Data-driven sensitivity analysis for Matching estimators

Giovanni Cerulli ¹

¹IRCrES-CNR, Research Institute on Sustainable Economic Growth

London Stata Conference 2018
Cass Business School
September 6-7
Summary

- Motivation and objective
- Current approaches
- The LOCO approach
- Stata implementation via sensimatch
- Application
- Conclusion
Motivation and objective

- Under "unobservable selection" Matching is an inconsistent estimator of the ATET

- Unobservables are context-dependent (genuine and/or contingent unobservables)

- Alternative methods: instrumental-variables (IV), selection models (SM), and quasi-natural approaches (regression discontinuity design, RD), Diff-in-diffs

- Costly alternatives require extra information and assumptions, rarely available, not accessible, often unreliable

- Sensitivity analysis helps to detect whether Matching is robust to unobservable selection
Motivation and objective

This paper:

- proposes a (novel) sensitivity analysis for **unobservable selection** in Matching estimation based on a “leave-one-covariate-out” (**LOCO**) approach

- rooted in the **Machine Learning** literature

- based on a bootstrap over different **subsets** of covariates

- simulates **estimation scenarios** and compares them with the baseline Matching estimated by the analyst

- introduces **sensimatch**, a Stata routine I developed to run this method

- provides an instructional application on real data
Io intendo scultura, quella che si fa per forza di **levare**: quella che si fa per via di **porre**, è simile alla pittura

(*I mean sculpture, the one that one does by force of removing: what one does by posing, is similar to painting*)

**Michelangelo Buonarroti**

*“Letter to Sir Benedetto Varchi”*

*Florence, XVI Century*
Sensitivity analysis: the study of how the uncertainty in the output of a model or system can be explained by different sources of uncertainty in its inputs
Sensitivity approaches in the Matching literature

Two Matching sensitivity tests for the possible presence of unobservable selection:

- The Rosenbaum (1987) test \(\Rightarrow\) based on the Wilcoxon’s signed rank statistic

- The Ichino, Mealli, and Nannicini (IMN, 2008) test \(\Rightarrow\) based simulating the (possible) presence of unobeservable
Rosenbaum approach

- Assume perfect randomization (as restored after Matching)

- Define $\Gamma = \text{"PS ratio between treated and untreated"}$ $\Rightarrow$ same odds under randomization

- Perturbate randomization by increasing $\Gamma \Rightarrow$ larger departure from randomization

- Look at what $\Gamma$ the effect (ATET) is no longer significant (result overturning)

- A high level of critical $\Gamma$ is a signal of Matching robustness
IMN approach

- Consider the baseline Matching estimates

- Define $d$ and $s$ as two probability ratios increasing with unobservable selection: 1. $d$: UCs effect on the outcome; 2. $s$: UCs effect on the treatment

- As soon as both $d$ and $s$ increase, ATET goes to zero

- Tabulate increasing values of $d$ and $s$ until ATET is no longer significant.

- A high level of critical $d$ and $s$ is a signal of Matching robustness
The logic of LOCO

- Previous methods follow a **posing** logic ⇒ what happens when one perturbates the baseline model by adding up UCs

- LOCO follows a different but **specular** logic: “if the baseline model results are poorly (strongly) sensitive to adding up UCs, it is likely to be poorly (strongly) sensitive to removing them”

- We can obtain a specular result by **removing**, instead of **posing**
The LOCO algorithm

1. Start from running a Matching model using $x = \{x_1, x_2, \ldots, x_K\}$ observable confounders, thus estimating one single ATET, and take this as the baseline estimate.

2. Starting from the $K$ observables, select a subset size $S$ with $S = 1, 2, \ldots, j, \ldots, M$, and $M < K$.

3. Draw $H$ times at random and without replacement a set of covariates of size $S$ from the original set of observables $x$.

4. Run $H$ Matching models of size $S$ thus obtaining a number of $H$ ATET point estimates, standard errors, and confidence intervals.

5. For each size $S$, average the obtained estimates over $H$, and check whether the results are sensibly changed by reducing $S$ from $K - 1$ to 1.
The Stata module `sensimatch`

**Title**

`sensimatch` – Data-driven sensitivity analysis to assess Matching robustness to unobservable selection

**Syntax**

```plaintext
sensimatch outcome treatment [varlist] ,
   sims(#) mod(modeltype) seed(#) fac(varlist_f)
   vce(vcetype) graph_options(options)
```

*modeltype*

- `reg`: Ordinary Least Squares
- `match`: Nearest-neighbour propensity-score Matching
Application on real data

- **Dataset**: National Longitudinal Survey of Mature and Young Women (NLSW) in 1988

- **Objective**: Detecting the effect of “unionization” on hourly “wage” on 2,246 American women

