An Example

Statistics and Econometrics

Question

We use the data in meap93.RData for this exercise. This data was collected from the Michigan Department of Education web site in 1993. Suppose we want to explore the relationship between the math pass rate ($math10$) and spending per student ($expend$). $math10$ denote the percentage of tenth graders at a high school receiving a passing score on a standardized mathematics exam. $expend$ indicate the average spending per student at the high school, including instruction employee salaries and benefits, pupil support, instruction staff support, school administration and etc.

1. Do you think each additional dollar spent has the same effect on the pass rate, or does a diminishing effect seem more appropriate? Explain.
2. In the population model
$$math10 = \beta_0 + \beta_1 \log(expend) + u,$$
argue that $\beta_1/10$ is the percentage point change in $math10$ given a 10% increase in $expend$.
3. Use the data in meap93.RData to estimate the model from part 2. Report the results.
4. How big is the estimated spending effect? Namely, if spending increases by 10%, what is the estimated percentage point increase in $math10$?
5. One might worry that regression analysis can produce fitted values for $math10$ that are greater than 100. Why is this not much of a worry in this data set?

Solutions

1. It seems plausible that another dollar of spending has a larger effect for low-spending schools than for high-spending schools. At low-spending schools, more money can go toward purchasing more books, computers, and for hiring better qualified teachers. At high levels of spending, we would expect little, if any, effect because the high-spending schools already have high-quality teachers, nice facilities, plenty of books, and so on.
2. If we take changes as usual, we obtain
$$\Delta math10 = \beta_1 \Delta \log(expend) \approx (\beta_1/100)(\%\Delta expend).$$
So, if $\%\Delta expend = 10$, $\Delta math10 = \beta_1/10$.
3.

```r
load("meap93.RData")
model.1 <- lm(math10 ~ log(expend), data = data)
stargazer(model.1, header = FALSE, type = 'latex', title = "Question 1.3")
```

4. If $expend$ increases by 10 percent, $\widehat{math10}$ increases by about 1.1 percentage points. This is not a huge effect, but it is not trivial for low-spending schools, where a 10 percent increase in spending might be a fairly small dollar amount.
5. In this data set, the largest value of $math10$ is 66.7, which is not especially close to 100. In fact, the largest fitted values is only about 30.2.

Table 1: Question 1.3

	Dependent variable:
	math10
log(expend)	11.164^{***}
	(3.169)
Constant	-69.341^{***}
	(26.530)
Observations	408
R^2	0.030
Adjusted R^2	0.027
Residual Std. Error	10.350 (df = 406)
F Statistic	12.411^{***} (df = 1; 406)
Note:	*p<0.1; **p<0.05; ***p<0.01