

Homework

Using the first 100 observations of the caschool data, explore the influence and leverage properties with respect to least squares. Consider the model

$$testscr = X\beta + u$$

where X contains a constant, str , el_pct , and $avginc$.

1. [Gauss] Estimate the model using least squares. Obtain the residuals, compute h_i , and $\hat{\alpha}_i$ for each observation and list them together, along with observation numbers, in a table. Which 3 observations have highest leverage? Which 3 have the highest influence?
2. [Gauss] Using equation 2.63, how much will the coefficient estimates change if each of the three most influential points are omitted from the data? (Omit one at a time with replacement).
3. [Stata] Verify (2) using Stata.
4. [algebra] Exercise 2.29 in ETM.