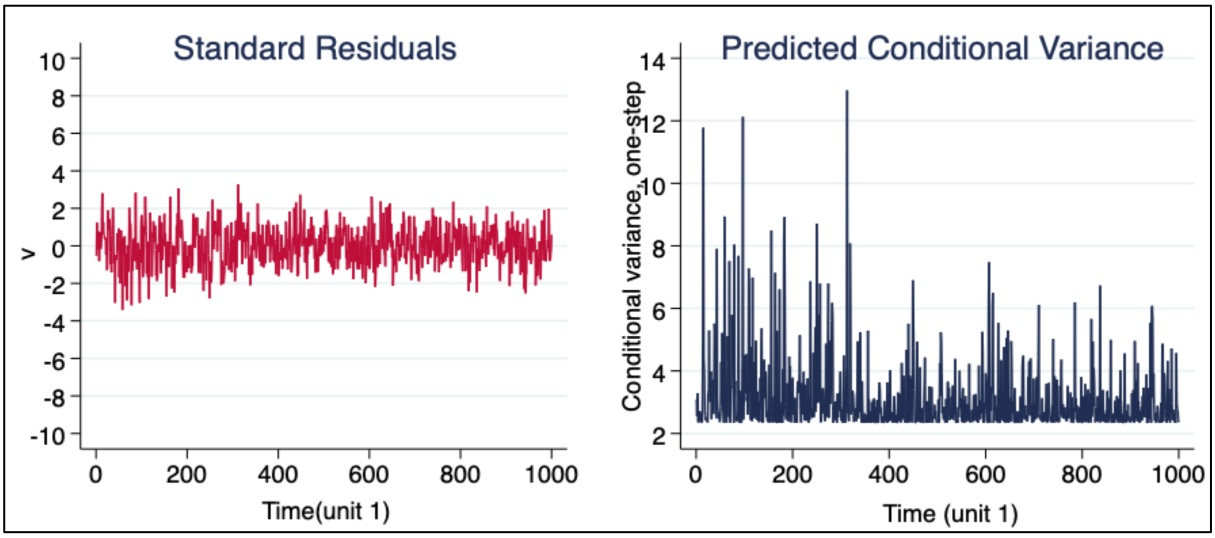


Figure_10:_ARCH(1)_Graphs

Histogram in Figure_11 shows standardised residuals are normally distributed. This is in line with the p-value of 5.9%, hence the null hypothesis (H0: Normality) of Jarque-Bera is accepted. Concluding, the SEs are consistent and unbiased. Other models showed similar results.

