1. **Ursula Beijer is the portfolio manager of Scheldt Equities Fund. The fund holds stocks from emerging markets, and Beijer is concerned about the risk of government debt default. She asks Roland Drenth, a quantitative analyst at the firm, to build a multiple regression model based on macroeconomic factors to predict whether or not a country will default. Drenth explains that the objective of the model is to predict a discrete variable that takes the value of:**

    - **1 when the model predicts a debt default or**
    - **0 when a default is not expected.**

    Beijer also wants to investigate how the returns of the Scheldt fund are influenced by distinct market factors. She asks Drenth to build a regression model of the fund's excess returns against the three factors in the Fama-French model. Drenth presents the following multiple regression equation:

    $$SEF_i = b_0 + b_1 MKTRF_i + b_2 SMB_i + b_3 HML_i + \varepsilon_i, \text{ where:}$$

    - **SEF = Excess return of the Scheldt Equities Fund**
    - **MKTRF = Market excess return over the risk-free rate**
    - **SMB = Small minus big (size effect)**
    - **HML = High minus low (value premium)**

    Beijer asks Drenth how to interpret the regression variables and estimated coefficients. Drenth makes the following statements:

    Statement 1: **The intercept b_0 is the fund's expected excess return if the error term ε_i is zero.**

    Statement 2: **The slope b_1 describes how SEF is affected by a unit change in MKTRF while holding SMB and HML constant.**

    Statement 3: **MKTRF, SMB, and HML constitute the stochastic part of the model.**

    Drenth then runs the regression with 60 monthly return observations, stated in whole percentages (ie, 1 = 1%), and arrives at the following equation with the estimated coefficients:

    $$SEF = 0.06 + 0.51 MKTRF - 0.72 SMB + 0.30 HML$$

    Beijer asks for a numeric example, and Drenth estimates Scheldt's excess return in a hypothetical month, given the following assumptions:

    - **MKTRF = 1.00%,**
    - **SMB = 1.00%, and**
    - **HML = 0.50%.**

    Beijer wants to use the regression model to predict the fund's performance. However, Drenth advises Beijer that several additional steps must be taken before deeming the

model acceptable for making predictions and performing further analysis.

Question 1 of 4

To estimate whether or not a country will default, Drenth should *most appropriately* use a:

a. logistic regression model.
b. restricted regression model.
c. regression model with a dummy independent variable.

---

2. Question 2 of 4

Which of the following statements about variables and coefficients is *correct*?

a. Statement 1
b. Statement 2
c. Statement 3

---

3. Question 3 of 4

Based on the estimated regression equation and on Drenth's assumptions, Scheldt's estimated excess return is *closest* to:

a. −0.06%
b. 0.00%
c. 0.15%

---

4. Question 4 of 4

After estimating the regression model, Drenth's *most appropriate* next step is to:

a. analyze the residuals.
b. measure the goodness of fit.
c. test the model's performance with out-of-sample data.

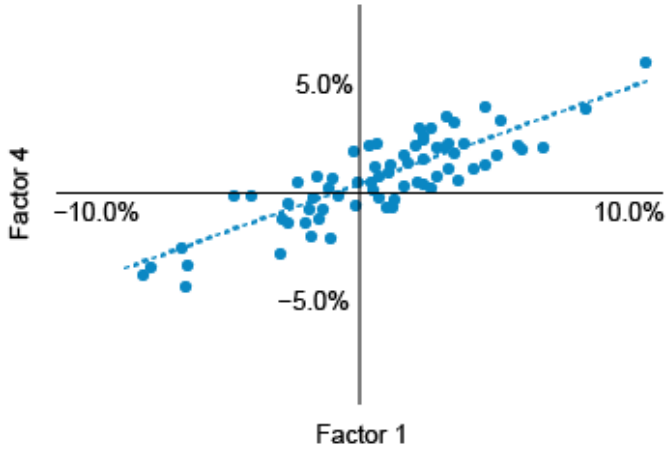
---

5. Henrik Lippert is a junior quantitative analyst at Regnitz. Lippert is running a multiple regression to predict revenue growth for an energy stock he covers and presents his findings to Tanja Vogl, his supervisor at Regnitz. Vogl analyzes the residual plot and notices a pattern: the variance of residuals changes across observations and is correlated with one of the independent variables. She tells Lippert that he needs to adjust the model since it appears to violate a multiple linear regression assumption.

