Combining Difference-in-difference and Matching for Panel Data Analysis

Weihua An

Departments of Sociology and Statistics
Indiana University

July 28, 2016

Research Interests

- Network Analysis
  - Social influence and networks
  - Network and measurement
  - Text networks (social media, citation, biographies, sports records)
- Causal Inference
  - Matching and propensity score methods
  - Instrumental variable methods
  - Causal inference under interference
- Applied Research
  - Social policy (e.g., network and neighborhood)
  - Organizations (e.g., network and cognition)

Outline

- Previous Parametric Models for Panel Data Analysis
- Combining Difference-in-difference and Matching
- One Example
- Summary

Previous Parametric Models

Assume that in the absence of treatment, the potential outcome for subject i at time t is

$$Y_{it}^0 = \lambda_t + X_{it}\gamma + C_i + e_{it},$$

Assuming an additive constant treatment δ, we can write the potential outcome for subject i at time t under treatment as

$$Y_{it}^1 = Y_{it}^0 + \delta.$$

This implies that using the observed outcome, we can write

$$Y_{it} = \lambda_t + \delta D_{it} + X_{it}\gamma + C_i + e_{it}, \tag{1}$$

The detail of the models and methods can be found in chapter 9 of Morgan and Winship (2007) and Wooldridge (2001).

Parametric Models

- Random effects: $C_i \sim N(0, \sigma^2)$
- Fixed effects: using differenced outcomes to remove C_i.
  - FE model: $Y_{it} - \bar{Y}_i = \delta(D_{it} - \bar{D}_i) + (X_{it} - \bar{X}_i)\gamma + (e_{it} - \bar{e}_i)$
  - FD model: $\Delta Y_{it} = \Delta\lambda_t + \delta\Delta D_{it} + \Delta X_{it}\gamma + \Delta e_{it}$
- Random trend and slope: $Y_{it} = \lambda_t + g_i t + \delta D_{it} + h_i D_{it} + X_{it}\gamma + C_i + e_{it}$
- MA(1): $e_{it} = \rho e_{it-1} + v_{it}$
- AR(1): $Y_{it} = \lambda_t + \theta Y_{it-1} + \delta D_{it} + \delta_1 D_{it-1} + X_{it}\gamma + C_i + e_{it}$

Matching

Definitions of treatment effects:

- ATE: $\delta = E[Y^1 - Y^0]$
- ATT: $\delta_1 = E[Y^1 - Y^0 | D = 1]$

The basic idea of matching is to match units with exact covariates in the opposite treatment group to impute the missing potential outcomes.

The key of matching is to measure similarity between units. The distance between two units in their covariate values is often measured according to

$$d(X_i, X_j) = \sqrt{(X_i - X_j)^T W^{-1} (X_i - X_j)}.$$

There are two popular choices of the weight matrix: (1) the variance-covariance matrix of X; and (2) the sample variances of X.

Estimating ATE:

$$\hat{\delta} = \frac{1}{N} \sum_{i=1}^{N} (\hat{Y}_i^1 - \hat{Y}_i^0) = \frac{1}{N} \sum_{i=1}^{N} (2D_i - 1)(1 + K_i) Y_i$$

$$\sigma_{\hat{\delta}}^2 = \frac{1}{N^2} \sum_{i=1}^{N} \{1 + K_i\}^2 \sigma^2 (Y_i | X_i, D_i)$$

where K_i is the number of times unit i serves as a match to other units.

Difference-in-difference (DID)

DID assumes that in the absence of treatment the original difference between treated units (denoted by (1)) and control units (denoted by (0)) in the outcome will remain constant over time.

$$E[Y^0_{t-1}(1) - Y^0_{t-1}(0)] = E[Y^0_t(1) - Y^0_t(0)]$$

which implies $E[Y^0_t(1) - Y^0_{t-1}(1)] = E[Y^0_t(0) - Y^0_{t-1}(0)]$, or say, $\Delta Y^0_t \perp D_t$.

With this assumption, we can estimate the ATT as $\Delta Y_t(1) - \Delta Y_t(0)$.

$$\begin{aligned}
E[\Delta Y_t(1) - \Delta Y_t(0)] &= E[Y_t(1) - Y_{t-1}(1) - (Y_t(0) - Y_{t-1}(0))] \\
&= E[Y^1_t(1) - Y^0_{t-1}(1) - (Y^0_t(0) - Y^0_{t-1}(0))] \\
&= E[Y^1_t(1) - Y^0_{t-1}(1) - (Y^0_t(1) - Y^0_{t-1}(1))] \\
&= E[Y^1_t(1) - Y^0_t(1)] \\
&= ATT
\end{aligned}$$

DID + Matching

Assumptions:

- Strong form of ignorability

$$\Delta Y_t^1, \Delta Y_t^0 \perp D_t \mid \overrightarrow{X_t}, \overrightarrow{D_{t-1}}. \tag{2}$$

- Weak form of ignorability:

$$\Delta Y_t^1, \Delta Y_t^0 \perp D_t \mid \overrightarrow{X_{t,s}}, \overrightarrow{D_{t-1,s}}. \tag{3}$$

In short, what we propose is to first-difference the outcome and then apply matching to estimate treatment effects at each wave.

