Causal Inference from Observational Data using Propensity Scores

Ehsan Karim

Department of Statistics
University of British Columbia

March 17, 2011
Outline

- Background
- Motivating Example
- Propensity Score
- Software and References

Causal Inference using Propensity Scores
### Background > Individual treatment effect

#### One Individual $i$

<table>
<thead>
<tr>
<th></th>
<th>our world</th>
<th>parallel world</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>treatment 1</strong></td>
<td><img src="image1.png" alt="Image of treatment 1" /></td>
<td><img src="image2.png" alt="Image of treatment 0" /></td>
</tr>
<tr>
<td><strong>outcome due to</strong></td>
<td><img src="image3.png" alt="Image of positive outcome" /></td>
<td><img src="image4.png" alt="Image of negative outcome" /></td>
</tr>
<tr>
<td><strong>treatment 1</strong>, $Y(1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>treatment 0</strong>, $Y(0)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Treatment effect</strong> $\tau_{i,TE} = Y_i(1) - Y_i(0)$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Background > Counterfactual

**Table:** Counterfactual Items for average treatment effect under stable unit treatment value assumption

<table>
<thead>
<tr>
<th>Population</th>
<th>Population 1</th>
<th>Population 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Outcome</td>
<td>Y(1)</td>
<td>Y'(1)</td>
</tr>
<tr>
<td></td>
<td>Y(0)</td>
<td>Y'(0)</td>
</tr>
</tbody>
</table>

\[ \tau_{ATE} = E\{Y(1) - Y(0)\} \] in the population 1.

Similarly, \[ \tau_{ATE} = E\{Y'(1) - Y'(0)\} \] in the population 2.
Table: Observed Items for the average treatment effect under stable unit treatment value assumption

<table>
<thead>
<tr>
<th>Population</th>
<th>Population 1</th>
<th>Population 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>1</td>
<td>×</td>
</tr>
<tr>
<td>Outcome</td>
<td>Y(1)</td>
<td>×</td>
</tr>
</tbody>
</table>

\[ \tau'_{ATE} = E\{Y(1) - Y'(0)\} \] across populations.

If \( Y(0) \neq Y'(0) \), then confounding exists.
Figure: Simple Causal relationship and Confounding
Background > Randomization

- Claim: Ability to discriminate: whether the milk was added first or the tea was added first

- Common practice: Control for all covariates or confounders.
- Fisher’s proposal: Control for nothing! Randomization.
### Background > Situation for Randomized

From Duncan et al. (2003) “Randomized Clinical Trial of Therapeutic Exercise in Subacute Stroke”