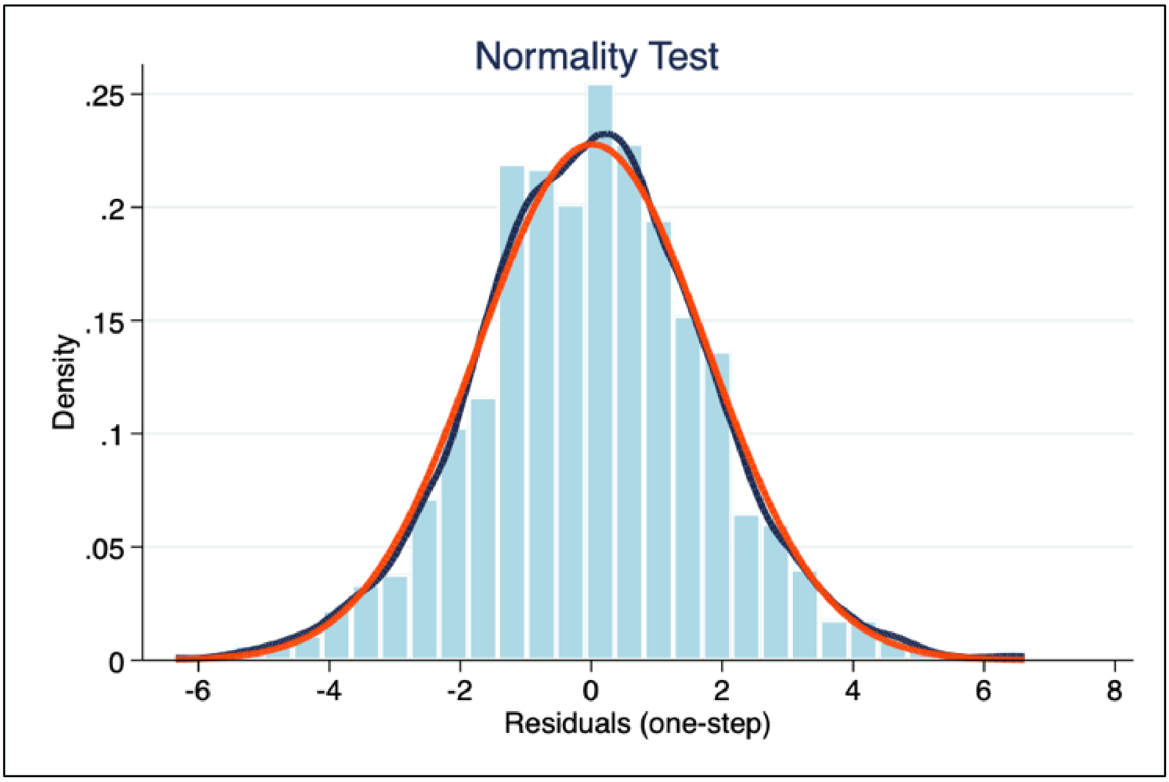


Figure_11:_Normality_Test

Positivity and stationarity conditions are also required for GARCH(1,1) and GARCH(2,2) models. If these conditions are met, then GARCH can be interpreted as ARCH(∞), else the model is not suitable for forecasting. To test for positive constant term and stationarity, the commands in Appendix_5 can be used. These tests involve two steps to break down the two-sided tests towards a single-sided test. Results are provided in table_4.

Table_4:_Two_Step_Positivity_and_Stationarity_Tests

			GARCH (1,1)	GARCH (2,2)
Positivity	2 sided	Chi2	41.24	1.32
		p-value	0	0.2503
	1 sided	H0	0	0
		p-value	6.71E-11	0.1251
Stationarity	2 sided	chi2	38.999	1.34
		p-value	0	0.247
	1 sided	H0	0	0
		p-value	2.13E-10	0.1234

Table_4 shows the positivity and stationarity for GARCH(1,1) are satisfied, whereas GARCH(2,2) fails to do so. Table_5 shows the results for eight different models.

Table_5:ARCH_&_GARCH_Models

	ARCH (1)	**ARCH (3)**	**GARCH (1,1)**	**IGARCH**	**GARCH (2,2)**	**EGARCH**	**GJR**	**GARCH-M**
μ	3.364***	3.364***	3.359***	3.411***	3.383***	3.371***	3.349***	3.171***
	-0.05592	0.05619	0.05587	0.0526	0.0533	0.0564	0.05577	0.19047
δ								0.058
								0.0661
a1	0.243***	0.2513***	0.2487***	0.0295***	0.2093***	0.3897***	0.2092***	0.2469***
	0.0505	0.05101	0.0509	0.00875	0.0511	0.6776	0.642	0.05142
a2		-0.02808			(-0.184)***			
		0.03101			0.0501			
a3		0.2907						
		0.03441						
β1			-0.0981	0.9705***	0.6393***	0.1013	-0.1127	-0.09276
			0.1193	0.00871	0.1569	0.1886	0.1184	0.1205
β2					0.3254**			
					0.1514			
γ**1**						0.02745	0.06715	
						0.4701	0.0868	
a0	2.337***	2.314***	2.623***	0.00319	0.02928	0.9764***	2.678***	2.607***
	0.1388	0.1789	0.4085	0.00409	0.0255	0.2105	0.40916	0.4068
# of Obs.	1000	1000	1000	1000	1000	1000	1000	1000
AIC	3924	3926	3925	3928	3906	3932	3926	3926
SBIC	3938	3951	3945	3943	3935	3956	39501	3950
likelihood	-1959	-1958	-1958	-1961	-1947	-1961	-1958	-1957
# of sig coeff	3	3	3	3	5	3	3	3
Positivity	Yes	No	Yes	Yes	No	Always	No	No
Stationary	Yes	Yes	Yes	No	No	Yes	Yes	Yes