- **Confounders**: age: age of the woman; race: race of the woman (white, black, other); married: married vs. non-married; never_married: whether or not never married; grade: grade obtained at school final exam; south: whether of not the woman comes from the South; smsa: whether she lives in SMSA; city: whether of not she lives in central city; collgrad: whether she is college graduated; hours: usual hours worked; ttl_exp: total work experience; tenure: job tenure in years; industry: type of industry; occupation: type of occupation.
Baseline propensity-score Matching results - psmatch2

use nlsw88, clear

global y "wage"
global w "union"
global xvars age race married never_married grade south smsa c_city collgrad hours ttl_exp tenure
global factors "industry occupation"

xi: psmatch2 $w $xvars i.industry i.occupation , out($y) common

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>C</th>
<th>Diff</th>
<th>S.E.</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIM</td>
<td>8.67</td>
<td>7.25</td>
<td>1.44</td>
<td>.22</td>
<td>6.44</td>
</tr>
<tr>
<td>ATET</td>
<td>8.67</td>
<td>7.65</td>
<td>1.02</td>
<td>.37</td>
<td>2.76</td>
</tr>
</tbody>
</table>
Using `rbounds`

```stata
.xi: psmatch2 $w $xvars i.ind i.occ , out($y) common
.gen delta = $y - _wage if _treated==1 & _support==1
.rbounds delta, gamma(1 (0.01) 2)
```
### Sensitivity analysis for Matching

**Motivation and objective**

**Current approaches**

**The LOCO approach**

**The Stata module sensimatch**

**Application**

**Conclusion**

Rosenbaum sensitivity analysis - rbounds - #2

<table>
<thead>
<tr>
<th>Gamma</th>
<th>sig+</th>
<th>sig-</th>
<th>t-hat+</th>
<th>t-hat-</th>
<th>CI+</th>
<th>CI-</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.6e-06</td>
<td>2.6e-06</td>
<td>1.08293</td>
<td>1.08293</td>
<td>.619968</td>
<td>1.53784</td>
</tr>
<tr>
<td>1.01</td>
<td>4.0e-06</td>
<td>1.7e-06</td>
<td>1.05878</td>
<td>1.10306</td>
<td>.595817</td>
<td>1.55797</td>
</tr>
<tr>
<td>1.02</td>
<td>6.1e-06</td>
<td>1.1e-06</td>
<td>1.03772</td>
<td>1.12319</td>
<td>.575685</td>
<td>1.58212</td>
</tr>
<tr>
<td>1.03</td>
<td>9.2e-06</td>
<td>6.9e-07</td>
<td>1.0145</td>
<td>1.14331</td>
<td>.556793</td>
<td>1.60628</td>
</tr>
<tr>
<td>1.04</td>
<td>6.1e-06</td>
<td>4.4e-07</td>
<td>.994364</td>
<td>1.16345</td>
<td>.539451</td>
<td>1.62641</td>
</tr>
<tr>
<td>1.05</td>
<td>9.2e-06</td>
<td>2.8e-07</td>
<td>.974235</td>
<td>1.1876</td>
<td>.515301</td>
<td>1.64654</td>
</tr>
<tr>
<td>1.06</td>
<td>1.00029</td>
<td>1.8e-07</td>
<td>.954105</td>
<td>1.2037</td>
<td>.495169</td>
<td>1.66667</td>
</tr>
<tr>
<td>1.07</td>
<td>1.00042</td>
<td>4.1e-07</td>
<td>.933976</td>
<td>1.22474</td>
<td>.47504</td>
<td>1.6868</td>
</tr>
<tr>
<td>1.08</td>
<td>1.00059</td>
<td>6.9e-08</td>
<td>.913847</td>
<td>1.24798</td>
<td>.458934</td>
<td>1.70692</td>
</tr>
<tr>
<td>1.09</td>
<td>1.00083</td>
<td>4.3e-08</td>
<td>.893721</td>
<td>1.26811</td>
<td>.434783</td>
<td>1.72705</td>
</tr>
<tr>
<td>1.1</td>
<td>1.00116</td>
<td>2.7e-08</td>
<td>.873592</td>
<td>1.28422</td>
<td>.414655</td>
<td>1.74641</td>
</tr>
<tr>
<td>1.11</td>
<td>1.00159</td>
<td>1.7e-08</td>
<td>.857484</td>
<td>1.30435</td>
<td>.394527</td>
<td>1.76731</td>
</tr>
<tr>
<td>1.12</td>
<td>1.00218</td>
<td>1.0e-08</td>
<td>.837229</td>
<td>1.32448</td>
<td>.378421</td>
<td>1.78342</td>
</tr>
<tr>
<td>1.13</td>
<td>1.00294</td>
<td>6.4e-09</td>
<td>.817228</td>
<td>1.34213</td>
<td>.358293</td>
<td>1.80354</td>
</tr>
<tr>
<td>1.14</td>
<td>1.00394</td>
<td>3.9e-09</td>
<td>.797103</td>
<td>1.36071</td>
<td>.334139</td>
<td>1.81965</td>
</tr>
<tr>
<td>1.15</td>
<td>1.00523</td>
<td>2.4e-09</td>
<td>.776974</td>
<td>1.38083</td>
<td>.314009</td>
<td>1.83978</td>
</tr>
</tbody>
</table>
### Sensitivity analysis for Matching

