**Exercises (Chapters 2 – 4) – Solutions**

1. First, import the data into Eviews. Then, calculate the excess returns of the two stocks and the market portfolio by writing on the command window:

Series exmsft = msft – rf

The variable exmsft defines the excess return of Microsoft stock.[[1]](#footnote-1) Similarly, calculate the excess returns of GE and the market portfolio.

1. Group MSFT and GE creating a new object. In the new object go to View -> Descriptive Stats -> Common Sample. Table 1 presents the descriptive statistics of the two stocks excess returns. To create a joint plot of the two stocks, go to View -> Graph and click OK. Figure 1 shows a joint plot of the time-series observations of MSFT and GE excess return.

|  |  |  |
| --- | --- | --- |
| **Table 1: Descriptive statistics of GE and MSFT excess returns** | | |
|  | EXGE | EXMSFT |
| Mean | -0.030742 | -0.023547 |
| Median | -0.037149 | -0.023156 |
| Maximum | 0.143972 | 0.359511 |
| Minimum | -0.237442 | -0.396229 |
| Std. Dev. | 0.069072 | 0.108938 |
| Skewness | 0.069981 | 0.269196 |
| Kurtosis | 3.378953 | 4.610460 |
|  |  |  |
| Jarque-Bera | 0.897570 | 15.85896 |
| Probability | 0.638403 | 0.000360 |
|  |  |  |
| Sum | -4.057966 | -3.108141 |
| Sum Sq. Dev. | 0.624988 | 1.554641 |
|  |  |  |
| Observations | 132 | 132 |

GE stock has an average monthly return of -3%, ranging between -23% and 14%. Its monthly standard deviation is 6.9%. The stock empirical distribution exhibits positive skewness, indicating that the right tail is longer or fatter than the left one. It also has a kurtosis of 3.37, which is larger than the kurtosis of the normal distribution (equal to 3). Therefore, the tails of the empirical distribution are fatter to those predicted by the normal distribution.

MSFT stock has an average monthly return of -2.3%, ranging between -39% and 36%. Its monthly standard deviation is 10.9%, lower than the standard deviation of GE. The skewness of the empirical distribution, equal to 0.26, is positive, indicating that the right tail is longer or fatter than the left one. Once more, the kurtosis exceeds 3, thus the tails of the distribution are fatter to those predicted by the normal distribution.

The plots of Figure 1 indicate that the two stocks move closely in parallel. The first period examined, 1998 – 2002, is dominated by high variability, related to the burst of the DotCom bubble in 2000 – 2001. From, 2002 to 2008, the variability has significantly decreased without the two stocks providing significant returns to investors. The sample period ends in 2008 with a significant drop in the stocks’ value.



**Figure 1: Time-series plot of GE and MSFT excess returns, 1998-2008**

1. We first estimate the CAPM for MSFT. To do so, we go to Quick -> Estimate Equation. In the Equation specification field, we write:

exmsft c exmkt

where exmsft is the dependent variable of the regression, c denotes the intercept, and exmkt is the independent variable. In the Estimation settings we can choose the estimation method. The default value is LS – Least Squares, indicating that the model would be estimated using ordinary least squares (OLS). We can also choose the sample period. The default value is the entire sample period of the data. We can change that by choosing to estimate the model using a sub-period. Click OK and a new table (see Table 2) will appear showing the estimation results.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2: OLS estimates of the CAPM for MSFT stock** | | | | |
| Dependent Variable: EXMSFT | | | |  |
| Method: Least Squares | | |  |  |
| Sample: 1998M01 2008M12 | | |  |  |
| Included observations: 132 | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.013737 | 0.009061 | 1.516086 | 0.1319 |
| EXMKT | 1.259919 | 0.156861 | 8.032071 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.331668 | Mean dependent var | | -0.023547 |
| Adjusted R-squared | 0.326527 | S.D. dependent var | | 0.108938 |
| S.E. of regression | 0.089400 | Akaike info criterion | | -1.976347 |
| Sum squared resid | 1.039016 | Schwarz criterion | | -1.932669 |
| Log likelihood | 132.4389 | Hannan-Quinn criter. | | -1.958598 |
| F-statistic | 64.51416 | Durbin-Watson stat | | 2.348050 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

