Implementing ODA from Within Stata: A Reanalysis of the National Supported Work Experiment

Ariel Linden, Dr.P.H. and Paul R. Yarnold, Ph.D.
Linden Consulting Group, LLC
Optimal Data Analysis LLC

Data from the National Supported Work (NSW) randomized experiment have been used frequently over the past 30 years to demonstrate implementation of various non-experimental methods for drawing causal inferences about treatment effects. In this paper we reanalyze these data using the new Stata package for implementing ODA.

Studies in which participants are randomized to treatment are considered the gold standard for assessing causal inference because randomization putatively ensures that the study groups do not differ systematically in their characteristics, and consequently, treatment effects are assumed to be unbiased.1 If randomization is infeasible, investigators rely on statistical techniques which model treatment assignment in order to control for threats to validity which may compromise causal interpretation of the results.2-6

In this paper we reanalyze data from the National Supported Work (NSW) experiment which was originally discussed by LaLonde7 in the context of economic evaluation approaches, but has since then been utilized frequently to demonstrate the implementation of various non-experimental techniques, such as propensity scoring methods, for assessing causal inference.

Herein we apply the new Stata package called oda,8 that implements ODA from within the Stata environment, to these data to assess whether results are consistent with findings reported by Dehejia and Wahba.9

The oda package is a wrapper for the MegaODA software10, and the megaODA.exe file must be loaded on the computer for the oda package to work (ODA software is available at https://odajournal.com/resources/). To download the oda package, at the Stata command line type: “ssc install oda” without quotation marks.

Methods

Data

The NSW was a US federally- and privately funded program that aimed to provide work experience for individuals who had faced economic and social problems prior to enrollment in the program. Candidates for the experiment were selected on the basis of eligibility criteria, and then were either randomly assigned to, or
excluded from the training program. We use the same subset of NSW data used by Dehejia and Wahba\(^9\), joining the 185 treated units from the NSW experiment to comparison units from the 15,992 individuals in the Current Population Survey (CPS). Data were retrieved from:

http://users.nber.org/~rdehejia/nswdata2.html

Available variables included were age, education, black, hispanic, no degree, married, real earnings in 1974, 1975 and 1978 (all adjusted to 1982 US dollars), and indicators for unemployed status in 1974 and 1975. The outcome (primary model attribute) was real earnings in 1978, and the treatment (class) variable indicates whether individuals participated in the NSW intervention or were untreated from the CPS data.

**Analysis**

Dehejia and Wahba\(^9\) estimated a propensity score in which the binary treatment indicator was regressed on age, age\(^2\), age\(^3\), education, education\(^2\), married, no degree, black, Hispanic, earnings in 1974 and 1975, unemployed in 1974 and 1975, and an interaction of education and earnings in 1974. They then used various matching algorithms and compared the outcomes across the methods. Here we replicate their propensity score estimation and the 1:1 matching without replacement method.

* Estimate the propensity score
  
  \[
  \text{logit treat age c.age#c.age} \\
  \text{c.age#c.age#c.age educ c.educ#c.educ} \\
  \text{married nодеgree black hispan re74 re75 u74 u75 c.educ#c.re74}
  \]

* save predictions
  
  \text{predict pscore}

* use psmatch2 to perform 1:1 matching
  
  \text{psmatch2 treat, outcome(re78)}
  \text{pscore(pscore) neighbor(1) noreplace}

* Evaluate treatment effects using regression with robust standard errors and frequency weights from the psmatch2 output
  
  \[
  \text{regress re78 treat [fw = _weight], vce(robust)}
  \]

The regression produced a non-statistically significant estimated treatment effect of $1282 (95\% \text{ CI: } -204, 2768, P = 0.091).

We now evaluate these data using oda with the following syntax (see the help file for oda for a complete description of the syntax options):

\[
\text{oda treat re78 if _weight != ., pathoda("C:\ODA\") store("C:\ODA\output") iter(10000) loo, seed(1234)}
\]

The above syntax is explained as follows: The variable “treat” is the class variable; the outcome variable “re78” (earnings in 1987) is the attribute; the [if] statement indicates that the sample should be limited to observations with a non-missing weight (i.e., matched treated and control observations), the directory path where the megaODA.exe file is located on my computer is "C:\ODA\"; the directory path where the output and other files generated during the analysis should be stored is "C:\ODA\output"; the number of iterations (repetitions) for computing a permutation P-value is 10,000; leave-one-out (LOO) analysis should be performed, and the seed should be set to 1234 to ensure replication of the permutation results.

The oda package produces an extract of the total output produced by the ODA software (the complete output is stored in the specified directory with the extension “.out”).

As shown in the oda output below, the ODA model can be interpreted as follows: "if real earnings in 1978 <= $1237.291, then predict that the treatment group is 0 (controls). If the earnings are > $1237.291, then predict that the treatment group is 1 (treatment)."
The effect strength for sensitivity (ESS) is labelled in the output as “Effect Strength PAC”. In the training data the ESS is 16.76% and in the LOO analysis it is 15.68% (a very weak effect). The permutation P-value for the training sample was 0.010 and for the LOO analysis was 0.005. In contrast to the regression which found no statistical difference between treatment groups (i.e., no treatment effect), ODA found a statistically significant treatment effect in these data, although the model found by ODA had a difficult time in discriminating between treatment groups.

ODA model:

IF RE78 < 1237.291 THEN TREAT = 0
IF 1237.291 < RE78 THEN TREAT = 1

Summary for Class TREAT Attribute RE78

<table>
<thead>
<tr>
<th>Performance Index</th>
<th>Train LOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Accuracy</td>
<td>58.38% 57.84%</td>
</tr>
<tr>
<td>PAC TREAT=0</td>
<td>49.19% 49.19%</td>
</tr>
<tr>
<td>PAC TREAT=1</td>
<td>67.57% 66.48%</td>
</tr>
<tr>
<td>Effect Strength PAC</td>
<td>16.76% 15.68%</td>
</tr>
<tr>
<td>PV TREAT=0</td>
<td>60.26% 59.48%</td>
</tr>
<tr>
<td>PV TREAT=1</td>
<td>57.08% 56.68%</td>
</tr>
<tr>
<td>Effect Strength PV</td>
<td>17.34% 16.16%</td>
</tr>
<tr>
<td>Effect Strength Total</td>
<td>17.05% 15.92%</td>
</tr>
</tbody>
</table>

Monte Carlo summary (Fisher randomization):

- Iterations: 10000
- Estimated p: 0.010400

Results of leave-one-out analysis

370 observations

Fisher’s exact test (directional) classification table p = .001530

Discussion

In this paper, we demonstrated how the new Stata package oda can be used in conjunction with a matching algorithm to evaluate treatment effects in observational data. ODA should be considered the preferred approach over commonly-used parametric models because ODA avoids the assumptions required of parametric models, is insensitive to skewed data or outliers, and has the ability to handle any variable metric including categorical, Likert-type integer, and real number measurement scales. Moreover, in contrast to regression models, ODA also has the distinct ability to ascertain where the optimal (maximum-accuracy) cutoffs are on the outcome variable, which in turn, facilitates the use of measures of predictive accuracy. Moreover, ODA can perform cross-validation using LOO which allows for assessing the cross-generalizability of the model to potentially new study participants or non-participants.

Finally, the findings continue to support our recommendation to employ the ODA and CTA frameworks to evaluate the efficacy of health-improvement interventions and policy initiatives.

References


**Author Notes**

No conflict of interest was reported.