13. Propensity Score Matching Technique

The basic idea behind propensity score matching (PSM) is to match each participant with an identical nonparticipant and then measure the average difference in the outcome variable between the participants and the nonparticipants. This exercise illustrates how to implement PSM in the Stata program.

The estimation command in Stata is "pscore.ado," developed by Becker and Ichino (2002). The "pscore" command estimates the propensity score, which is the probability of getting a treatment for each household, and tests the balancing property—that is, observations with the same propensity score must have the same distribution of observable characteristics independent of treatment status. After balancing is done, different commands can be used to carry out different types of matching and then derive the average treatment effect.

Propensity Score Equation: Satisfying the Balancing Property

The first step in PSM is to determine the propensity score and satisfy the balancing property. It is done using the "pscore" command in Stata. Use the 1998 data, hh_98.dta. Start with the male program participation variable "dmmfd" as the treatment variable. The following command shows the application of the "pscore" command:

```
pscore dmmfd sexhead agehead educhead lnland vaccess pcirr rice
wheat milk oil egg [pw=weight], pscore(ps98) blockid(blockf1)
comsup level(0.001);
```

The results include probit regression output, the estimation and description of the propensity scores, the number of blocks and stratification using propensity scores, and the balancing property test. The area of common support is those propensity scores within the range of the lowest and highest estimated values for households in the treatment group.

The following output shows that the identified region of common support is [.00180123, .50022341], the final number of blocks is 4, and the balancing property is not satisfied. The most important element to look for in the output is the list of variables that cause the balancing property not to be satisfied. The output shows the "egg" variable is not balanced in block 2. The solution to this problem is to use a different set of covariates and rerun the "pscore" command.

```
******************************************
Algorithm to estimate the propensity score
******************************************

The treatment is dmmfd

HH has male |
microcredit |
participant |
 : 1=Y, 0=N |      Freq.     Percent        Cum.
------------+----------------------------------
          0 |        909       80.51      80.51
          1 |        220       19.49     100.00
------------+----------------------------------
      Total |      1,129      100.00

Estimation of the propensity score

(sum of wgt is  1.1260e+03)
Iteration 0:  log pseudolikelihood = -424.61883
Iteration 1:  log pseudolikelihood = -390.85321
Iteration 2:  log pseudolikelihood = -389.10243
Iteration 3:  log pseudolikelihood = -389.05511
Iteration 4:  log pseudolikelihood = -389.05501

Probit estimates                           Number of obs     =    1129
                                           Wald chi2(11)     =   64.36
                                           Prob > chi2       =  0.0000
Log pseudolikelihood = -389.05501          Pseudo R2         =  0.0838
----------------------------------------------------------------------
            |             Robust
      dmmfd |     Coef.  Std. Err.     z   P>|z|  [95% Conf. Interval]
------------+---------------------------------------------------------
    sexhead |   .915108   .2432905   3.76  0.000   .4382675   1.391949
    agehead | -.0036952   .0046186  -0.80  0.424  -.0127475    .005357
   educhead |  .0161662   .0170125   0.95  0.342  -.0171777     .04951
     lnland | -.3341691   .1113146  -3.00  0.003  -.5523417  -.1159965
    vaccess | -.0752904   .1770457  -0.43  0.671  -.4222935   .2717128
      pcirr |  .2088394   .1753383   1.19  0.234  -.1348174   .5524961
       rice |   .145771   .0384417   3.79  0.000   .0704268   .2211153
      wheat |  .0465751   .0648087   0.72  0.472  -.0804475   .1735977
       milk | -.0017358    .023861  -0.07  0.942  -.0485026    .045031
        oil | -.0249797   .0135856  -1.84  0.066   -.051607   .0016476
        egg | -.7687454   .2311995  -3.33  0.001  -1.221888  -.3156028
      _cons | -1.188481   .8358266  -1.42  0.155  -2.826671   .4497088
----------------------------------------------------------------------

Note: the common support option has been selected
The region of common support is [.00180123, .50022341]

Description of the estimated propensity score
in region of common support
```

```
                 Estimated propensity score
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0055359      .0018012
 5%     .0170022      .0020871
10%     .0346036      .0026732      Obs                1127
25%      .069733      .0028227      Sum of Wgt.        1127
50%     .1206795                    Mean           .1339801
                       Largest      Std. Dev.      .0850809
75%     .1811405      .4698302
90%     .2527064       .472444      Variance       .0072388
95%     .2965199      .4735467      Skewness       .8931864
99%     .3903884      .5002234      Kurtosis       3.942122

*****************************************************
Step 1: Identification of the optimal number of blocks
Use option detail if you want more detailed output
*****************************************************


The final number of blocks is 4

This number of blocks ensures that the mean propensity score
is not different for treated and controls in each blocks

**********************************************************
Step 2: Test of balancing property of the propensity score
Use option detail if you want more detailed output
**********************************************************

Variable egg is not balanced in block 2

The balancing property is not satisfied

Try a different specification of the propensity score

           |   HH has male
  Inferior |   microcredit
  of block | participant: 1=Y, 0=N
 of pscore |      0           1 |     Total
-----------+--------------------+----------
         0 |    380          49 |       429
        .1 |    382          97 |       479
        .2 |    140          70 |       210
        .4 |      5           4 |         9
-----------+--------------------+----------
     Total |    907         220 |     1,127

Note: the common support option has been selected

*******************************************
End of the algorithm to estimate the pscore
*******************************************
```

