Bertrand, Duflo and Mulaihattan:

- Get female wage , employment data from CPS from 1979 to 1999 between ages 25 and 50.
- generate “fake” policy: choose a random year t from 1985 to 1999. Select 25 out of 50 states to designate those to have policy change from year t.
- Do DD estimation and see whether those fake policies are estimated to have a significant effect on female earnings. Repeat the policy simulation and estimation 200 times.
- The null of no policy effect is rejected 67.5 % of the time at 5%, significance level, when serial correlation was not taken into account for deriving the standard error.

- Sample size of women with positive earnings is 540,000. But the large sample at the individual level did not help.

$$y_{igt} = z_{igt}\gamma_{gt} + \lambda_t + \alpha_g + x_{gt}\beta + \nu_{gt} + u_{igt}$$

z_{igt}: individual specific covariates: age, etc.
g: group: states.
t: year
λ_t: year dummy coefficients
α_g: state dummy coefficients.
x_{gt}: policy variable. e.g. $I(t \in T_P) \times I(g \in G_P)$

The error term:

$$\nu_{gt} + u_{igt}$$

ν_{gt}: unobserved state/year effects.
Because of large sample size in i for each year and state cell, γ_{gt} can be estimated consistently. Then,

$$\hat{\delta}_{gt} = \overline{y}_{it} - \overline{z}_{igt}\hat{\gamma}_{gt} \approx \lambda_t + \alpha_g + x_{gt}\beta + \nu_{gt}$$

because of the large sample size within state/time cell, It is the serial correlation of ν_{gt} that creates overrejection. And large individual sample size does not alleviate the problem.

Variance Covariance Matrix of ν_{gt}

- Either the number of states g or number of times periods t needs to be large.
- For large g, one can flexibly estimate the Ω_{ν} as T by T matrix by assuming that ν_{gt} are uncorrelated across groups.
- For small g large T, one can estimate the standard error of the coefficients by using *HAC* estimation, where times series asymptotics are used.

Small Number of Policy Changes

Let

$$\hat{\delta}_{gt} = \overline{y}_{it} - \overline{z}_{igt}\hat{\gamma}_{gt} \approx \lambda_t + \alpha_g + x_{gt}\beta + \gamma d_{gt} + \nu_{gt}$$

- The effective sample size of policy variation $d_{gt}\beta$ is small, e.g., the sample size of $I(t \in T_P) \times I(g \in G_P) = 1$ is small.

Let N_0 be the number of first groups that face policy change during the sample period. Net N_1 be the number of the rest of the groups without policy change. Let

$$\overline{Z}_g = \frac{1}{T}\sum_{t=1}^{T} Z_{gt}$$

$$\overline{Z}_t = \frac{1}{N_0 + N_1}\sum_{g=1}^{N_0+N_1} Z_{gt}$$

$$\overline{Z} = \frac{1}{T}\frac{1}{N_0 + N_1}\sum_{t=1}^{T}\sum_{g=1}^{N_0+N_1} Z_{gt}$$

Let

$$\widetilde{Z_{gt}} = Z_{gt} - \overline{Z_g} - \overline{Z_t} + \overline{Z}$$

Then, after taking out the state and time fixed effects,

$$\widetilde{\delta_{gt}} = \widetilde{x_{gt}}\beta + \widetilde{d_{gt}}\gamma + \widetilde{\nu_{gt}}$$

- Assumption 1: Cross Section Independence: (x_{gt}, ν_{gt}) are independently and identically distributed across groups g.
- Variation in x left after controling for group and time dummies.

$$\frac{1}{T}\frac{1}{N_0 + N_1}\sum_{t=1}^{T}\sum_{g=1}^{N_0+N_1}\widetilde{x_{gt}}\widetilde{x_{gt}}' \rightarrow \Sigma_x$$

  Σ_x is finite and full rank.