If the summation of coefficients is close to 1, then the model is ought to perform well. ARCH(3) has a summation of coefficients of 0.5139, whereas integrated GARCH, IGARCH, equals 1. This indicates the shocks to conditional variance are strongly persistent for IGARCH. Ultimately, a large sum of coefficients means a big positive/negative variance is forecasted. Note, only if the series are non-stationary, IGARCH can be used, otherwise it is unreliable. This can be found since the stationarity condition has not been met.

Since the modelling of EGARCH is achieved using logs, the signs do not matter as the volatility is always positive. EGARCH's likelihood is highly nonlinear and often it fails to reach the local optimum. the positive sign of $\gamma 1$ in GJR and EGARCH means a positive shock has a larger effect than a negative shock.

Disadvantage of ARCH/GARCH models are their symmetry, meaning the absolute value of shocks are important but not the sign. Asymmetric models allow for situations when the unexpected fall in price of an asset (bad news) has a stronger impact on future volatilities than the unexpected surge (good news) of same/similar magnitude.

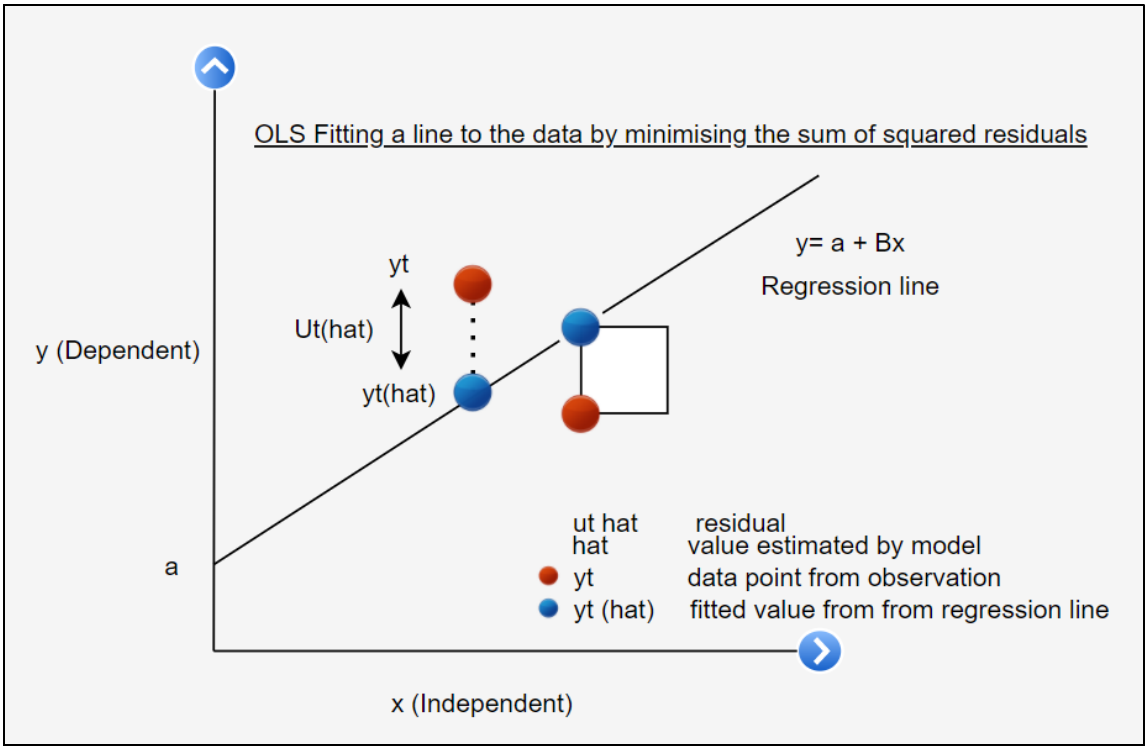
GARCH(2,2) does not satisfy the positivity nor the stationarity conditions, hence it is not a good model. GARCH-M and GJR models fail to satisfy the positivity test. Although this is not ideal, the models can still be used as stationarity condition, which is a more important requirement, is satisfied. According to likelihood parameter, GARCH-M is the best model as it has the largest value with -1957. AIC and SBIC however, suggest ARCH(1) is the optimum model for this TS as it has the lowest number of 3924 and 3938, respectively.