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Intervention Group (n=44)</th>
<th>Usual Care Group (n=48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>68.5 (9.0)</td>
<td>70.2 (11.4)</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>23 (52.3)</td>
<td>27 (56.3)</td>
</tr>
<tr>
<td>Race (white), n (%)</td>
<td>37 (84.1)</td>
<td>37 (77.1)</td>
</tr>
<tr>
<td>Insurance coverage, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>67</td>
<td>61</td>
</tr>
<tr>
<td>Private</td>
<td>79</td>
<td>90</td>
</tr>
<tr>
<td>Medicaid</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary outcomes at baseline</th>
<th>Intervention Group (n=44)</th>
<th>Usual Care Group (n=48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankle isometric dorsiflexion, Nm</td>
<td>19.5 (10.0)</td>
<td>20.1 (10.0)</td>
</tr>
<tr>
<td>Knee isometric extension, Nm</td>
<td>59.4 (27.4)</td>
<td>54.8 (22.0)</td>
</tr>
<tr>
<td>Lower-extremity Fugl-Meyer score</td>
<td>24.1 (3.7)</td>
<td>23.7 (3.5)</td>
</tr>
<tr>
<td>Grip strength, kg</td>
<td>15.4 (10.2)</td>
<td>13.4 (9.5)</td>
</tr>
<tr>
<td>Berg Balance Score</td>
<td>42.8 (7.2)</td>
<td>43.1 (9.0)</td>
</tr>
<tr>
<td>Functional reach, cm</td>
<td>25.3 (6.6)</td>
<td>22.3 (6.1)</td>
</tr>
<tr>
<td>Duration of bike exercise, min</td>
<td>5.1 (2.8)</td>
<td>5.2 (2.8)</td>
</tr>
<tr>
<td>Peak $\dot{V}O_2$, mL · kg⁻¹ · min⁻¹</td>
<td>11.7 (3.3)</td>
<td>11.2 (2.9)</td>
</tr>
<tr>
<td>10-m gait velocity, m/s</td>
<td>0.7 (0.3)</td>
<td>0.6 (0.3)</td>
</tr>
<tr>
<td>6-min walk distance, m</td>
<td>238.0 (103.9)</td>
<td>215.6 (94.8)</td>
</tr>
<tr>
<td>Wolf Motor Function time, s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Including all subjects*</td>
<td>13.8 (21.9)</td>
<td>12.0 (18.8)</td>
</tr>
<tr>
<td>Excluding those unable to do tasks</td>
<td>3.5 (1.9)</td>
<td>4.5 (2.9)</td>
</tr>
<tr>
<td>Upper-extremity Fugl-Meyer score</td>
<td>45.8 (12.8)</td>
<td>43.3 (11.9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comorbid conditions at baseline, n (%)</th>
<th>Intervention Group (n=44)</th>
<th>Usual Care Group (n=48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>History of angina</td>
<td>5 (11)</td>
<td>3 (6.2)</td>
</tr>
<tr>
<td>History of heart attack</td>
<td>2 (4.5)</td>
<td>5 (10)</td>
</tr>
<tr>
<td>Irregular heart rhythm</td>
<td>4 (9.1)</td>
<td>6 (12)</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>30 (69)</td>
<td>37 (77)</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>4 (9.1)</td>
<td>5 (10)</td>
</tr>
<tr>
<td>Lung disease</td>
<td>5 (11)</td>
<td>3 (6.2)</td>
</tr>
<tr>
<td>Arthritis</td>
<td>19 (43)</td>
<td>22 (46)</td>
</tr>
<tr>
<td>Chronic pain</td>
<td>14 (32)</td>
<td>17 (35)</td>
</tr>
<tr>
<td>Depression or anxiety</td>
<td>12 (27)</td>
<td>12 (25)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>11 (25)</td>
<td>13 (27)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stroke characteristics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Olney Prognostic Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right hemisphere</td>
<td>22 (50.0)</td>
<td>22 (45.8)</td>
</tr>
<tr>
<td>Left hemisphere</td>
<td>18 (40.9)</td>
<td>22 (45.8)</td>
</tr>
<tr>
<td>Brain stem/other</td>
<td>4 (9.1)</td>
<td>4 (8.3)</td>
</tr>
<tr>
<td>Stroke type, ischemic</td>
<td>39 (88.6)</td>
<td>44 (91.7)</td>
</tr>
<tr>
<td>Time since stroke, d</td>
<td>77.5 (28.7)</td>
<td>73.5 (27.1)</td>
</tr>
</tbody>
</table>

### Causal Inference using Propensity Scores
Background > Solutions for Non-randomized data

To bring covariate balance,

- Popular solution is matching
- subclassification
- Regression adjustment as covariate
- Regression adjustment by weighting
The data set we will use to illustrate is a subset of the job training program analyzed in Lalonde (1986). We deal with a subsample of the original data consisting of

treated group the National Supported Work Demonstration (NSW) and

comparison sample the Population Survey of Income Dynamics (PSID).
The variables in this dataset include

**Outcome:** re78 1978 real earnings

**Treatment Indicator:** treat participation in the job training program from March 1975-June 1977 (= 1 if participated, 0 otherwise)

**Pre-treatment Covariates:**
- educ years of education
- age age
- race It has three categories white as baseline
  - black (= 1 if black, and 0 otherwise)
  - hispan (= 1 if hispanic, and 0 otherwise)
- married marital status (= 1 if married, 0 otherwise)
- nodegree high school degree (= 1 if no degree, 0 otherwise)
- re74 1974 real earnings
- re75 1975 real earnings
### Example > Data View

10 variables and 614 observations (slightly modified)

<table>
<thead>
<tr>
<th></th>
<th>treat</th>
<th>age</th>
<th>educ</th>
<th>black</th>
<th>hispan</th>
<th>married</th>
<th>nodegree</th>
<th>re74</th>
<th>re75</th>
<th>re78</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW1</td>
<td>1</td>
<td>37</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5345</td>
<td>4665</td>
<td>9930.05</td>
</tr>
<tr>
<td>NSW2</td>
<td>1</td>
<td>22</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>354</td>
<td>3255</td>
<td>3595.89</td>
</tr>
<tr>
<td>NSW3</td>
<td>1</td>
<td>30</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>67</td>
<td>0</td>
<td>24909.45</td>
</tr>
<tr>
<td>NSW4</td>
<td>1</td>
<td>27</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7506.15</td>
</tr>
<tr>
<td>NSW5</td>
<td>1</td>
<td>33</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>25</td>
<td>289.79</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSID424</td>
<td>0</td>
<td>25</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>235</td>
<td>643</td>
<td>0.00</td>
</tr>
<tr>
<td>PSID425</td>
<td>0</td>
<td>18</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>44</td>
<td>10150.50</td>
</tr>
<tr>
<td>PSID426</td>
<td>0</td>
<td>24</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8457</td>
<td>346</td>
<td>19464.61</td>
</tr>
<tr>
<td>PSID427</td>
<td>0</td>
<td>21</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>676</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>PSID428</td>
<td>0</td>
<td>32</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>325</td>
<td>0</td>
<td>187.67</td>
</tr>
<tr>
<td>PSID429</td>
<td>0</td>
<td>16</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>897</td>
<td>453</td>
<td>1495.46</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For data lalonde in package MatchIt, we obtain 13 Exact Subclasses having following $2 \times 2$ table:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>429</td>
<td>185</td>
</tr>
<tr>
<td>Matched</td>
<td>413</td>
<td>185</td>
</tr>
</tbody>
</table>

