Matching Methods for Causal Inference with Time-Series Cross-Sectional Data

Kosuke Imai    In Song Kim    Erik Wang
Harvard    MIT    Princeton

Keynote Talk at
Latin American Political Methodology Meeting

November 9, 2018
Motivation and Overview

- Matching methods have become part of toolkit for social scientists
  1. reduces model dependence in observational studies
  2. provides diagnostics through balance checks
  3. clarifies comparison between treated and control units

- Yet, almost all existing matching methods deal with cross-sectional data

- We propose a matching method for time-series cross-sectional data
  1. create a matched set for each treated observation
  2. refine the matched set via any matching or weighting method
  3. compute the difference-in-differences estimator

- Provide a model-based standard error

- Develop an open-source software package PanelMatch

- Empirical applications:
  - Democracy and economic growth (Acemoglu et al.)
  - Interstate war and inheritance tax (Scheve & Stasavage)
Democracy and Economic Growth

- Acemoglu et al. (2017): an up-to-date empirical study of the long-standing question in political economy
- TSCS data set: 184 countries from 1960 to 2010
- Dynamic linear regression model with fixed effects:

\[ Y_{it} = \alpha_i + \gamma_t + \beta X_{it} + \sum_{\ell=1}^{4} \left\{ \rho_{\ell} Y_{i,t-\ell} + \zeta_{\ell}^T Z_{i,t-\ell} \right\} + \epsilon_{it} \]

- \( X_{it} \): binary democracy indicator
- \( Y_{it} \): log real GDP per capita
- \( Z_{it} \): time-varying covariates (population, trade, social unrest, etc.)

- Sequential exogeneity assumption:

\[ \mathbb{E}(\epsilon_{it} | \{ Y_{it'} \}_{t'=1}^{t-1}, \{ X_{it'} \}_{t'=1}^{t}, \{ Z_{it'} \}_{t'=1}^{t-1}, \alpha_i, \gamma_t) = 0 \]

- Nickell bias \( \Rightarrow \) GMM estimation with instruments (Arellano & Bond)
### Regression Results

<table>
<thead>
<tr>
<th></th>
<th>ATE ($\hat{\beta}$)</th>
<th>0.787 (0.226)</th>
<th>0.875 (0.374)</th>
<th>0.666 (0.307)</th>
<th>0.917 (0.461)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}_1$</td>
<td>1.238</td>
<td>1.204 (0.041)</td>
<td>1.100 (0.042)</td>
<td>1.046 (0.043)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.207</td>
<td>-0.193 (0.038)</td>
<td>-0.133 (0.042)</td>
<td>-0.121 (0.043)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\rho}_2$</td>
<td>-0.026</td>
<td>-0.028 (0.045)</td>
<td>0.005 (0.041)</td>
<td>0.014 (0.038)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\rho}_3$</td>
<td>-0.043</td>
<td>-0.036 (0.028)</td>
<td>0.003 (0.030)</td>
<td>-0.018 (0.029)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\rho}_4$</td>
<td>-0.026</td>
<td>-0.020 (0.017)</td>
<td>0.024 (0.024)</td>
<td>-0.023 (0.023)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>country FE</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>time FE</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>time trends</td>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>covariates</td>
<td></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>estimation</td>
<td></td>
<td>OLS</td>
<td>GMM</td>
<td>OLS</td>
<td>GMM</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>6,336</td>
<td>4,416</td>
<td>6,161</td>
<td>4,245</td>
</tr>
</tbody>
</table>
Regression models does not tell us where the variation comes from.

Estimation of counterfactual outcomes requires comparison between treated and control observations.

Identification strategy:
- within-unit over-time variation
- within-time across-units variation
Inheritance tax plays a central role in wealth accumulations and income inequality.

Scheve and Stasavage (2012): war increases inheritance taxation.

TSCS Data: 19 countries over 185 years from 1816 to 2000.