- Using the differenced outcome helps remove the effects of time-invariant confounding factors.
- Matching is nonparametric and helps balance covariates and create a more focused causal inference.

Aggregating Treatment Effects across Waves

$$\hat{\delta}_W = \sum_{t=2}^{T} \frac{N_t}{N} \hat{\delta}_t$$

$$\hat{\sigma}^2_{\hat{\delta}_W} = \sum_{t=2}^{T} \frac{N_t^2}{N^2} \hat{\sigma}^2_{\hat{\delta}_t} + 2 \sum_{2 \leq g < h \leq T} \frac{N_g N_h}{N^2} \mathrm{Cov}(\hat{\delta}_g, \hat{\delta}_h)$$

where $\mathrm{Cov}(\hat{\delta}_g, \hat{\delta}_h)$ is the covariance of treatment effects across waves g and h.

Estimating the covariance terms:

$$\hat{\delta}_g = \frac{1}{N_g} \sum_{i=1}^{N_g} (2D_{ig} - 1)(1 + K_{ig}) \Delta Y_{ig}$$

$$\hat{\delta}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} (2D_{ih} - 1)(1 + K_{ih}) \Delta Y_{ig}$$

$$\mathrm{Cov}(\hat{\delta}_g, \hat{\delta}_h) = \mathrm{Cov}(\frac{1}{n} \sum_{i=1}^{n} J_{ig} \Delta Y_{ig}, \frac{1}{n} \sum_{i=1}^{n} J_{ih} \Delta Y_{ih}) = \frac{1}{n^2} \sum_{i=1}^{n} J_{ig} J_{ih} \mathrm{Cov}(\Delta Y_{ig}, \Delta Y_{ih})$$

where n is the number of common observations in waves g and h,
$J_{ig} = (2D_{ig} - 1)(1 + K_{ig})$, and $J_{ih} = (2D_{ih} - 1)(1 + K_{ih})$.

One Example

- The goal is to study race-of-interviewer effects (ROIE) in the General Social Survey (GSS). Specifically, we are interested in whether an interviewer's race will affect respondents' responses.
- Ten outcome measures: intelligence gap between blacks and whites, respondents' views on the government's responsibility and spending and their confidence in political and financial institutions.
- The major explanatory variable: interviewer's race($1 =$ black; $0 =$ otherwise). Other interviewer's characteristics include age, sex, and years of working for the GSS.
- Control variables: respondent's age, sex, race, employment status, family income, number of children, marital status, years of education, party identification, religious denomination, residential place (e.g., large city vs. small town), and residential region.

Table 3. Estimated Race of Interviewer Effects on Ten Selected Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Outcomes	RE	FE	FD	RTS	MA(1)	AR(1)
Perceived Intelligence Gap	0.43***	0.36***	0.32**	0.43***	0.43***	0.27
	0.07	0.09	0.10	0.07	0.07	0.15
	2,616	2,617	1,407	2,616	2,616	658
Confidence in Executive Branch of Fed. Govt.	0.03	-0.00	0.02	0.03	0.04	-0.09
	0.04	0.05	0.06	0.04	0.04	0.12
	2,695	2,696	1,449	2,695	2,695	678
Confidence in Congress	0.07*	0.06	0.08	0.07*	0.07*	0.09
	0.04	0.04	0.05	0.04	0.04	0.09
	2,700	2,701	1,451	2,700	2,700	683
Confidence in Supreme Court	0.07	0.03	0.07	0.07	0.08*	-0.10
	0.04	0.05	0.05	0.04	0.04	0.09
	2,684	2,685	1,432	2,684	2,684	668
Confidence in Bank and Financial Institutions	0.05	-0.03	-0.06	0.05	0.05	-0.02
	0.04	0.05	0.05	0.04	0.04	0.11
	2,713	2,714	1,461	2,713	2,713	688
Spending on Welfare	0.19***	0.19**	0.16*	0.19***	0.19***	0.26*
	0.05	0.06	0.07	0.05	0.05	0.11
	1,990	1,991	1,059	1,990	1,990	493
Spending on Blacks	0.32***	0.29***	0.28***	0.32***	0.32***	0.30*
	0.04	0.05	0.05	0.04	0.04	0.13
	1,858	1,859	938	1,858	1,858	422
Should Help Blacks	0.57***	0.53***	0.47***	0.56***	0.57***	0.46***
	0.07	0.09	0.09	0.07	0.07	0.16
	2,662	2,663	1,405	2,662	2,662	645
Should Help Poor	0.08	0.01	0.00	0.08	0.08	-0.02
	0.07	0.08	0.08	0.07	0.06	0.15
	2,676	2,677	1,423	2,676	2,676	653
Should Help Sick	0.11	0.07	0.02	0.11	0.11	0.03
	0.07	0.09	0.10	0.07	0.07	0.17
	2,683	2,684	1,428	2,683	2,683	659