After a few iterations, you will find that dropping “egg” and “lnland” allows the “pscore” command to be rerun with the balancing property satisfied. So “pscore” on “dfmfd” is run again, this time excluding the “egg” and “lnland” variables. Before rerunning the “pscore” command, it is important to drop the “ps98” and “blockf1” variables that were created as a result of the earlier run. Because female program participation is of more interest, the “pscore” command is shown here with female participation only.

```
pscore dfmfd sexhead agehead educhead lnland vaccess pcirr rice
wheat milk oil egg [pw=weight], pscore(ps98) blockid(blockf1)
comsup level(0.001);
```

This time the balancing property is satisfied, as shown here:

```
****************************************************
Algorithm to estimate the propensity score
****************************************************

The treatment is dfmfd

     HH has |
     female |
microcredit |
participant |
 : 1=Y, 0=N |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        534       47.30       47.30
          1 |        595       52.70      100.00
------------+-----------------------------------
      Total |      1,129      100.00

Estimation of the propensity score

(sum of wgt is  1.1260e+03)
Iteration 0:  log pseudolikelihood = -750.38718
Iteration 1:  log pseudolikelihood = -682.82636
Iteration 2:  log pseudolikelihood = -680.63459
Iteration 3:  log pseudolikelihood = -680.62452
Iteration 4:  log pseudolikelihood = -680.62452

Probit estimates                           Number of obs     =    1129
                                           Wald chi2(11)     =   85.21
                                           Prob > chi2       =  0.0000
Log pseudolikelihood = -680.62452          Pseudo R2         =  0.0930
----------------------------------------------------------------------
            |             Robust
      dmmfd |     Coef.  Std. Err.      z   P>|z|  [95% Conf. Interval]
------------+---------------------------------------------------------
    sexhead |  -.037986   .1662857  -0.23  0.819     -.3639   .287928
    agehead |  .0013931   .0037305   0.37  0.709  -.0059185   .0087047
   educhead | -.0465567   .0151559  -3.07  0.002  -.0762618  -.0168516
     lnland | -.6662184    .101586  -6.56  0.000  -.8653232  -.4671136
    vaccess | -.1173796     .13358  -0.88  0.380  -.3791916   .1444323
      pcirr |  .4304416    .154365   2.79  0.005   .1278917   .7329915
```

```
        rice |   .0571981   .0307982    1.86   0.063  -.0031652    .1175615
       wheat |  -.0055393    .056959   -0.10   0.923  -.1171769    .1060982
        milk |    .015395   .0184184    0.84   0.403  -.0207044    .0514944
         oil |   .0235048     .01239    1.90   0.058   -.000779    .0477887
         egg |  -.1114687   .1647319   -0.68   0.499  -.4343373    .2113999
       _cons |  -1.483823   .7367316   -2.01   0.044  -2.927791   -.0398558
------------------------------------------------------------------------------

Note: the common support option has been selected
The region of common support is [.02576077, .71555996]

------------OUTPUT OMITTED----------

**********************************************************
Step 2: Test of balancing property of the propensity score
Use option detail if you want more detailed output
**********************************************************

The balancing property is satisfied

------------OUTPUT OMITTED----------

*******************************************
End of the algorithm to estimate the pscore
*******************************************
```

With the propensity scores generated, the outcomes of interest (such as total per capita expenditure) between the treatment group and the matched control group are now compared to see whether the microcredit programs affect the outcome of interest. The following sections estimate the treatment effect of microcredit program participation, using different matching techniques that are available.