Static model:

\[ Y_{it} = \alpha_i + \gamma_t + \beta X_{i,t-1} + \delta^\top Z_{i,t-1} + \lambda_i t + \epsilon_{it} \]

- \( X_{i,t-1} \): interstate war for country \( i \) in year \( t - 1 \)
- \( Y_{it} \): top rate of inheritance tax
- \( Z_{i,t-1} \): time-varying covariates (leftist executive, a binary variable for the universal male suffrage, and logged real GDP per capita)
- Strict exogeneity:

\[ \mathbb{E}(\epsilon_{it} \mid X_i, Z_i, \alpha_i, \gamma_t, \lambda_i) = 0 \]

where \( X_i = (X_{i1}, X_{i2}, \ldots, X_{iT}) \) and \( Z_i = (Z_{i1}^\top, Z_{i2}^\top, \ldots, Z_{iT}^\top)^\top \)
Dynamic model without country fixed effects:

\[ Y_{it} = \gamma_t + \beta X_{i,t-1} + \rho Y_{i,t-1} + \delta Z_{i,t-1} + \lambda_i t + \epsilon_{it} \]

where the strict exogeneity assumption is now given by,

\[ \mathbb{E}(\epsilon_{it} \mid X_i, Z_i, Y_{i,t-1}, \gamma_t, \lambda_i) = 0 \]

Regression results:

<table>
<thead>
<tr>
<th></th>
<th>ATE ((\hat{\beta}))</th>
<th>(\hat{\rho}_1)</th>
<th>country FE</th>
<th>time FE</th>
<th>time trends</th>
<th>covariates</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.775 (2.392)</td>
<td>0.908 (0.014)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>2,780</td>
</tr>
<tr>
<td></td>
<td>1.745 (0.729)</td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>2,537</td>
</tr>
<tr>
<td></td>
<td>5.970 (2.081)</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>2,779</td>
</tr>
<tr>
<td></td>
<td>1.636 (0.757)</td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>2,536</td>
</tr>
</tbody>
</table>
Treatment is concentrated in a few years

How should we estimate counterfactual outcomes?
Quantity of Interest and Assumptions

- Choose number of lags $L = 2, \ldots$, for confounder adjustment
- Choose number of leads, $F = 0, 1, \ldots$, for short or long term effects
- Average Treatment Effect of Policy Change for the Treated (ATT):

$$\mathbb{E} \left\{ Y_{i,t+F} \left( X_{it} = 1, X_{i,t-1} = 0, \{ X_{i,t-\ell} \}_{\ell=2}^L \right) - Y_{i,t+F} \left( X_{it} = 0, X_{i,t-1} = 0, \{ X_{i,t-\ell} \}_{\ell=2}^L \right) \mid X_{it} = 1, X_{i,t-1} = 0 \right\}$$

Assumptions:

1. No spillover effect
2. Limited carryover effect (up to $L$ time periods)
3. Parallel trend after conditioning:

$$\mathbb{E} \left[ Y_{i,t+F} \left( X_{it} = X_{i,t-1} = 0, \{ X_{i,t-\ell} \}_{\ell=2}^L \right) - Y_{i,t-1} \mid X_{it} = 1, X_{i,t-1} = 0, \{ X_{i,t-\ell}, Y_{i,t-\ell} \}_{\ell=2}^L, \{ Z_{i,t-\ell} \}_{\ell=0}^L \right]$$

$$= \mathbb{E} \left[ Y_{i,t+F} \left( X_{it} = X_{i,t-1} = 0, \{ X_{i,t-\ell} \}_{\ell=2}^L \right) - Y_{i,t-1} \mid X_{it} = 0, X_{i,t-1} = 0, \{ X_{i,t-\ell}, Y_{i,t-\ell} \}_{\ell=2}^L, \{ Z_{i,t-\ell} \}_{\ell=0}^L \right]$$

Constructing Matched Sets

- Control units with identical treatment history from time \( t - L \) to \( t - 1 \)
- Construct a matched set for each treated observation
- Formal definition:

\[
\mathcal{M}_{it} = \{ i' : i' \neq i, X_{i't} = 0, X_{i't'} = X_{it'} \text{ for all } t' = t - 1, \ldots, t - L \}
\]

- Some treated observations have no matched control
  \( \Rightarrow \) change the quantity of interest by dropping them
- Similar to the risk set of Li et al. (2001) but we do not exclude those who already receive the treatment
## An Example of Matched Set

