Impact Evaluation Methods (2)

2. Difference-in-differences

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy experiment." *American Economic Review* 91(4): 795–813.

	Years of education		
	Level of program in region of birth		Difference
	Low (C)	High (T)	(T - C)
Before (12-17 in 1974)	9.40	8.02	-1.38
After (2-6 in 1974)	9.76	8.49	-1.27
Before-After changes	0.36	0.47	0.11

<u>Data needed</u>: Observations before and after the implementation of the program, for both the treatment group and the comparison group.

Key assumption for the validity of the method: the difference between before and after in the comparison group is a good counterfactual for the treatment group.

a) Compute the difference before-after for the comparison group:

$\bar{y}_{C1} - \bar{y}_{C0}$.

Represents the change in outcome due to natural trend and all other events.

b) Compute the difference before-after for the treatment group:

$\bar{y}_{T1} - \bar{y}_{T0}$.

Represents the change in outcome due to natural trend and all other events, and the program

c) The impact of the program:

$$\boxed{\text{Impact } = \left(\bar{y}_{T1} - \bar{y}_{T0}\right) - \left(\bar{y}_{C1} - \bar{y}_{C0}\right)}$$

In a regression framework:

$y_i = \beta_0 + \beta_1 After_i + \beta_2 T_i + \beta_3 T_i After_i + u_i$

$\hat{y}_i = 9.40 + 0.36 After_i - 1.39 T_i + 0.12 T_i After_i$

(0.04) (0.04) (0.07) (0.089)

Using: the number of school constructed per 1000 children as a measure of T : $\hat{\beta}_3 = 0.124$ (0.025)

Adding control variables $\hat{\beta}_3 = 0.188$ (0.0289)

Tests in support of the validity of the method

Verify that before the program, the Control and Treatment groups had the same trend ("parallel trends")
Diff-in-Diffs estimation for the 2 periods prior to the program:

	Years of education		
	Level of program in region of birth		Difference
	Low (C)	High (T)	(T - C)
6 years before (18-24 in 1974)	9.12	7.70	-1.42
Before (12-17 in 1974)	9.40	8.02	-1.38
Pre-program changes	0.28	0.32	0.04

$\hat{y}_i = 9.12 + 0.28 After_i - 1.42 T_i + 0.04 T_i After_i$

(0.04) (0.06) (0.07) (0.10)

Using: the number of school constructed per 1000 children as a measure of T : $\hat{\beta}_3 = 0.009$ (0.026)

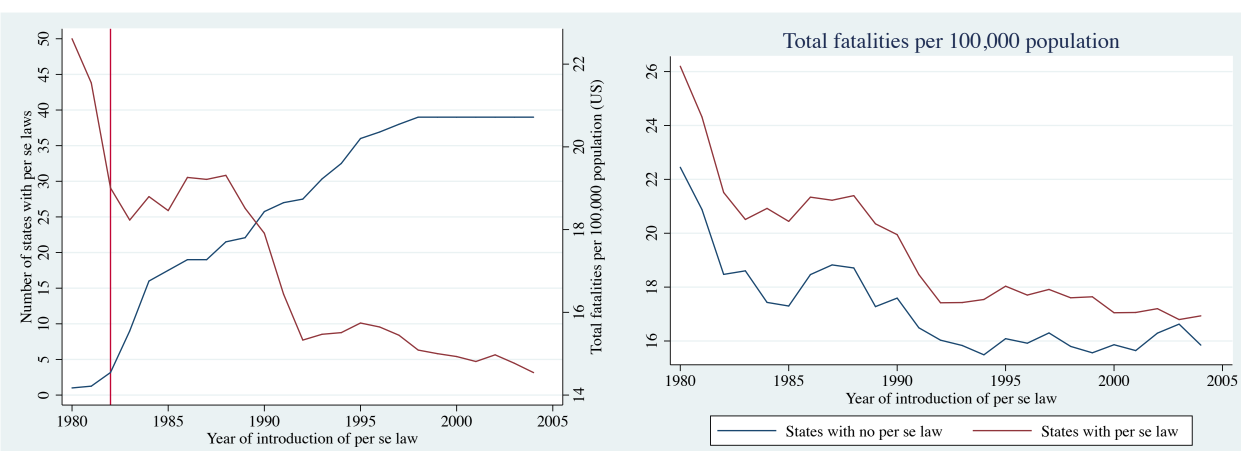
Adding control variables $\hat{\beta}_3 = 0.0075$ (0.0297)

3. Extension of Diff-in-diffs: Rollout of policies and panel data

Do change in traffic laws affect traffic fatalities?
Freeman, D.G. (2007) "Drunk Driving Legislation and Traffic Fatalities: New evidence on the BAC 08 Laws" *Contempory Economic Policy* 25, 293-308
State level analysis, using 25 years of data from 1980 to 2004, with changes in various state drunk driving, seat belt, and speed limit laws. Data set in Wooldridge "driving.dta"

```
perse           administrative license revocation (per se law)
totfatrte       total fatalities per 100,000 population
year            1980 to 2004
state           48 continental states, alphabetical

. sum perse totfatrte
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       perse |      1200    .5470833    .4928654          0          1
   totfatrte |      1200    18.91856    6.367407        6.2      53.32
```



Note: States with higher level of fatalities introduced these laws. Obvious time trend, even in states with no introduction of laws. Can we attribute their faster decline in fatalities to the law?

