

```
. /* Workshop of impact evaluation. MEASURE Evaluation-INSP, 2015*/
```

```
*****
*****
```

```
DEMO: Propensity Score Matching Analysis
```

```
. *Description of the program:
```

```
. /*We want to evaluate the impact of a program on production abilities and the likelihood
that women engage in farming or animal raising. The program was implemented through the
participation of various NGOs, who presented the project for funding. Only women assigned to
the NGOs were the potential beneficiaries to the program.
```

```
To evaluate the program we had a non-experimental design, which included all beneficiary women
from the program. Women in the comparison group were selected within the communities where
women were participating in the intervention group and by interviewing their closest neighbor.
The treated/non-treated ratio was approximately 1:1.*/
```

```
. use demo_psm_SA, clear
```

```
.
```

```
. des id-nse3
```

variable name	storage type	display format	value label	variable label
id	str8	%9s		(ID) woman
state	str2	%9s		entidad federativa
municipality	float	%9.0g		group(entidad municip)
locality	float	%9.0g		group(entidad municip localid)
program	float	%9.0g	afirma	program
osc	float	%9.0g		Societal Civil Organization
osc1	byte	%8.0g		osc== 1.0000
osc2	byte	%8.0g		osc== 2.0000
osc3	byte	%8.0g		osc== 3.0000
osc4	byte	%8.0g		osc== 4.0000
osc5	byte	%8.0g		osc== 5.0000
osc6	byte	%8.0g		osc== 6.0000
osc7	byte	%8.0g		osc== 7.0000
osc8	byte	%8.0g		osc== 8.0000
osc9	byte	%8.0g		osc== 9.0000
osc10	byte	%8.0g		osc== 10.0000
age	byte	%9.0g		
agec	float	%13.0g	agec	categorical age
agec1	byte	%8.0g		agec==age 18-30 yrs
agec2	byte	%8.0g		agec==age 31-40 yrs
agec3	byte	%8.0g		agec==age 41-50 yrs
agec4	byte	%8.0g		agec==age 50 o more
kinship	float	%9.0g	kin	
kinship1	byte	%8.0g		kinship==Head
kinship2	byte	%8.0g		kinship==wife
kinship3	byte	%8.0g		kinship==Other
agehead	byte	%9.0g	m02d05	age head of HH
literacy	byte	%9.0g	afirma	Literacy: read & write spanish
language	byte	%9.0g	afirma	speak indigenous language
school	float	%17.0g	school	Schooling
school1	byte	%8.0g		school==illiteracy
school2	byte	%8.0g		school==primary
school3	byte	%8.0g		school==secondary or more
headschooll	float	%17.0g	school	schooling head of HH
headschooll1	byte	%8.0g		headschooll==illiteracy
headschooll2	byte	%8.0g		headschooll==primary
headschooll3	byte	%8.0g		headschooll==secondary or more
farming	byte	%9.0g	afirma	self-employ in farm activities
bluecollar	byte	%9.0g	afirma	work at factory
selling	byte	%9.0g	afirma	work in selling products
homework	byte	%9.0g	afirma	Work only at home
oportunidades	byte	%9.0g	afirma	Beneficiary of Oportunidades program
procampo	byte	%9.0g	afirma	(hh) beneficiary of Procampo program

```

loctime      byte    %11.0g    m02d37    years living in the town
voting       byte    %11.0g    afirma    If vote at presidential elections 2000
pregnancy    float  %9.0g    afirma    She was pregnant 1 year before
children     byte    %9.0g    afirma    She have children
children12   byte    %9.0g    afirma    She have children<12 yrs
HHsize       byte    %9.0g                (hh) size
child05      float  %9.0g    afirma    (hh) with children 0-5 yrs
pers517      float  %9.0g    afirma    (hh) with persons 5-17 yrs
pers60       float  %9.0g    afirma    (hh) with persons 60 o more
HHperswk     float  %9.0g    work      (hh) working persons
hhwk1        byte    %8.0g                HHperswk== 0.0000
hhwk2        byte    %8.0g                HHperswk== 1.0000
hhwk3        byte    %8.0g                HHperswk== 2.0000
discap       byte    %19.0g   afirma    (hh) any person with disability
foodexp      float  %9.0g                (hh) monthly food expenditure
totalexp     float  %9.0g                (hh) monthly total expenditure
oproductive  float  %9.0g    afrima    Participating in productive organizations
before-program
osocial      float  %9.0g    afirma    Participating in social organizations before-
program
NSE          float  %9.0g                Socio-economic level
nse_st       float  %12.0g   nse       NSE stratified
nse1         byte    %8.0g                nse_st==xxbajo
nse2         byte    %8.0g                nse_st==xbajo
nse3         byte    %8.0g                nse_st==bajo y medio