<table>
<thead>
<tr>
<th></th>
<th>Country</th>
<th>Year</th>
<th>Democracy</th>
<th>logGDP</th>
<th>Population</th>
<th>Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Argentina</td>
<td>1974</td>
<td>1</td>
<td>888.20</td>
<td>29.11</td>
<td>14.45</td>
</tr>
<tr>
<td>2</td>
<td>Argentina</td>
<td>1975</td>
<td>1</td>
<td>886.53</td>
<td>29.11</td>
<td>12.61</td>
</tr>
<tr>
<td>3</td>
<td>Argentina</td>
<td>1976</td>
<td>0</td>
<td>882.91</td>
<td>29.15</td>
<td>12.11</td>
</tr>
<tr>
<td>4</td>
<td>Argentina</td>
<td>1977</td>
<td>0</td>
<td>888.09</td>
<td>29.32</td>
<td>15.15</td>
</tr>
<tr>
<td>5</td>
<td>Argentina</td>
<td>1978</td>
<td>0</td>
<td>881.99</td>
<td>29.57</td>
<td>15.54</td>
</tr>
<tr>
<td>6</td>
<td>Argentina</td>
<td>1979</td>
<td>0</td>
<td>890.24</td>
<td>29.85</td>
<td>15.93</td>
</tr>
<tr>
<td>7</td>
<td>Argentina</td>
<td>1980</td>
<td>0</td>
<td>892.81</td>
<td>30.12</td>
<td>12.23</td>
</tr>
<tr>
<td>8</td>
<td>Argentina</td>
<td>1981</td>
<td>0</td>
<td>885.43</td>
<td>30.33</td>
<td>11.39</td>
</tr>
<tr>
<td>9</td>
<td>Argentina</td>
<td>1982</td>
<td>0</td>
<td>878.82</td>
<td>30.62</td>
<td>13.40</td>
</tr>
<tr>
<td>10</td>
<td>Thailand</td>
<td>1974</td>
<td>1</td>
<td>637.24</td>
<td>43.32</td>
<td>37.76</td>
</tr>
<tr>
<td>11</td>
<td>Thailand</td>
<td>1975</td>
<td>1</td>
<td>639.51</td>
<td>42.90</td>
<td>41.63</td>
</tr>
<tr>
<td>12</td>
<td>Thailand</td>
<td>1976</td>
<td>0</td>
<td>645.97</td>
<td>42.44</td>
<td>42.33</td>
</tr>
<tr>
<td>13</td>
<td>Thailand</td>
<td>1977</td>
<td>0</td>
<td>653.02</td>
<td>41.92</td>
<td>43.21</td>
</tr>
<tr>
<td>14</td>
<td>Thailand</td>
<td>1978</td>
<td>1</td>
<td>660.57</td>
<td>41.39</td>
<td>42.66</td>
</tr>
<tr>
<td>15</td>
<td>Thailand</td>
<td>1979</td>
<td>1</td>
<td>663.64</td>
<td>40.82</td>
<td>45.27</td>
</tr>
<tr>
<td>16</td>
<td>Thailand</td>
<td>1980</td>
<td>1</td>
<td>666.57</td>
<td>40.18</td>
<td>46.69</td>
</tr>
<tr>
<td>17</td>
<td>Thailand</td>
<td>1981</td>
<td>1</td>
<td>670.27</td>
<td>39.44</td>
<td>53.40</td>
</tr>
<tr>
<td>18</td>
<td>Thailand</td>
<td>1982</td>
<td>1</td>
<td>673.52</td>
<td>38.59</td>
<td>54.22</td>
</tr>
</tbody>
</table>
Refining Matched Sets

- Make additional adjustments for past outcomes and confounders
- Use any matching or weighting method
- **Mahalanobis distance matching:**
  1. Compute the distance between treated and matched control obs.

\[
S_{it}(i') = \frac{1}{L} \sum_{\ell=1}^{L} \sqrt{\left(V_{i,t-\ell} - V_{i',t-\ell}\right)^\top \Sigma_{i,t-\ell}^{-1} \left(V_{i,t-\ell} - V_{i',t-\ell}\right)}
\]

where \(V_{it'} = (Y_{it'}, Z_{i,t'+1}^\top)^\top\) and \(\Sigma_{it'} = \text{Cov}(V_{it'})\)

2. Match the most similar \(J\) matched control observations

- **Propensity score weighting:**
  1. Estimate propensity score

\[
e_{it}(\{V_{i,t-\ell}\}_{\ell=1}^{L}) = \Pr(X_{it} = 1 | \{V_{i,t-\ell}\}_{\ell=1}^{L})
\]