The results of the previous table indicate the followings: First, the beta of the CAPM for MSFT is estimated to be equal to 1.26. This coefficient is statistically significant at the 1% level. Second, the intercept is estimated to be equal to 0.013 (or 1.3%) but it is insignificant. Third, the coefficient is equal to 33%, indicating that 33% of the variation of MSFT excess returns during the sample period is explained by the model, whereas the remaining 67% remains unexplained. Finally, the S.E of regression provides an estimate of (i.e., the standard deviation of the error term), equal to 8.9%. This can also be considered as an estimate of *idiosyncratic volatility* for MSFT. Given that beta coefficient exceeds one we may conclude that MSFT is an aggressive stock.

We repeat the same exercise for GE. The results are shown in Table 3. They indicate the followings: First, the beta of the CAPM for GE is estimated to be equal to 0.86 This coefficient is statistically significant at the 1% level. Second, the intercept is estimated to be equal to -0.005 (or -0.5%) but it is insignificant. Third, the coefficient is equal to 38%, indicating that 38% of the variation of GE excess returns during the sample period is explained by the model, whereas the remaining 62% remains unexplained. Finally, the idiosyncratic volatility is estimated to be equal to 5.4%. Given that beta coefficient is lower one we may conclude that GE is a defensive stock.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 3: OLS estimates of the CAPM for GE stock** | | | | |
| Dependent Variable: EXGE | | | |  |
| Method: Least Squares | | |  |  |
| Sample: 1998M01 2008M12 | | |  |  |
| Included observations: 132 | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -0.005324 | 0.005518 | -0.964827 | 0.3364 |
| EXMKT | 0.858974 | 0.095525 | 8.992169 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.383475 | Mean dependent var | | -0.030742 |
| Adjusted R-squared | 0.378732 | S.D. dependent var | | 0.069072 |
| S.E. of regression | 0.054443 | Akaike info criterion | | -2.968300 |
| Sum squared resid | 0.385321 | Schwarz criterion | | -2.924621 |
| Log likelihood | 197.9078 | Hannan-Quinn criter. | | -2.950551 |
| F-statistic | 80.85911 | Durbin-Watson stat | | 2.255717 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. To construct confidence intervals in Eviews in the Equation object go to View -> Coefficient Diagnostics -> Confidence Intervals. In the confidence levels field, you can choose all three conventional levels (the default value) or just the 95%. In the latter case you obtain the 95% confidence intervals for all the parameters of the model. Therefore, the confidence interval for MSFT beta is equal to . Analytically, we can obtain the 95% confidence interval as follows:

We estimate with 95% confidence that the beta of MSFT falls in the above interval. It is possible that MSFT beta falls outside this interval, but we would be surprised of it since this procedure works 95% of time.

Similarly, we find that the 95% confidence interval of GE beta is equal to:

1. For MSFT the hypothesis to test is: . To perform this test in Eviews in the Equation object go to View -> Coefficient Diagnostics -> Wald test. In the new window field write the null hypothesis of your test. By default, Eviews perform only two-tailed tests, so practically you can only test for . To do so, write:

C(2)=1

The t-statistic of the two-tailed test is 1.65 and the p-value is equal to 9.9%. You can now easily derive the p-value of the one-sided test. It’s half the p-value of the two-tailed test, i.e., 4.95%. Therefore, the null hypothesis cannot be rejected at the 1% level. Therefore, we cannot conclude that MSFT is an aggressive stock.

Analytically, we can calculate the t-statistic as follows:

The critical value is . Therefore, , so we cannot reject the null hypothesis.

For GE the hypothesis to test is: . The t-statistic of the test is:

The critical value is . Therefore, , and again the null hypothesis cannot be rejected at the 1% significance level.