```

```

. /*Checking assignment to program*/

```

```

. tab program

```

program	Freq.	Percent	Cum.
No	672	52.58	52.58
yes	606	47.42	100.00
Total	1,278	100.00	

We suspect the presence of selection bias in the participation to the NGOs. Therefore the women in the intervention and control groups could likely be not comparable in observable and unobservable characteristics. We explore this hypothesis with the observed X/

```

/*We check the distribution of the outcome variable (farming), by program status*/

```

```

. tab program, sum(farming)

```

program	Summary of self-employ in farm activities		
	Mean	Std. Dev.	Freq.
No	.32985075	.47050999	670
yes	.52640264	.49971489	606
Total	.42319749	.49425988	1276

```

. logit farming program, cluster(locality)

```

```

Iteration 0:  log pseudolikelihood = -869.34275
Iteration 1:  log pseudolikelihood = -844.05583
Iteration 2:  log pseudolikelihood = -844.03075
Iteration 3:  log pseudolikelihood = -844.03075

```

```

Logistic regression                                Number of obs =      1276
                                                    Wald chi2(1)      =      29.44
                                                    Prob > chi2       =      0.0000
Log pseudolikelihood = -844.03075                Pseudo R2        =      0.0291

```

(Std. Err. adjusted for 51 clusters in locality)

farming	Robust			
	Coef.	Std. Err.	z	P> z
				[95% Conf. Interval]

program	.8145691	.1501186	5.43	0.000	.5203421	1.108796
_cons	-.7088602	.1823298	-3.89	0.000	-1.06622	-.3515004

. /*This would be the effect of the program if the assignment of the treatment was random*/

. disp .526-.3299
.1961

. /*Due to the non experimental design, we need to control for potential factors influencing participation in the program. The challenge is to identify a comparison group as similar to the intervention group.

An alternative technique is the Propensity Score Matching method*/

. /*STEP 1. Estimate the probability of participation to program using all the information, this is the PROPENSITY SCORE*/

. probit program agec2 agec3 agec4 kinship2 kinship3 agehead language school2 school3 pregnancy oportunitades procampo oproductive osocial loctime voting children12 child05 pers517 pers60 HHsize hhwk2 hhwk3 discap nse2 nse3 osc2 osc3 osc4 osc5 osc6 osc7 osc8 osc9 osc10 munc2-munc33

(some output ommited)

Iteration 0: log likelihood = -860.52644
Iteration 1: log likelihood = -531.87361
Iteration 2: log likelihood = -525.24599
Iteration 3: log likelihood = -525.15128
Iteration 4: log likelihood = -525.13571
Iteration 5: log likelihood = -525.13287
Iteration 6: log likelihood = -525.13231
Iteration 7: log likelihood = -525.13219
Iteration 8: log likelihood = -525.13218