2. Weight each matched control observation
An Example of Refinement

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Democracy</th>
<th>logGDP</th>
<th>Population</th>
<th>Trade</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>1979</td>
<td>0</td>
<td>890.24</td>
<td>29.85</td>
<td>15.93</td>
<td>1.00</td>
</tr>
<tr>
<td>Argentina</td>
<td>1980</td>
<td>0</td>
<td>892.81</td>
<td>30.12</td>
<td>12.23</td>
<td>1.00</td>
</tr>
<tr>
<td>Argentina</td>
<td>1981</td>
<td>0</td>
<td>885.43</td>
<td>30.33</td>
<td>11.39</td>
<td>1.00</td>
</tr>
<tr>
<td>Argentina</td>
<td>1982</td>
<td>0</td>
<td>878.82</td>
<td>30.62</td>
<td>13.40</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Argentina</strong></td>
<td><strong>1983</strong></td>
<td><strong>1</strong></td>
<td><strong>881.09</strong></td>
<td><strong>30.75</strong></td>
<td><strong>16.46</strong></td>
<td><strong>1.00</strong></td>
</tr>
<tr>
<td>Argentina</td>
<td>1984</td>
<td>1</td>
<td>881.76</td>
<td>30.77</td>
<td>15.67</td>
<td>1.00</td>
</tr>
<tr>
<td>Mali</td>
<td>1979</td>
<td>0</td>
<td>542.02</td>
<td>43.80</td>
<td>31.18</td>
<td>0.26</td>
</tr>
<tr>
<td>Mali</td>
<td>1980</td>
<td>0</td>
<td>535.65</td>
<td>43.96</td>
<td>41.82</td>
<td>0.26</td>
</tr>
<tr>
<td>Mali</td>
<td>1981</td>
<td>0</td>
<td>529.10</td>
<td>44.07</td>
<td>41.92</td>
<td>0.26</td>
</tr>
<tr>
<td>Mali</td>
<td>1982</td>
<td>0</td>
<td>522.25</td>
<td>44.45</td>
<td>42.53</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>Mali</strong></td>
<td><strong>1983</strong></td>
<td><strong>0</strong></td>
<td><strong>524.84</strong></td>
<td><strong>44.74</strong></td>
<td><strong>43.65</strong></td>
<td><strong>0.26</strong></td>
</tr>
<tr>
<td>Mali</td>
<td>1984</td>
<td>0</td>
<td>527.13</td>
<td>44.95</td>
<td>45.92</td>
<td>0.26</td>
</tr>
<tr>
<td>Chad</td>
<td>1979</td>
<td>0</td>
<td>506.71</td>
<td>44.61</td>
<td>44.80</td>
<td>0.27</td>
</tr>
<tr>
<td>Chad</td>
<td>1980</td>
<td>0</td>
<td>498.36</td>
<td>44.84</td>
<td>45.75</td>
<td>0.27</td>
</tr>
<tr>
<td>Chad</td>
<td>1981</td>
<td>0</td>
<td>497.18</td>
<td>45.07</td>
<td>51.58</td>
<td>0.27</td>
</tr>
<tr>
<td>Chad</td>
<td>1982</td>
<td>0</td>
<td>500.07</td>
<td>45.44</td>
<td>43.97</td>
<td>0.27</td>
</tr>
<tr>
<td><strong>Chad</strong></td>
<td><strong>1983</strong></td>
<td><strong>0</strong></td>
<td><strong>512.20</strong></td>
<td><strong>45.76</strong></td>
<td><strong>29.22</strong></td>
<td><strong>0.27</strong></td>
</tr>
<tr>
<td>Chad</td>
<td>1984</td>
<td>0</td>
<td>511.63</td>
<td>46.04</td>
<td>29.91</td>
<td>0.27</td>
</tr>
<tr>
<td>Uruguay</td>
<td>1979</td>
<td>0</td>
<td>858.39</td>
<td>27.23</td>
<td>41.51</td>
<td>0.47</td>
</tr>
<tr>
<td>Uruguay</td>
<td>1980</td>
<td>0</td>
<td>863.39</td>
<td>27.04</td>
<td>37.99</td>
<td>0.47</td>
</tr>
<tr>
<td>Uruguay</td>
<td>1981</td>
<td>0</td>
<td>864.28</td>
<td>26.93</td>
<td>36.20</td>
<td>0.47</td>
</tr>
<tr>
<td>Uruguay</td>
<td>1982</td>
<td>0</td>
<td>853.36</td>
<td>26.86</td>
<td>35.84</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Uruguay</strong></td>
<td><strong>1983</strong></td>
<td><strong>0</strong></td>
<td><strong>841.87</strong></td>
<td><strong>26.83</strong></td>
<td><strong>33.36</strong></td>
<td><strong>0.47</strong></td>
</tr>
<tr>
<td>Uruguay</td>
<td>1984</td>
<td>0</td>
<td>840.08</td>
<td>26.82</td>
<td>42.98</td>
<td>0.47</td>
</tr>
</tbody>
</table>
The Difference-in-Differences Estimator