Average Treatment Effect Using Nearest-Neighbor Matching

The command to estimate the average treatment effect on the treated group using nearest-neighbor matching is "attnd." Following is the application of the "attnd" command to estimate the average treatment effect of female participation in microcredit programs on per capita total expenditure using nearest-neighbor matching:

```
attnd lexptot dfmfd [pweight=weight], pscore(ps98) comsup;
```

Estimating "attnd" with or without weights does not affect the results. Just for the purpose of this exercise, "attnd" was shown with weights estimation.

As the following output shows, female microcredit participation does have a significant impact on household per capita expenditure with the nearest-neighborhood matching method ($t = 3.256$). The average treatment of the treated (ATT) on per capita expenditure for female program participation is 13.6 percent.

```
ATT estimation with Nearest Neighbor Matching method
(random draw version)
Analytical standard errors
---------------------------------------------------------
n. treat.     n. contr.        ATT      Std. Err.       t
---------------------------------------------------------
      595           293      0.136          0.042   3.256
---------------------------------------------------------
Note: the numbers of treated and controls refer to actual
nearest-neighbor matches
```

Average Treatment Effect Using Stratification Matching

The "atts" command calculates the average treatment effect on the treated using stratification matching. To estimate the average treatment effect of female participation on the treated for per capita total expenditure, use the following:

```
atts lexptot dfmfd, pscore(ps98) blockid(blockf1) comsup
```

The result that follows shows a 9.9 percent increase in per capita expenditure because of women's participation in the microcredit programs. The impact is significant at the 5 percent level ($t = 3.320$).

```
ATT estimation with the Stratification method
Analytical standard errors
---------------------------------------------------------
n. treat.     n. contr.        ATT      Std. Err.       t
---------------------------------------------------------
      595           529      0.099          0.030   3.320
---------------------------------------------------------
```

Average Treatment Effect Using Radius Matching

The "attr" command calculates the average treatment effect on the treated using radius matching. Following is a demonstration:

```
. attr lexptot dfmfd, pscore(ps98) radius(0.001) comsup
```

The result shows an increased impact (14.6 percent) with high significance ($t = 3.793$) of women's microcredit participation on per capita expenditure:

```
ATT estimation with the Radius Matching method
Analytical standard errors
---------------------------------------------------------
n. treat.     n. contr.        ATT      Std. Err.       t
---------------------------------------------------------
      478           386      0.146          0.039   3.793
---------------------------------------------------------
Note: the numbers of treated and controls refer to actual
matches within radius
```

Average Treatment Effect Using Kernel Matching

The "attk" command computes the average treatment effect using kernel-based matching. The "reps" option performs the bootstrapping 50 times.

```
attk lexptot dfmfd, pscore(ps98) comsup bootstrap reps(50)
```

Results are consistent with earlier findings. Women's participation increases per capita expenditure by 4 percent at a 5 percent significance level.

```
ATT estimation with the Kernel Matching method
Bootstrapped standard errors
---------------------------------------------------------
n. treat.     n. contr.        ATT      Std. Err.       t
---------------------------------------------------------
      595       529           0.107          0.032    3.331
---------------------------------------------------------
```

Checking Robustness of Average Treatment Effect

There are several ways to check robustness of the findings. One approach is to estimate the propensity score equation and then use the different matching methods previously discussed to compare the results. The findings with different matching techniques are quite consistent.

Another way to check robustness is to apply *direct* nearest-neighbor matching instead of estimating the propensity score equation first. Stata has a command ("nnmatch") to do that. If both methods give similar results, then the findings are assumed to be more reliable.

The following Stata command will estimate the average treatment effect on the outcome of interest using direct nearest-neighbor matching with one match per treatment. The "m" option specifies the number of matches closest to the treated observations.

```
nnmatch lexptot dfmfd sexhead agehead educhead lnland vaccess
pcirr rice wheat milk oil egg, tc(att) m(1);
```

Results are again consistent with earlier findings. A 13.6 percent positive impact of microcredit participation is seen at a 5 percent significance level.

```
Matching estimator: Average Treatment Effect for the Treated

Weighting matrix: inverse variance     Number of obs          =  1129
                                       Number of matches (m)  =     1
----------------------------------------------------------------------
   lexptot |     Coef.   Std. Err.      z   P>|z|  [95% Conf. Interval]
-----------+----------------------------------------------------------
      SATT |  .1360462    .0377988   3.60  0.000     .061962  .2101304
----------------------------------------------------------------------

Matching variables: sexhead agehead educhead lnland vaccess pcirr rice
wheat milk oil egg
```

Further Exercises

Do the same exercise using male participation ("dmmfd"). Discuss your results.

Reference

Becker, Sascha, and Andrea Ichino. 2002. "Estimation of Average Treatment Effects Based on Propensity Scores." *Stata Journal* 2 (4): 358–77.