
PO12Q - Introduction to Quantitative Political Analysis II:

Indicative Answers to Self-Reflection Questions

Dr Florian Reiche
F.Reiche@warwick.ac.uk



1. Compare bivariate regression and multiple regression.

- multiple regression allows multiple independent variables, bivariate only one
- this means that whilst the conditional expected value can be written as $E(y|x)$ in a bivariate model, it needs to be extended to $E(y|x_1, x_2, \dots, x_k)$ in multiple regression
- the slope coefficients in multiple regression are partial slope coefficients. Their interpretation is “ceteris paribus”, meaning all other things (i.e. the values of all other variables included in the model) being equal.
- To measure model fit, we can use R^2 in bivariate regression, but we need to use Adjusted R^2 (\bar{R}^2) in multiple regression.

2. Give an example of the relationship in which you could apply multiple linear regression.

- Any in which you have more than one explanatory factor, so pretty much any situation in social sciences. For example, using modernisation theory, we could try to explain the level of democracy through per capita GDP, infant mortality rates, and literacy rates.

3. How do you interpret partial slope coefficients?

- As indicated under 1., partial slope coefficients have been purified from the influence of the other variables included in the model. So we need to indicate this by adding to the interpretation “holding the values of the other variables constant”. You should make this as specific as possible, and replace “the other variables” by the actual variables in the model. Look at Model 3 on lecture slide 13, for example. The correct interpretation of the “Ethnic Minorities” coefficient would be: “Holding household income constant, for every additional percentage of ethnic minorities in a constituency, the average GCSE score of that constituency would decrease by 0.073, on average”.

4. How do you interpret adjusted R-Squared in regression analysis?

- You interpret this pretty much the same way as R^2 in bivariate regression, in the sense that \bar{R}^2 tells us how much variation of the dependent variable is explained through the independent variables.

5. Prepare to interpret the following regression output substantively (what do the coefficients mean, how good is the model fit, etc.) income represents the median household income in a constituency (in £), and turnout represents the turnout at the 2012 mayoral election (in %).

- **Intercept:** This is statistically significant at a 95% confidence level, so we can interpret this. If the median household income in a London ward was equal to zero, on average the voter turnout in the 2012 mayoral elections would be 19.55%
 - **Slope:** This is also statistically significant at a 95% confidence level, so we can interpret this. For every additional £ in median household income, the turnout in the 2012 mayoral elections increases by 0.000372 percentage points, on average.
 - We can obtain the t-values of both coefficients by dividing the coefficient by its respective standard error.
 - **R^2 :** Median household income is able to explain 26.73% of the variation in the turnout at the 2012 mayoral election.
 - **Degrees of Freedom:** 623 degrees of freedom tells us that there were 625 observations which – assuming there is no missing data – would be equal to the number of London wards.
-