Conclusively, alternative models (2,4,5 and 7) introduced did not fit the data better. ARCH(1) has the lowest AIC and SBIC, and being one of the only two models satisfying both positivity and stationarity conditions, and the SEs are much lower for ARCH(1), therefore it is the best model.

Words: 2993

Appendices

Appendix_1: Ordinary Least Squares



Appendix_2: ARCH Effects

```
. estat archlm, lags(7)
LM test for autoregressive conditional heteroskedasticity (ARCH)
```

lags(p)	chi2	df	Prob > chi2
7	49.715	7	0.0000

Chi value is statistically insignificant and the null is rejected. This proves that there is, indeed, ARCH effects.

Appendix_4: CLRM violations Tests

		CAPM		Fama-French-3		Fama-French-5	
	Null-Hypo	Prob>Chi	Result	Prob>Chi	Result	Prob>chi	Result
Hetereoskedasticity							
Breusch-Pagan	Constant Variance	0.8593	No Hettest	0.3528	No hettest	0.3031	No hettest
Autocorrelation							
B Godfrey LM	No serial correlation	0.4523	No serial correlation	0.5749	No serial corerlation	0.4391	No serial correlation
D Watson	No serial correlation	1.947	No serial correlation	2.007	No serial corerlation	2.024	No serial correlation
Portmanteau White noise	Whitenoise present	0.4435	Whitenoise	0.4435	Whitenoise	0.4435	Whitenoise
Structural Break							
Wald-test	No Break	0.3687	No break	0.1673	No Break	0.6472	No Break
Corrolation							
Corrolation	70% <		No		No		Possbile
Omitted Variables							
Ramsey	No Ovs	0.610.	No Ovs	0.2721	No Ovs	0.9475	No Ovs

Appendix_5: Command Lists

Commands for Question_1

```
. tsset time
. gen e_portfolio = r_portfolio - r_rf
. gen e_mktport = r_mktport - r_rf
. scatter e_portfolio e_mktport || lfit e_portfolio e_mktport

. regress e_portfolio e_mktport
. regress e_portfolio e_mktport smb hml
. regress e_portfolio e_mktport smb hml rmw cma

. predict resid, residuals
. twoway (tsline resid)
. estat hettest, rhs
. regress e_portfolio e_mktport smb hml rmw cma, vce (robust)
. estat dwatson
. estat bgodfrey , lags(12)
. histogram resid
. histogram resid, kdensity normal
. sktest resid
. estat sbknown, break(tm(1995m12))
```

```
. correlate e_portfolio e_mktport
. estat ovtest
```

Commands for Question_2

```
. Ac yt, lags(12)
. pac yt, lags(12)
. arima yt, ar(1/2) ma(1/4)
. Predict r_arma24, r
. wntestq r_arma24, lags(12)
. estat ic

. corrgram yt, lags(4)
. generate tdate = _n
. format tdate %tm
. tsset tdate
. dfuller
. pperron
. kpss
. arima yt, ar(1)
. predict r_arma10, r
. wntestq r_arma10, lags(4)
. estat ic

. arima yt, ar(1/2)
. arima yt, ar(1/3)
. arima yt, ar(1/4)
. arima yt, ar(1) ma(1)
. arima yt, ar(1/2) ma(1)
. arima yt, ar(1) ma(1/2)
. arima yt, ar(1) ma(1/3)
. arima yt, ar(1/2) ma(1/2)
. arima yt, ar(1/2) ma(1/3)
. arima yt, ma(1)
. arima yt, ma(1/2)
. arima yt, ma(1/3)
. arima yt, ar(1/3) ma(1)
. arima yt, ar(1/3) ma(1/2)
. arima yt, ar(1/3) ma(1/3)
. arima yt, ar(1/4) ma(1/4)

. predict resid20, resid
. ac resid20
. predict resid32, resid
. ac resid32
```