1. We have already tested for the significance of the intercept term in question (b). For both stocks the intercept term is insignificant, therefore we cannot reject that . This result indicates that the sample data are consistent with the prediction of economic theory that the intercept term is equal to 0.
   1. First, import the data into Eviews. Group , , and creating a new object. In the new object go to View -> Descriptive Stats -> Common Sample. Table 1 presents the descriptive statistics of the three variables.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1: Descriptive statistics** | | | |
|  | STARTS | SOLD | FIXED\_RATE |
| Mean | 1506.696 | 67.82609 | 7.650815 |
| Median | 1525.500 | 66.00000 | 7.675000 |
| Maximum | 2228.000 | 128.0000 | 10.22000 |
| Minimum | 798.0000 | 29.00000 | 5.500000 |
| Std. Dev. | 272.9083 | 19.08515 | 1.146265 |
| Skewness | 0.052947 | 0.569399 | 0.306674 |
| Kurtosis | 2.849400 | 3.215838 | 2.640599 |
|  |  |  |  |
| Jarque-Bera | 0.259854 | 10.29975 | 3.874468 |
| Probability | 0.878160 | 0.005800 | 0.144102 |
|  |  |  |  |
| Sum | 277232.0 | 12480.00 | 1407.750 |
| Sum Sq. Dev. | 13629645 | 66656.43 | 240.4480 |
|  |  |  |  |
| Observations | 184 | 184 | 184 |

During 1990-2005, the average number of new housings start per month is 1,507,000 ranging between 798,000 and 2,228,000. The standard deviation is 272,908. On the same period, the average number of houses sold per month is 67,820 ranging between 29,000 and 128,000. The standard deviation is 19,085. The average 30-year fixed mortgage rate is 7.65% with a standard deviation of 1.14%. Its value ranges between 5.5% and 10.22%.

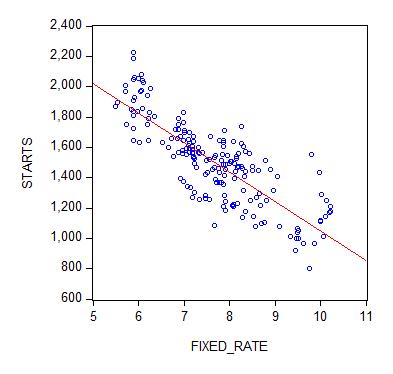
* 1. To estimate the regression model, go to Quick -> Estimate Equation. In the Equation specification field, write:

starts c fixed\_rate

Click OK. Table 2 shows the estimation results. These results indicate the followings: First, the intercept term is estimated to be equal to 2992.739. Thus, if the fixed rate goes to zero, the new housings start would be 2,992,739. This coefficient is statistically significant at the 1% level. Second, the coefficient of the FIXED\_RATE variable is estimated to be equal to -194.2334. If the mortgage fixed rate increases (decreases) by 1 percentage point, the number of new housings start will decrease (increase) by 194,233. This coefficient is also statistically significant at the 1% level, indicating a statistically significant negative relation between mortgage interest rates and new housings start. Finally, the coefficient of the regression is 67%, indicating that 67% of the variation in new housings start is explained by the model, while the remaining 33% remains unexplained.

To create a scatter plot of the two variables along with the regression line first group the two variables. Then go to View -> Graph. In the field Specific choose Scatter, and in the field Fit lines choose Regression line. Figure 1 presents the scatter plot of the two variables along with the estimated regression line. The plot supports our previous findings. There is a strong negative relation between and .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2: OLS estimates regressing STARTS on FIXED\_RATE** | | | | |
| Dependent Variable: STARTS | | | |  |
| Method: Least Squares | | |  |  |
| Sample: 1990M01 2005M04 | | |  |  |
| Included observations: 184 | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 2992.739 | 78.95146 | 37.90607 | 0.0000 |
| FIXED\_RATE | -194.2334 | 10.20606 | -19.03119 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.665556 | Mean dependent var | | 1506.696 |
| Adjusted R-squared | 0.663718 | S.D. dependent var | | 272.9083 |
| S.E. of regression | 158.2590 | Akaike info criterion | | 12.97715 |
| Sum squared resid | 4558359. | Schwarz criterion | | 13.01210 |
| Log likelihood | -1191.898 | Hannan-Quinn criter. | | 12.99132 |
| F-statistic | 362.1861 | Durbin-Watson stat | | 0.392418 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |



**Figure 1: Scatter plot of FIXED RATE and STARTS**

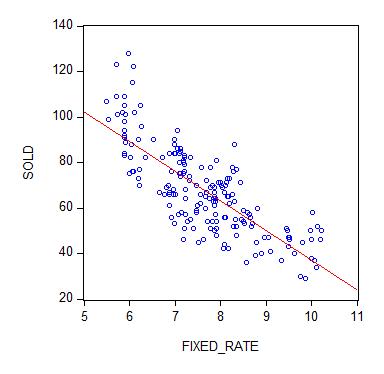
* 1. To estimate the regression model, go to Quick -> Estimate Equation. In the Equation specification field, write:

sold c fixed\_rate

Click OK. Table 3 reports the results indicating the followings: First, the intercept term is estimated to be equal to 167.5479. Thus, if the fixed rate goes to zero, the houses sold would be 167,547. This coefficient is statistically significant at the 1% level. Second, the coefficient of the FIXED\_RATE variable is estimated to be equal to -13.0341. If the mortgage fixed rate increases (decreases) by 1 percentage point, the number of houses sold will decrease (increase) by 13,034. This coefficient is also statistically significant at the 1% level, indicating a statistically significant negative relation between mortgage interest rates and houses sold. Finally, the coefficient of the regression is 61%, indicating that 61% of the variation in houses sold is explained by the model, while the remaining 39% remains unexplained.

To create a scatter plot of the two variables along with the regression line first group the two variables. Then go to View -> Graph. In the field Specific choose Scatter, and in the field Fit lines choose Regression line. Figure 2 presents the scatter plot of the two variables along with the estimated regression line. The plot supports our previous findings. There is a strong negative relation between and .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 3: OLS estimates regressing SOLD on FIXED RATE** | | | | |
| Dependent Variable: SOLD | | | |  |
| Method: Least Squares | | |  |  |
| Sample: 1990M01 2005M04 | | |  |  |
| Included observations: 184 | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 167.5479 | 5.940521 | 28.20425 | 0.0000 |
| FIXED\_RATE | -13.03415 | 0.767931 | -16.97306 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.612836 | Mean dependent var | | 67.82609 |
| Adjusted R-squared | 0.610709 | S.D. dependent var | | 19.08515 |
| S.E. of regression | 11.90784 | Akaike info criterion | | 7.803081 |
| Sum squared resid | 25806.98 | Schwarz criterion | | 7.838026 |
| Log likelihood | -715.8834 | Hannan-Quinn criter. | | 7.817244 |
| F-statistic | 288.0849 | Durbin-Watson stat | | 0.371369 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |



**Figure 2: Scatter plot of FIXED RATE and SOLD**

* 1. We formulate the null and alternative hypothesis as follows: . The t-statistic of this test is:

The critical values of the test are and . We observe that , so the null hypothesis is rejected. Therefore, an increase of mortgage rates by 1 percentage point would not decrease house starts by 150,000.

* 1. The 95% confidence interval is given as:

In 95% of repeated samples from the same population the true value of lies between . In our sample this is equal to the interval . We are 95% confident that if mortgage interest rates increase by 1 percentage point the number of houses sold would decrease by 11,520 to 14,540.

* 1. The estimated regression model predicts that the number of housings start on May 2005 is:

Similarly, the number of housings start predicted in June 2005 is:

* 1. To calculate the prediction intervals, we need first to calculate the standard error. The forecast error variance is given as:

For May 2015 the forecast error variance is:

The standard error is .