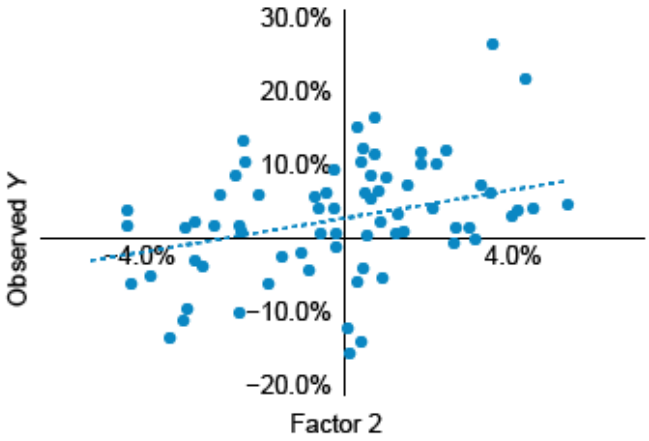
Lippert uses 72 observations to recalculate the regression of the dependent variable against four factors, and he plots the three graphs presented in Exhibit 1:

Exhibit 1 Selected Scatterplots for Multiple Regression Analysis

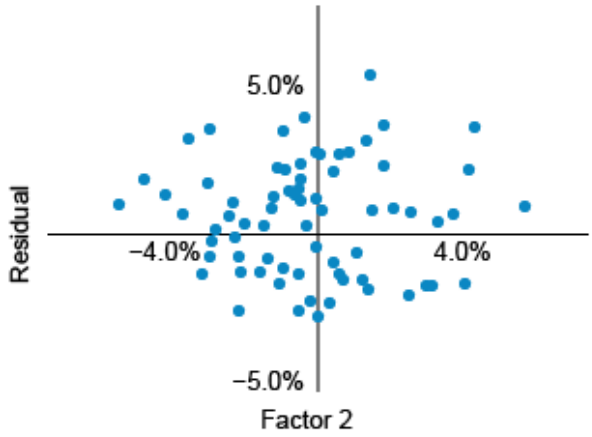
Graph 1: Factor 4 against Factor 1



Graph 2: Observed *Y* against Factor 2



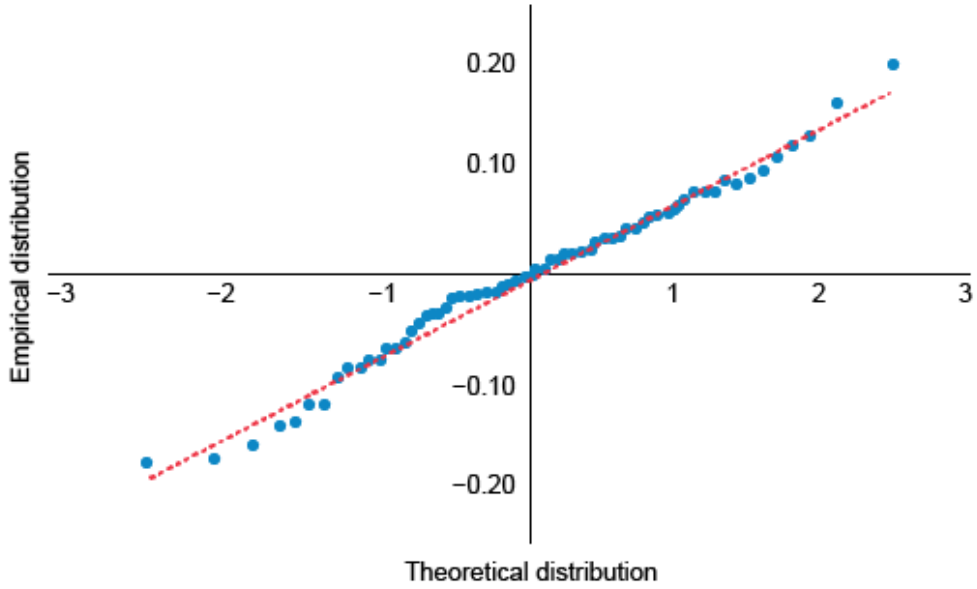
Graph 3: Regression residuals against Factor 3



Vogl analyzes the graphs and says that Graph 1 suggests a potential violation of another multiple linear regression assumption. She adds that scatterplots can also be used to identify the violation of other assumptions, such as linearity.

Finally, Vogl asks Lippert to create a Q–Q plot of regression residuals; he presents the graph in Exhibit 2:

Exhibit 2 Q–Q Plot of Regression Residuals



Question 1 of 4

Which of the following violations is *most likely* indicated by the changes in the residuals' variances across observations?

a. Nonlinearity
b. Autocorrelation
c. Heteroskedasticity

---

6. Question 2 of 4

Graph 1 in Exhibit 1 indicates a possible violation of the assumption of independence of:

a. errors.
b. observations.
c. independent variables.

---

7. Question 3 of 4

Based on Exhibit 1, which of the following graphs is *most appropriately* used to identify a violation of the linearity assumption?

a. Graph 1
b. Graph 2
c. Graph 3

---

8. Question 4 of 4

Based on Exhibit 2, which of the following multiple linear regression assumptions is *most likely* violated?

a. Linearity
b. Normality
c. Homoskedasticity