- Compute the weighted average of difference-in-differences among matched control observations
- Weights are based on refinement
- A synthetic control for each treated observation

The DiD estimator:

$$\frac{1}{N \sum_{i=1}^{T-F} D_{it}} \sum_{i=1}^{N} \sum_{t=L+1}^{T-F} D_{it} \left\{ (Y_{i,t+F} - Y_{i,t-1}) - \sum_{i' \in M_{it}} W_{it}' (Y_{i',t+F} - Y_{i',t-1}) \right\}$$

- Equivalent to the weighted two-way fixed effects estimator:

$$\arg\min_{\beta} \sum_{i=1}^{N} \sum_{t=1}^{T} W_{it} \left\{ (Y_{it} - \bar{Y}_i - \bar{Y}_t + \bar{Y}) - \beta (X_{it} - \bar{X}_i - \bar{X}_t + \bar{X}) \right\}^2$$
Balance for covariate $j$ at time $t - \ell$ in each matched set:

$$B_{it}(j, \ell) = \frac{V_{i, t-\ell, j} - \sum_{i' \in M_{it}} w_{it}^{i'} V_{i', t-\ell, j}}{\sqrt{\frac{1}{N_1 - 1} \sum_{i'=1}^{N} \sum_{t'=L+1}^{T-F} D_{it'} (V_{i', t'-\ell, j} - \overline{V}_{t'-\ell, j})^2}}$$

Average this measure across all treated observations:

$$\overline{B}(j, \ell) = \frac{1}{N_1} \sum_{i=1}^{N} \sum_{t=L+1}^{T-F} D_{it} B_{it}(j, \ell)$$

Standard error calculation \(\Rightarrow\) consider weight as a covariate

1. Block bootstrap
2. Model-based cluster robust standard error within the GMM framework
Empirical Application (1)

- ATT with $L = 4$ and $F = 1, 2, 3, 4$
- We consider democratization and authoritarian reversal
- Examine the number of matched control units
- 18 (13) treated observations have no matched control
Improved Covariate Balance

- Standardized Mean Differences for Democratization
- Standardized Mean Differences for Authoritarian Reversal
- Standardized Mean Differences for Starting War

Before Matching
Before Refinement
Mahalanobis Distance Matching
Propensity Score Matching
Propensity Score Weighting

Imai, Kim, and Wang (HU/MIT/PU)
Estimated Causal Effects

Mahalanobis Matching
Up to 5 matches
Up to 10 matches

Propensity Score Matching
Up to 5 matches
Up to 10 matches

Propensity Score Weighting

Estimated Effect of Democratization

Estimated Effect of Authoritarian Reversal

Years relative to the administration of treatment

Imai, Kim, and Wang (HU/MIT/PU)
Empirical Application (2)

Starting War

Ending War

Standardized Mean Differences for Starting War

Standardized Mean Differences for Ending War

Years relative to the administration of treatment

Number of matched control units

Frequency

Acemoglu et al. (2018)
Scheve & Stasavage (2012)

Imai, Kim, and Wang (HU/MIT/PU)

Matching for TSCS Data

LatamPolMeth (Nov 9, 2018)
Estimated Causal Effects

Mahalanobis Matching

Up to 1 matches

Treatment reversal allowed
Treatment reversal not allowed

One Year Lag

Estimated effect of war

-8 -4 0 4 8 12 16
0 1 2 3 4

Propensity Score Matching

Up to 1 matches

Up to 3 matches

Propensity Score Weighting

Up to 1 matches

Up to 3 matches

Y ears relative to the administration of treatment

Years relative to the administration of treatment

Imai, Kim, and Wang (HU/MIT/PU)
Concluding Remarks

- Matching as transparent and simple methods for causal inference
- Yet, matching has not been applied to time-series cross-sectional data

- We propose a matching framework for TSCS data
  1. construct matched sets
  2. refine matched sets
  3. compute difference-in-differences estimator

- Checking covariates and computing standard errors
- R package PanelMatch implements all of these methods

- Future research: addressing possible spillover effects