For June 2005 the forecast error variance is:

The standard error is .

The 95% prediction interval for May 2005 is:

The 95% prediction interval for June 2005 is:

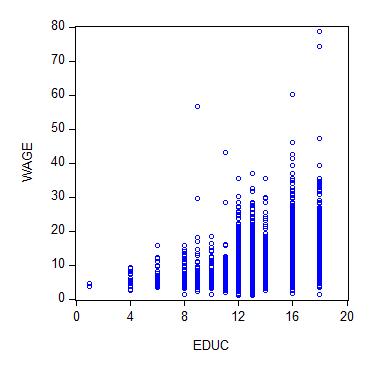
In both cases the interval contains the true value of new housings start.

* 1. First, import the data into Eviews. Then, group the two variables creating a new object. In the new object go to View -> Descriptive Stats -> Common Sample. Table 1 presents the descriptive statistics of EDUC and WAGE.

|  |  |  |
| --- | --- | --- |
| **Table 1: Descriptive statistics** | | |
|  | EDUC | WAGE |
| Mean | 13.30235 | 10.18720 |
| Median | 13.00000 | 8.530000 |
| Maximum | 18.00000 | 78.71000 |
| Minimum | 1.000000 | 1.050000 |
| Std. Dev. | 2.356101 | 6.213761 |
| Skewness | -0.087973 | 2.109494 |
| Kurtosis | 4.421100 | 12.46459 |
|  |  |  |
| Jarque-Bera | 404.3724 | 21175.88 |
| Probability | 0.000000 | 0.000000 |
|  |  |  |
| Sum | 62960.00 | 48216.01 |
| Sum Sq. Dev. | 26268.34 | 182706.4 |
|  |  |  |
| Observations | 4733 | 4733 |

The average and median years of education in the sample is 13.30 and 13, respectively, ranging between 1 and 18 years. The standard deviation is 2.35 years. The average hourly wage is $10.18, ranging between $1.05 and $78.71. The standard deviation is $6.21.

To create a scatter plot of the two variables, go to View -> Graph. In the field Specific choose Scatter. Figure 1 presents the scatter plot of EDUC and WAGE. The plot shows a positive relation between years of education and wage rates. It also indicates an increased variability of wage rates for higher levels of years of education.



**Figure 1: Scatter plot of EDUC and WAGE**

* 1. To estimate the regression model, go to Quick -> Estimate Equation. In the Equation specification field, write:

wage c educ

Click OK. Table 2 shows the estimation results.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2: Linear regression model of WAGE on EDUC** | | | | |
| Dependent Variable: WAGE | | | |  |
| Method: Least Squares | | |  |  |
| Sample: 1 4733 | | |  |  |
| Included observations: 4733 | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -5.202605 | 0.465486 | -11.17672 | 0.0000 |
| EDUC | 1.156924 | 0.034457 | 33.57628 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.192437 | Mean dependent var | | 10.18720 |
| Adjusted R-squared | 0.192266 | S.D. dependent var | | 6.213761 |
| S.E. of regression | 5.584556 | Akaike info criterion | | 6.278309 |
| Sum squared resid | 147546.9 | Schwarz criterion | | 6.281040 |
| Log likelihood | -14855.62 | Hannan-Quinn criter. | | 6.279269 |
| F-statistic | 1127.367 | Durbin-Watson stat | | 0.358700 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

The results of the table indicate the followings: First, the intercept term is estimated to be equal to -5.20, indicating that workers with 0 years of education have an average wage rate of -$5.20. Even if, this coefficient is statistically significant it is unrealistic; negative wages do not exist. Second, the coefficient of EDUC is positive, equal to 1.15, and statistically significant at the 1% significance level. Therefore, a 1-year increase in education would increase hourly wage rate by $1.15. Finally, the coefficient is 19%, indicating that 19% of the variation in wage rates in the sample can be explained by the model. The rest, 81%, remains unexplained.