Then, as $N_1 \to \infty$

$$\begin{aligned} \hat{\beta} &\xrightarrow{P} \beta \\ \hat{\gamma} - \gamma &\xrightarrow{P} \frac{\sum_{g=1}^{N_0} \sum_{t=1}^{T} \left(d_{gt} - \overline{d_g}\right)\left(\nu_{gt} - \overline{\nu_g}\right)}{\sum_{g=1}^{N_0} \sum_{t=1}^{T} \left(d_{gt} - \overline{d_g}\right)^2} \end{aligned}$$

If either $N_0 \to \infty$ or $T \to \infty$ then $\hat{\gamma} - \gamma \to 0$. But otherwise, RHS of $\hat{\gamma} - \gamma$ has a nondegenerate distribution.

How do we derive the distribution of

$$\frac{\sum_{g=1}^{N_0}\sum_{t=1}^{T}\left(d_{gt}-\overline{d_g}\right)\left(\nu_{gt}-\overline{\nu_g}\right)}{\sum_{g=1}^{N_0}\sum_{t=1}^{T}\left(d_{gt}-\overline{d_g}\right)^2} \ ?$$

For that, we need the distribution of the error term ν_{gt}. We get that from the group without policy change, whose sample size is large by the assumption $N_1 \to \infty$
Derive the sample distribution of

$$\widetilde{\delta_{gt}} - \widetilde{x_{gt}}\beta = \widetilde{\nu_{gt}}$$

using only sample in group 1, those without policy change.

Procedure for Inference

Let $\gamma = \gamma_0$ be the null hypothesis.

Step 1 Estimate $\hat{\beta}$ and $\hat{\gamma}$ from the data.

Step 2 Randomly draw N_0 groups $g_{1,1}, \ldots, g_{1,N_0}$ from group 1 without policy intervention.

Step 3 Calculate

$$(\gamma - \gamma_0)_m = \frac{\sum_{g=1}^{N_0} \sum_{t=1}^{T} \left(d_{gt} - \overline{d_g}\right) \left(\nu_{1,gt} - \overline{\nu_{1,g}}\right)}{\sum_{g=1}^{N_0} \sum_{t=1}^{T} \left(d_{gt} - \overline{d_g}\right)^2}$$

where

$$\nu_{1,gt} - \overline{\nu_{1,g}} = \widetilde{\delta_{1,gt}} - \widetilde{x_{1,gt}} \hat{\beta}$$

the residual from the regression of group 1.

Step 4 Repeat Step 2, Step 3, 100 times. Then, derive the cutoff point of highest 2.5 % ($(\gamma - \gamma_0)_H$) and lowest 2.5 % ($(\gamma - \gamma_0)_L$) cutoff points. If

$$(\gamma - \gamma_0)_L \leq \gamma - \gamma_0 \leq (\gamma - \gamma_0)_H$$

then, the hypothesis is not rejected. Otherwise, the hypothesis is rejected at 5% significance level.

Nonparametric Difference in Differences

Linear Difference in Differences with 2 groups and periods.

- Assume there are only 2 periods, $t = 1$ is policy period and $t = 0$ is not.
- Assume there are only 2 groups, treatment group $g = 1$ and control group $g = 0$.

Then, the linear DD model:

$$y_{igt} = \beta_0 + \beta_1 I_1(t) + \beta_2 I_1(g) + \beta_3 I_1(g) \times I_1(t) + \epsilon_{it}$$

Then,

$$\begin{aligned} E[y_{01}] - E[y_{00}] &= \beta_1 \\ E[y_{11}] - E[y_{10}] &= \beta_1 + \beta_3 \end{aligned}$$