Regression: $y_{it} = \beta T_{it} + a_i + \delta_t + u_{it}$

```
. xtreg  totfatrte  perse i.year, fe i( state)

Fixed-effects (within) regression               Number of obs      =      1200
Group variable: state                           Number of groups   =        48

R-sq:  within  = 0.5354                         Obs per group: min =        25
       between = 0.0273                                        avg =      25.0
       overall = 0.1017                                        max =        25

                                                F(25,1127)         =     51.94
corr(u_i, Xb)  = -0.0659                        Prob > F           =    0.0000

------------------------------------------------------------------------------
   totfatrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       perse |  -1.848261   .2423821    -7.63   0.000    -2.323831    -1.37269
             |
        year |
       1981  |  -1.814749   .4565585    -3.97   0.000    -2.710549   -.9189488
       1982  |  -4.468642   .4566879    -9.78   0.000    -5.364697   -3.572588
       1983  |  -5.033624   .4583405   -10.98   0.000     -5.93292   -4.134327
       1984  |  -4.649502   .4627972   -10.05   0.000    -5.557543   -3.741461
       1985  |  -5.007786   .4640971   -10.79   0.000    -5.918377   -4.097194
..........
       2002  |  -7.001794   .4952416   -14.14   0.000    -7.973493   -6.030095
       2003  |  -7.267836   .4952416   -14.68   0.000    -8.239535   -6.296137
       2004  |  -7.302419   .4952416   -14.75   0.000    -8.274118    -6.33072
             |
       _cons |   25.53309   .3228739    79.08   0.000     24.89959    26.16659
-------------+----------------------------------------------------------------
     sigma_u |  5.7016328
     sigma_e |  2.2366622
         rho |  .86663588   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(47, 1127) =   155.62            Prob > F = 0.0000
```

Key assumption for the validity of the method:
The annual change in the comparison group is a good counterfactual for the annual change in the treatment group

Tests in support of the validity of the method:
(i). The entry into the treatment is not correlated with a differential trend in the performance of the unit in the pre-treatment period.

```
Define the change in fatality rate:  totfatrte(t)=totfatrte(t) - totfatrte(t-1)
Define the year of introduction of the law: perseyear
Regress the change in fatality rate on the year in which the law was passed


.  reg  dtotfatrte  perseyear  d82 if year<1983

      Source |       SS       df       MS              Number of obs =      78
-------------+------------------------------           F(  2,    75) =    0.82
       Model |  16.7209756     2  8.36048778           Prob > F      =  0.4437
    Residual |   763.33703    75  10.1778271           R-squared     =  0.0214
-------------+------------------------------           Adj R-squared = -0.0047
       Total |  780.058006    77  10.1306234           Root MSE      =  3.1903

------------------------------------------------------------------------------
  dtotfatrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   perseyear |   .0118334   .0712763     0.17   0.869    -.1301563    .1538232
         d82 |  -.9182045    .722454    -1.27   0.208    -2.357407    .5209977
       _cons |  -25.41392   141.7312    -0.18   0.858    -307.7569    256.9291
------------------------------------------------------------------------------
```

What we obtain is a **precise zero** on the variable `perseyear`: This means that you fail to reject the value 0 and your estimator has a very small standard error (so that you would reject any big value).
Here with a 0.01 point estimate with se .07, and 75 degrees of freedom, the 95% CI is .01± 2(.07)=[−.13,.15]. These are small numbers compared to the estimated effect of -1.84 for *perse*.

(ii). Absence of Ashenfelter dip: The second verification to be done is that the law were not passed "in reaction" to a sharp increase in fatalities
If this was the case, what we may measure as effect of the program may be simply absence of the shock the next year.
We construct two dummy variables for the year prior to and 2 years before the law was passed:

```
gen perse_1=(year==perseyear-1)
gen perse_2=(year==perseyear-2)
```

and add them in the panel regression

```
. xtreg  totfatrte  perse perse_1 perse_2 i.year, fe i(state)

------------------------------------------------------------------------------
   totfatrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       perse |  -1.984322    .260309    -7.62   0.000    -2.495068   -1.473577
     perse_1 |    -.67682   .3903417    -1.73   0.083      -1.4427    .0890595
     perse_2 |  -.3457241   .4076816    -0.85   0.397    -1.145626    .4541778
             |
        year |
       1981  |  -1.785535   .4567127    -3.91   0.000    -2.681639   -.8894303
       1982  |  -4.355375   .4646388    -9.37   0.000    -5.267031   -3.443719
.................
```

We can see that they are not significantly different from 0 and in addition negative.

(iii). Finally robustness check, adding other policies that may be responsible for the decline in fatalities

```
seatbelt        =0 if none, =1 if primary, =2 if secondary
minage          minimum drinking age
bac10           blood alcohol limit .10
slnone          no speed limit
zerotol         zero tolerance law
gdl             graduated drivers license law

. xtreg  totfatrte  perse seatbelt minage slnone zerotol gdl i.year, fe i( state)

------------------------------------------------------------------------------
   totfatrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       perse |  -2.079465   .2494411    -8.34   0.000    -2.568889   -1.590041
    seatbelt |   .1725957   .1263679     1.37   0.172    -.0753485      .42054
      minage |   .3597417   .1146316     3.14   0.002     .1348252    .5846583
       bac10 |  -.2905969   .1939357    -1.50   0.134    -.6711148    .0899209
      slnone |  -.2599742   .9542762    -0.27   0.785    -2.132343    1.612394
     zerotol |    1.18105   .2877223     4.10   0.000     .6165153    1.745585
         gdl |  -.4026001   .3219036    -1.25   0.211    -1.034202    .2290014
.................
```

Conclusion: There is evidence that the administrative license revocation laws passed in different states between 1980 and 2004 had a strong effect in the reduction of traffic fatalities. It reduces fatalities by 1.8-2 per 100,000 population, over an average of 19 during this period. The result is robust to adding controls for other traffic laws passed by these states.