#### Motivation and objective

#### Current approaches

#### The LOCO approach

#### The Stata module sensimatch

#### Application

#### Conclusion

---

**Rosenbaum sensitivity analysis - rbounds - #3**

<table>
<thead>
<tr>
<th>1.35</th>
<th>0.033501</th>
<th>7.9e-14</th>
<th>438807</th>
<th>1.72593</th>
<th>-0.036234</th>
<th>2.18196</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.36</td>
<td>0.038743</td>
<td>4.6e-14</td>
<td>421621</td>
<td>1.73913</td>
<td>-0.052334</td>
<td>2.19659</td>
</tr>
<tr>
<td>1.37</td>
<td>0.044587</td>
<td>2.7e-14</td>
<td>406602</td>
<td>1.75523</td>
<td>-0.068438</td>
<td>2.21417</td>
</tr>
</tbody>
</table>

| 1.38 | 0.051068 | 1.6e-14 | 3905   | 1.77523 | -0.08454  | 2.23027 |

<table>
<thead>
<tr>
<th>1.39</th>
<th>0.058221</th>
<th>9.0e-15</th>
<th>378419</th>
<th>1.78744</th>
<th>-0.10064</th>
<th>2.24235</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4</td>
<td>0.066076</td>
<td>5.2e-15</td>
<td>362316</td>
<td>1.79952</td>
<td>-0.116748</td>
<td>2.25845</td>
</tr>
<tr>
<td>1.41</td>
<td>0.074661</td>
<td>3.0e-15</td>
<td>342191</td>
<td>1.81562</td>
<td>-0.132852</td>
<td>2.27455</td>
</tr>
<tr>
<td>1.42</td>
<td>0.083999</td>
<td>1.8e-15</td>
<td>326085</td>
<td>1.83172</td>
<td>-0.152974</td>
<td>2.29054</td>
</tr>
<tr>
<td>1.43</td>
<td>0.094111</td>
<td>1.0e-15</td>
<td>309982</td>
<td>1.84523</td>
<td>-0.165056</td>
<td>2.30274</td>
</tr>
<tr>
<td>1.44</td>
<td>0.105012</td>
<td>5.6e-16</td>
<td>293881</td>
<td>1.8599</td>
<td>-0.17992</td>
<td>2.31884</td>
</tr>
</tbody>
</table>

---

**Unlikely circumstance ⇒ Matching robust to unobservable selection**
Using \texttt{sensimatch}

\begin{verbatim}
sensimatch \$y \$w \$xvars , mod(match) sims(50) ///
vce(robust) fac($factors) seed(1010)
\end{verbatim}
Sensitivity analysis for CMI on ATET

Number of simulations: 50
Reference ATET: 1.03
Model: Propensity-score Matching
Number of baseline covariates: 37
Dependent variable: Wage
Sensitivity analysis for CMI on T-Student

Number of simulations: 50
Reference T-student: 2.76
Model: Propensity-score Matching
Number of baseline covariates: 37
Dependent variable: Wage
As a possible measure of sensitivity to unobservable selection one can consider, for instance, “the ratio between the number of not removed covariates leading to lose $\alpha$–significance and the number of the baseline covariates”:

$$\rho_\alpha = \frac{S_{\text{critical},\alpha}}{K}$$

As long as $\rho_\alpha$ increases, Matching sensitivity to unobservable selection increases accordingly.
In our previous example we have that:

\[
\rho_1 = \frac{12}{37} = 0.33
\]

\[
\rho_1 = \frac{9}{37} = 0.24
\]

\[
\rho_1 = \frac{7}{37} = 0.18
\]

One can pre-fix a given **threshold** for the accepted level of uncertainty as, for example, a \( \rho \) not larger than 90%. A value of \( \rho \) larger than 90 may signal a **severe** sensitivity of Matching to unobservable selection.
The LOCO approach seems to lead to results **consistent** with those from the Rosenbaum approach.

- It has the advantage to be totally **data-driven** \(\rightarrow\) it is **model-free**.

- It can be **generalized** to whatever causal parameter and methods (for instance the IPW).

- It has the disadvantage to be **computationally intensive** and thus slower to provide results.
Many thanks !!!

See you next year for the London Stata Conference 2019 !