for the formula

matchit(treat ~ educ,
        data = lalonde, method = "exact")
Example > Matching

For data lalonde in package MatchIt, we obtain 20 Exact Subclasses having following $2 \times 2$ table:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>429</td>
<td>185</td>
</tr>
<tr>
<td>Matched</td>
<td>381</td>
<td>184</td>
</tr>
</tbody>
</table>

for the formula

```
machit(treat ~ educ + black, data = lalonde, method = "exact")
```
Example > Matching

For data `lalonde` in package `MatchIt`, we obtain 25 Exact Subclasses having following $2 \times 2$ table:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>429</td>
<td>185</td>
</tr>
<tr>
<td>Matched</td>
<td>376</td>
<td>184</td>
</tr>
</tbody>
</table>

for the formula

```
matchit(treat ~ educ + black + hispan, data = lalonde, method = "exact")
```
For data lalonde in package MatchIt, we obtain 35 Exact Subclasses having following $2 \times 2$ table:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>429</td>
<td>185</td>
</tr>
<tr>
<td>Matched</td>
<td>284</td>
<td>183</td>
</tr>
</tbody>
</table>

for the formula

```
matchit(treat ~ educ + black + hispan + married, 
        data = lalonde, method = "exact")
```
Example > Matching

For data lalonde in package MatchIt, we obtain 31 Exact Subclasses having following $2 \times 2$ table:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>429</td>
<td>185</td>
</tr>
<tr>
<td>Matched</td>
<td>59</td>
<td>54</td>
</tr>
</tbody>
</table>

for the formula

```
matchit(treat ~ educ+black+hispan+married+age, data = lalonde, method = "exact")
```
For the subclass 3 and 19, the following data are obtained:

<table>
<thead>
<tr>
<th>treat</th>
<th>age</th>
<th>educ</th>
<th>black</th>
<th>hispan</th>
<th>married</th>
<th>nodegree</th>
<th>re78</th>
<th>subclass</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSW11</td>
<td>1</td>
<td>19</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>8173.91</td>
<td>3</td>
</tr>
<tr>
<td>NSW123</td>
<td>1</td>
<td>19</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>17685.18</td>
<td>3</td>
</tr>
<tr>
<td>PSID140</td>
<td>0</td>
<td>19</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>14344.29</td>
<td>3</td>
</tr>
<tr>
<td>PSID355</td>
<td>0</td>
<td>19</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.00</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW89</td>
<td>1</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8048.60</td>
<td>19</td>
</tr>
<tr>
<td>NSW116</td>
<td>1</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12558.02</td>
<td>19</td>
</tr>
<tr>
<td>PSID103</td>
<td>0</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1687.56</td>
<td>19</td>
</tr>
<tr>
<td>PSID105</td>
<td>0</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5937.51</td>
<td>19</td>
</tr>
<tr>
<td>PSID137</td>
<td>0</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>19</td>
</tr>
<tr>
<td>PSID153</td>
<td>0</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4787.83</td>
<td>19</td>
</tr>
</tbody>
</table>
Rosenbaum and Rubin (1983) proposed a balancing score that will solve the curse of dimensionality.

Two subjects with same balancing score would mean they have similar covariate distribution, i.e., they are comparable.

One possible balancing score is “propensity score”:
- $P(\text{participating a program } W \mid \text{observed covariates } X)$
- $P(\text{receiving treatment } W \mid \text{observed covariates } X)$
## Development of Propensity Score

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Economics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counterfactual framework</td>
<td>Two step regression</td>
</tr>
<tr>
<td>Randomized Experiments</td>
<td>Structural equation model</td>
</tr>
<tr>
<td>Randomization &amp; Probability</td>
<td>sample selection</td>
</tr>
</tbody>
</table>
Propensity Score > Estimation

Following tools can be used to estimate propensity scores:

1. logistic link
2. probit link
3. complementary log-log link
4. Cauchy CDF link
5. boosted logistic regression
Here we set treat as dependent (W) and age, educ, black, hispan, nodegree, married, re74, re75 and their interactions and polynomials as covariates (X).
Once we estimate propensity score, then we need to match them:

1. Greedy matching
   - Mahalanobis metric including Propensity score
   - Nearest Neighborhood
   - Caliper
   - Nearest Neighborhood within Caliper
   - Nearest Mahalanobis metric within Caliper

2. Optimal matching (ex: .4, .5, .1, .9)
Propensity Score (Nearest Neighborhood algorithm) Matching (logit based)\(^1\) provides following\(^2\)

<table>
<thead>
<tr>
<th></th>
<th>treat</th>
<th>educ</th>
<th>black</th>
<th>hispan</th>
<th>age</th>
<th>re74</th>
<th>re75</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW35</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0.70933</td>
</tr>
<tr>
<td>NSW54</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0.70933</td>
</tr>
<tr>
<td>NSW74</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0.70933</td>
</tr>
<tr>
<td>NSW81</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0.70933</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW7</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0.70950</td>
</tr>
<tr>
<td>NSW38</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0.70950</td>
</tr>
<tr>
<td>NSW50</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0.70950</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)based on treat, educ, black, hispan, age, re74, re75

\(^2\)370 observations matched out of 614 and discards all units (treated and control) that are outside the support of the distance measure.
Propensity Score (Nearest Neighbor algorithm) Matching (probit based)$^3$ provides following$^4$

<table>
<thead>
<tr>
<th></th>
<th>treat</th>
<th>educ</th>
<th>black</th>
<th>hispan</th>
<th>age</th>
<th>re74</th>
<th>re75</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW25</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0.67778</td>
</tr>
<tr>
<td>PSID367</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0.67778</td>
</tr>
<tr>
<td>PSID411</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0.67778</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW29</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.67797</td>
</tr>
<tr>
<td>NSW31</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.67797</td>
</tr>
<tr>
<td>NSW56</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.67797</td>
</tr>
<tr>
<td>NSW92</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.67797</td>
</tr>
<tr>
<td>PSID392</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.67797</td>
</tr>
<tr>
<td>PSID401</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.67797</td>
</tr>
</tbody>
</table>

$^3$Based on treat, educ, black, hispan, age, re74, re75
$^4$370 observations matched out of 614 and discards all units (treated and control) that are outside the support of the distance measure.
Propensity Score \> Balance Checking

Summary of balance for matched data:

<table>
<thead>
<tr>
<th></th>
<th>Means Treated</th>
<th>Means Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>educ</td>
<td>10.35</td>
<td>10.45</td>
</tr>
<tr>
<td>black</td>
<td>0.84</td>
<td>0.46</td>
</tr>
<tr>
<td>hispan</td>
<td>0.06</td>
<td>0.26</td>
</tr>
<tr>
<td>age</td>
<td>25.82</td>
<td>26.48</td>
</tr>
<tr>
<td>re74</td>
<td>2095.57</td>
<td>2300.94</td>
</tr>
<tr>
<td>re75</td>
<td>1532.06</td>
<td>1808.56</td>
</tr>
</tbody>
</table>
**Figure:** Matching Balance Checking using Graph: Propensity score generated from logit – with educ, age, re74 as covariates.
**Figure:** The overall histograms of propensity scores in the treated and control groups.
Figure: The jitter plot shows the overall distribution of propensity scores in the treated and control groups.
**Figure:** The jitter plot shows the overall distribution of propensity scores in the treated and control groups (ratio being 2).
If **Balance** is not achieved, we need to go back to propensity score equation and modify it until satisfactory balance achieved.

Such model can not account for hidden bias: need to go through sensitivity analysis.

Once balance is achieved, the treatment effect can be estimated as

\[
\hat{\tau}_M = E(\hat{y}_{M,1}|w_M = 1) - E(\hat{y}_{M,0}|w_M = 0),
\]

where $M$ indicates matched sample, and $M, 1$ and $M, 0$ are sub-samples of it in treated and untreated group.
**Causal Inference using Propensity Scores**

**Software and Reference**

1. **R**
   - *MatchIt*: estimation, matching methods and graphics
   - *Matching*: automated, based on balance diagnostics
   - *Twang*: uses Generalized boosted regression
   - *rbounds*: Implements sensitivity analysis
   - *optmatch*: Variable ratio, optimal, and full matching
   - *PSAgraphics*: graphics

2. **Stata**
   - *psmatch2*: k:1, kernel, diagnostics, \( \hat{\tau} \)
   - *pscore*: k:1, caliper matching, and stratification
   - *match*: Primarily k:1 matching (with replacement)
   - *rbounds*: Implements sensitivity analysis

3. **SAS**
   - Greedy matching, Weighting
   - 1:1 Mahalanbois matching within propensity score calipers
   - Variable ratio matching, optimal matching algorithm

4. **SPSS**: Nearest neighbor matching

(Details in [http://www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html](http://www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html))
Causal Inference using Propensity Scores
Thank You!