Then, the policy effect is

$$\beta_3 = \{E[y_{11}] - E[y_{10}]\} - \{E[y_{01}] - E[y_{00}]\}$$

and can be estimated as

$$\hat{\beta}_3 = [\overline{y}_{11} - \overline{y}_{10}] - [\overline{y}_{01} - \overline{y}_{00}]$$

	group 0	group 1	difference
time 0	$\overline{y}_{00} = \hat{\beta}_0$	$\overline{y}_{10} = \hat{\beta}_0 + \hat{\beta}_2$	$\hat{\beta}_2$
time 1	$\overline{y}_{01} = \hat{\beta}_0 + \hat{\beta}_1$	$\overline{y}_{11} = \hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$	$\hat{\beta}_2 + \hat{\beta}_3$
difference	$\overline{y}_{01} - \overline{y}_{00} = \hat{\beta}_1$	$\overline{y}_{11} - \overline{y}_{10} = \hat{\beta}_1 + \hat{\beta}_3$	$\hat{\beta}_3$

That is, the outcome of individuals in period 1 if there was no policy would be

$$\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2$$

Nonparametric Analysis

Now, instead of a linear model, what we have in the nonparametric analysis, we have the four distributions.

	group 0	group 1
time 0	$F_{00}(y)$	$F_{10}(y)$
time 1	$F_{01}(y)$	$F_{11}(y)$

- In group 0 (control group) period 1, the individual had income y_{01}. What is the percentile of her income? $u = F_{01}(y_{01})$ percentile.
- What is the u percentile of income in period 0? Income which results in

$$F_{00}(y) = u = F_{01}(y_{01})$$

i.e.

$$y_{00} = F_{00}^{-1}\left(F_{01}(y_{01})\right)$$

- In group 0, u percentile income, which is y_{01} in period 1, changed from $y_{00} = F_{00}^{-1}(F_{01}(y_{01}))$ to y_{01}
- Let $F_{11}^{(0)}(y_{11})$ be the hypothetical distribution of group 1 period 1 income without policy. If we assume without policy, there is no difference in time trend between group 0 and group 1, then $v = F_{10}(y_{00})$ percentile income in group 1 should move from y_{00} to $y = y_{01}$.

		time 0	time 1
group 0	Dist.	$F_{00}()$	$F_{01}()$
	Percentile	u	$u = F_{01}(y)$
	income	$y_{00} = F_{00}^{-1}(F_{01}(y))$	y
group 1	Dist.	$F_{10}()$	$F_{11}^{(0)}()$
	Percentile	$v = F_{10}(y_{00})$	v
	income	$y_{00} = F_{00}^{-1}(F_{01}(y))$	y

Hence, percentile of income y in group 1 in period 1 without policy is

$$v = F_{10}\left(F_{00}^{-1}(F_{01}(y))\right)$$

Hence, hypothetical income distribution of group 1 period 1 without policy is

$$F_{11}^{(0)}(y) = F_{10}\left(F_{00}^{-1}(F_{01}(y))\right)$$

Another way to look

		time 0	time 1
group 0	Dist.	$F_{00}()$	$F_{01}()$
	Percentile	$u = F_{00}(y)$	u
	income	y	$y_{01} = F_{01}^{-1}(F_{00}(y))$
group 1	Dist.	$F_{10}()$	$F_{11}^{(0)}()$
	Percentile	$v = F_{10}(y_{10})$	v
	income	y_{10}	$y_{11}^{(0)} = F_{01}^{-1}(F_{00}(y_{10}))$

Time trend moves v th percentile income of group 1, which was y_{10} in period 0 to $y_{11}^{(0)} = F_{01}^{-1}(F_{00}(y_{10}))$ in period 1.

- The Average Treatment Effect of Change in Change framework is

$$\begin{aligned} \tau &= E\left[Y_{11}\right] - E\left[Y_{11}^{(0)}\right] \\ &= E\left[Y_{11}\right] - E\left[F_{01}^{-1}\left(F_{00}(Y_{10})\right)\right] \end{aligned}$$

- The estimator of the Average Treatment Effect is

$$\hat{\tau} = N_{11}^{-1}\sum_{i=1}^{N_{11}} Y_{11,i} - N_{10}^{-1}\sum_{i=1}^{N_{10}} \hat{F}_{01}^{-1}\left(\hat{F}_{00}(Y_{10,i})\right)$$