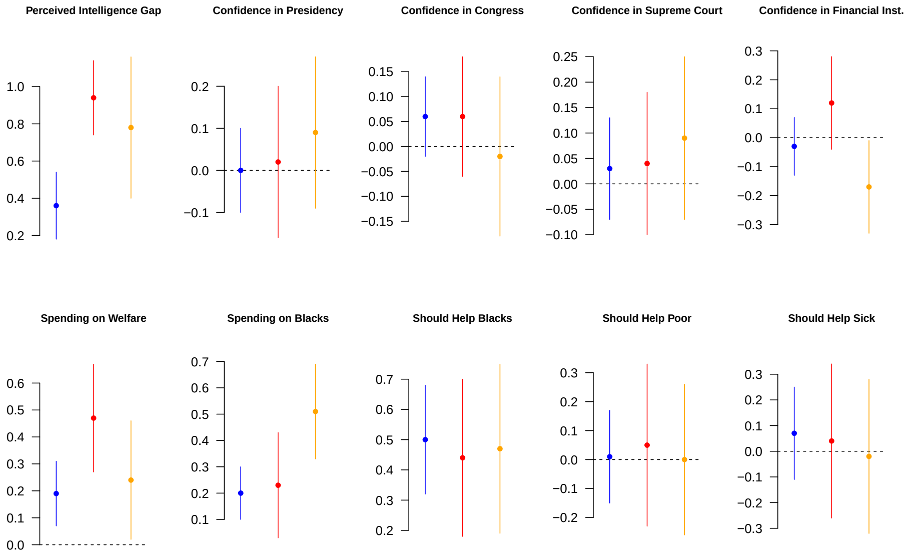
Table A4. Matching Estimates of the Race of Interviewer Effects at Waves 2 and 3

	ATE			ATT		
Outcomes	Est	SE	N	Est	SE	N
Wave 2						
Perceived Intelligence Gap	0.73	0.25**	734	1.29	0.32***	82
Confidence in Executive Branch of Fed. Govt.	0.05	0.11	763	-0.03	0.11	94
Confidence in Congress	0.12	0.09	762	0.02	0.10	94
Confidence in Supreme Court	0.13	0.10	753	0.00	0.12	92
Confidence in Bank and Financial Institutions	-0.09	0.11	771	-0.16	0.12	95
Spending on Welfare	0.16	0.19	559	0.13	0.16	66
Spending on Blacks	0.46	0.15**	494	0.54	0.13***	63
Should Help Blacks	0.52	0.20**	750	0.50	0.21*	89
Should Help Poor	0.07	0.17	756	-0.33	0.19	90
Should Help Sick	-0.19	0.21	763	-0.33	0.22	93
Wave 3						
Perceived Intelligence Gap	1.17	0.17***	672	0.32	0.21	91
Confidence in Executive Branch of Fed. Govt.	-0.01	0.15	686	0.21	0.13	102
Confidence in Congress	-0.01	0.09	689	-0.05	0.12	102
Confidence in Supreme Court	-0.06	0.10	679	0.17	0.11	100
Confidence in Bank and Financial Institutions	0.36	0.11***	690	-0.18	0.12	102
Spending on Welfare	0.81	0.17***	500	0.34	0.14*	74
Spending on Blacks	-0.03	0.12	444	0.49	0.12***	65
Should Help Blacks	0.35	0.17*	655	0.45	0.19*	97
Should Help Poor	0.03	0.22	667	0.30	0.19	99
Should Help Sick	0.31	0.22	665	0.27	0.20	102

Table 6. Combined Matching Estimates of the Race of Interviewer Effects from Waves 2 and 3

	ATE			ATT		
Outcomes	Est	SE	N	Est	SE	N
Perceived Intelligence Gap	0.94	0.15***	1,406	0.78	0.19***	173
Confidence in Executive Branch of Fed. Govt.	0.02	0.09	1,449	0.09	0.09	196
Confidence in Congress	0.06	0.06	1,451	-0.02	0.08	196
Confidence in Supreme Court	0.04	0.07	1,432	0.09	0.08	192
Confidence in Bank and Financial Institutions	0.12	0.08	1,461	-0.17	0.08*	197
Spending on Welfare	0.47	0.13***	1,059	0.24	0.11**	140
Spending on Blacks	0.23	0.10**	938	0.51	0.09***	128
Should Help Blacks	0.44	0.13***	1,405	0.47	0.14***	186
Should Help Poor	0.05	0.14	1,423	0.00	0.13	189
Should Help Sick	0.04	0.15	1,428	-0.02	0.15	195

Comparing the Results



Summary

- Race-of-interviewer effects seems to concentrate on racially charged survey items, but not on other items.
- Developing "**DIDMatch**" in *Stata* to implement the method. The key is to estimate $\text{Cov}(\hat{\delta}_g, \hat{\delta}_h) = \frac{1}{n^2}\sum_{i=1}^{n} J_{ig} J_{ih} \text{Cov}(\Delta Y_{ig}, \Delta Y_{ih})$.