Panel Data Estimators: Key

Suppose that our dependent variable is the median earnings of college graduates five years after graduation. We have data for all the NESCAC schools for all graduating classes from 1990 to 2007. Our data also includes U.S. GDP, and each school's endowment.

1. We will if there is unobserved heterogeneity by year that is correlated with some of our independent variables. We may expect, for example that graduates during the recession of the early 2000s may have reduced earnings. This is surely correlated with U.S. GDP. As long as it is not entirely captured by U.S. GDP (in which case pooled OLS may be plausible), time fixed effects make sense.

2. We will if there is unobserved heterogeneity by school. bates, for example, is obviously the best school in the NESCAC but this is not captured by our low endowment. So we need cross sectional dummies to capture this heterogeneity.

3. The Bates dummy will be super-duper-mega-large.

4. if all heterogeneity is captured by the independent variables. So labor market conditions are entirely captured by GDP, eliminating the need for time fixed effects, And if endowment captures all heterogeneity across schools.

5. The Hausman test does this.

6. We may run an f-test jointly testing the significance of all the fixed effects.

7. True. This is why fixed effects is also known as the Least Squared Dummy Variable Model.

8. True. In this case, OLS is consistent but inefficient.

9. We can run group fixed effects where these schools have the same fixed effects. This will be efficient.