More on Differences-in-Differences

14.32 Recitation 4/13/07

In lecture, we went over Card and Kreuger's estimating equation:

$$y_{it} = \alpha + \gamma NJ_i + \delta D_2 + \beta NJ_i * D_2 \qquad \#$$

Sometimes it's useful to use a 2x2 matrix for thinking about DD. For each of Card and Kreuger's four state-time categories, the conditional mean of y_{it} is:

	$T = 1$	$T = 2$	Diff across time
$NJ = 0$	α	$\alpha + \delta$	δ
$NJ = 1$	$\alpha + \gamma$	$\alpha + \gamma + \delta + \beta$	$\delta + \beta$
Diff. across states	γ	$\gamma + \beta$	β

Causality

The critical assumption in CK is that in the absence of the intervention, both states would have the same *time trend*. When using time series such as these, we can test the assumption using pre-trends. That is, check whether NJ and PA had the same time trend before the minimum wage was raised in NJ. Card and Kreuger don'bt report this in the main paper, but they do have some additional evidence on trends in the *Myth and Measurement* book.

Another DD application: Children's Test Scores:

Banerjee, **Cole**, **Duflo and Linden** (Forthcoming, Quarterly Journal of Economics) report the results of several randomized experiments in India. Randomization is considerably more "bulletproof" than cross-state comparisons. However, it's expensive collecting your own data and providing your own interventions. Development economists perform randomized trials a lot because

- Experiments are generally cheaper in developing countries
- It's harder to get good data prepackaged from developing countries

Computer-assisted learning/balsakhi experiment: schools were randomly assigned an extra teacher ("balsakhi") or were taught using computers (computer-assisted learning, or CAL), both, or nothing.

For the CAL experiment, schools were randomly assigned to either "treatment" (computer) or "control" (no computer) group. Computer classes were taught two hours a week to fourth graders and focused on math skills. Students were tested before the program began ("pretest") and after ("posttest").

Data were stacked for the analysis. That is, each child had two observations, one for the pretest and one for the posttest. The estimating equation for the effect of the program is

$$test_{it} = \alpha + \gamma treat_{it} + \delta post_{it} + \beta treat_{it} * post_{it}$$

where i, t represent child i, test t,
$treat = 1$ if the child was in the treatment group, 0 otherwise
$post = 1$ if the test was a posttest, 0 otherwise
$treat_{it} * post_{it}$ is the interaction of the two

We can create a similar 2x2 matrix as before:

	$post = 0$	$post = 1$	Diff across tests
$treat = 0$	α	$\alpha + \delta$	δ
$treat = 1$	$\alpha + \gamma$	$\alpha + \gamma + \delta + \beta$	$\delta + \beta$
Diff. across treatments	γ	$\gamma + \beta$	β

Questions:
- What is the critical assumption?
- How confident are we that the assumption holds?
- Was it important to have a pretest?