To estimate the log-linear regression model, go to Quick -> Estimate Equation. In the Equation specification field, write:

Log(wage) c educ

Click OK. Table 3 reports the results.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 3: Linear regression model of WAGE on EDUC** | | | | |
| Dependent Variable: LOG(WAGE) | | | |  |
| Method: Least Squares | | |  |  |
| Sample: 1 4733 | | |  |  |
| Included observations: 4733 | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.770472 | 0.040932 | 18.82324 | 0.0000 |
| EDUC | 0.104949 | 0.003030 | 34.63764 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.202295 | Mean dependent var | | 2.166535 |
| Adjusted R-squared | 0.202127 | S.D. dependent var | | 0.549766 |
| S.E. of regression | 0.491072 | Akaike info criterion | | 1.415969 |
| Sum squared resid | 1140.887 | Schwarz criterion | | 1.418699 |
| Log likelihood | -3348.890 | Hannan-Quinn criter. | | 1.416928 |
| F-statistic | 1199.766 | Durbin-Watson stat | | 0.379494 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

The results of the table indicate the followings: First, the intercept term is estimated to be equal to 0.77, indicating that workers with 0 years of education have an average log-wage rate of 0.77. Discarding convexity adjustment this implies an average wage rate of for workers with 0 years of education. This coefficient is statistically significant at the 1% level. Second, the coefficient of EDUC is positive, equal to 0.1, and statistically significant at the 1% significance level. Therefore, a 1-year increase in education would increase hourly wage rate by . Finally, the coefficient is 20%, indicating that 20% of the variation in log-wage rates in the sample can be explained by the model. The rest, 80%, remains unexplained.

* 1. For the log-linear model this is given by the estimate of , equal to 0.1. Therefore, a 1-year increase in education would increase hourly wage rate by . For the linear regression model this is given as:

Therefore, a 1-year increase in education would increase hourly wage rate by 11%.

* 1. To compute the generalized we need first to calculate the correlation coefficient between the fitted values of WAGE and the sample values of WAGE. To compute the fitted values of WAGE under the log-linear model use the Equation object. Go to Forecast and choose WAGE. Click OK. A new series (the default name is WAGEF) is created in the workfile. This series is the fitted values of WAGE. As a second step, open WAGE and WAGEF as a group. Go to View -> Covariance Analysis and choose correlation. The correlation matrix indicates that . Therefore, the generalized , or 21%. Compared, with the coefficient of 19% of the linear model, the log-linear model fits better to the data, though the improvement is small.
  2. Open EDUC and RESID as a group. Go to View -> Graph and choose Scatter. Figure 2 shows the scatter plot of EDUC and the residuals of the log-linear regression model. We observe that the absolute values of residuals increase in magnitude as EDUC increases suggesting the presence of heteroskedasticity. Also, we observe positive residuals in the early range of EDUC, suggesting that EDUC has an impact only after a minimum number of years of education.



**Figure 2: Scatter plot of EDUC and RESID**

* 1. We formulate the null and alternative hypothesis as follows: . The t-statistic of the test is:

The critical values of the test are -1.96 and 1.96. Therefore, the null hypothesis is rejected at the 5% significance level.

* 1. Using the linear regression model, the wage rate of a worker with 16 years of education is predicted as:

Using the log-linear regression model, the natural predictor is given as:

The corrected predictor is given as:

The average wage rate of all workers in the sample with 16 years of education is 13.5.[[2]](#footnote-2) We observe that both models can accurately predict the sample average.

1. Alternatively, we can calculate excess returns in Excel before importing the data into Eviews. [↑](#footnote-ref-1)
2. To calculate the average wage of workers with EDUC = 16, reduce the sample to them by writing on the Command Window smpl if educ = 16. The descriptive statistics tool shows that the average wage rate of workers with EDUC = 16 is 13.5. To restore the full sample, write on the Command Window smpl @all. [↑](#footnote-ref-2)