Commands for Question_3

```
. gen time = _n
. tsset t
. tsline yt
. regress yt
. estat archlm, lags(7)

. arch yt, arch(1)
. predict res, residuals
. predict sigma2, variance
. generate sigma= sqrt(sigma2)
```

. gen v = res/sigma
. tsline v, xlabel(, angle(verticlal))

. arch yt, arch(1/3)
. arch yt, arch(1/2) garch(1/2)
. arch yt, arch(1) garch(1) archmlags(1)
. arch yt, earch(1) egarch(1)
. arch yt, arch(1) garch(1) tarch(1)
. arch r_yt, arch(1) garch(1) archm
. constraint 1 [ARCH]l1.arch+[ARCH]l1.garch=1
. arch yt, arch(1) garch(1) constraints(1)
. arch r_gbp, arch(1) garch(1) tarch(1)
. arch yt, arch(1) garch(1) archm
. estat ic

. test [ARCH]L1.arch+[ARCH]L2.arch+[ARCH]L3.arch+[ARCH]L4.arch+[ARCH]L5.arch … == 1
if p-value is smaller than 5%, then it ARCH coefficients is smaller than 1. Satisfying the condition

. predict sigma2, variance
. tsline sigma2

Test if the constant term in conditional variance model is positive: (Null A0=0, A0 >0).
test [ARCH]_cons= 0

If 0, we reject the null.
This is a two sided one.
We need a two sided test:
(the null is smaller than or equal than the critical region is in the right tail).
. local sign_cons = sign([ARCH]_cons)

. local sign_cons22 = sign([ARCH]_cons)   <<<<<<<<<<< GARCH1/2 ½
. display "H_0: coef<=0 p-value = "1-normal(`sign_cons'*sqrt(number)

. display "H_0: coef<=0 p-value = "1-normal(`sign_cons22'*sqrt(number)
if p-value is zero, we reject the null; therefore, constant term of conditional variance model is positive.

To show that conditional variance is stationary, as A1+D1 < 1:
. test [ARCH]L.arch + [ARCH]L.garch == 1
. test [ARCH]L1.arch + [ARCH]L2.arch + [ARCH]L1.garch + [ARCH]L2.garch == 1
the outcome is p=0
but we need the calculations for single sided p-value
we, determine the sing of the test statistic:
. local sign_stationary = sign(-1 + [ARCH]L.arch + [ARCH]L.garch)
. local sign_stationary22= sign(-1 + [ARCH]L1.arch + [ARCH]L2.arch + [ARCH]L1.garch + [ARCH]L2.garch)
since the alternative is ‘smaller than’ the critical region is in lower tail.
the p-value of a one-sided z-stat is calculated:
. display "H_0: coef>=0 p-value = " normal(` sign_stationary '*sqrt(number)
. display "H_0: coef>=0 p-value = " normal(`sign_stationary22'*sqrt(1.34))
. display"H_0:coef>=0p-value="normal(`sign_stationary22'*sqrt(38.99))
the outcome, p-value is zero, then we reject the null, thus coefficient’s sum is smaller than 1.

Arch yt, arch(1) <<< Different models insert here
Predict res, residuals
predict sigma2, variance
Generate sigma = sqrt(sigma2)
Generate v = res/sigma
Tsline v, name (e1)
Histogram res
histogram res, kdensity normal name(e2)
graph combine e2 e2
Sktest res