Probit regression	Number of obs	=	1242
	LR chi2(57)	=	670.79
	Prob > chi2	=	0.0000
Log likelihood = -525.13218	Pseudo R2	=	0.3898

program	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
agec2	.1736858	.1371879	1.27	0.205	-.0951976 .4425692
agec3	-.0979386	.1861497	-0.53	0.599	-.4627854 .2669082
agec4	.2291391	.2351006	0.97	0.330	-.2316496 .6899279
kinship2	.3684817	.1333111	2.76	0.006	.1071968 .6297667
kinship3	.9273747	.2806919	3.30	0.001	.3772286 1.477521
agehead	-.0003433	.0047305	-0.07	0.942	-.0096149 .0089283
language	-.1875199	.1569745	-1.19	0.232	-.4951842 .1201444
school2	-.0173521	.117724	-0.15	0.883	-.2480869 .2133826
school3	.117212	.1758024	0.67	0.505	-.2273543 .4617783
pregnancy	.2950594	.1542258	1.91	0.056	-.0072177 .5973365
oportunida~s	1.184549	.1278025	9.27	0.000	.9340609 1.435037
procampo	.2991129	.1273348	2.35	0.019	.0495413 .5486845
oproductive	2.332739	.1337913	17.44	0.000	2.070513 2.594966
osocial	.1207714	.1169756	1.03	0.302	-.1084966 .3500394
loctime	.004189	.0039221	1.07	0.285	-.0034981 .0118761
voting	.2129515	.1232043	1.73	0.084	-.0285245 .4544274
children12	.0504109	.1338801	0.38	0.707	-.2119892 .3128111
child05	-.1890674	.1213808	-1.56	0.119	-.4269693 .0488346
pers517	.0842832	.1308217	0.64	0.519	-.1721227 .3406891
pers60	.0521885	.134913	0.39	0.699	-.2122362 .3166131
HHsize	-.0792429	.0305646	-2.59	0.010	-.1391483 -.0193374
hhwk2	-.0841241	.1744758	-0.48	0.630	-.4260904 .2578422
hhwk3	.0369288	.2095153	0.18	0.860	-.3737137 .4475713

```

discap | .2294439 .2092635 1.10 0.273 -.1807049 .6395927
nse2 | .0460404 .1250811 0.37 0.713 -.199114 .2911948
nse3 | -.0984854 .1903029 -0.52 0.605 -.4714723 .2745015
osc2 | 1.24663 .5811016 2.15 0.032 .1076915 2.385568
osc3 | 1.391183 .4520688 3.08 0.002 .5051449 2.277222
osc4 | -.8096703 .499725 -1.62 0.105 -1.789113 .1697728
osc5 | 3.45034 118.7532 0.03 0.977 -229.3017 236.2023
osc6 | .8931658 .4225099 2.11 0.035 .0650616 1.72127
osc7 | .8757222 .400065 2.19 0.029 .0916091 1.659835
osc8 | .6287982 .4378696 1.44 0.151 -.2294105 1.487007
osc9 | 2.513538 .6098582 4.12 0.000 1.318238 3.708838
osc10 | 1.228828 .3825317 3.21 0.001 .4790792 1.978576
munc2 | .1037635 .4328688 0.24 0.811 -.7446437 .9521708
munc3 | -.4860706 .3876873 -1.25 0.210 -1.245924 .2737825
munc4 | -1.285629 .4167301 -3.09 0.002 -2.102405 -.4688529
munc5 | .4710741 .3574622 1.32 0.188 -.2295389 1.171687
munc6 | 1.553711 .4111956 3.78 0.000 .7477826 2.35964
munc9 | -1.766728 118.7524 -0.01 0.988 -234.5171 230.9837
munc10 | -2.378333 118.7527 -0.02 0.984 -235.1293 230.3727
munc11 | .7612651 .4689084 1.62 0.104 -.1577786 1.680309
munc13 | -.6251579 .4280787 -1.46 0.144 -1.464177 .213861
munc15 | .465 .5472243 0.85 0.395 -.6075399 1.53754
munc17 | -1.413966 .5384064 -2.63 0.009 -2.469223 -.3587089
munc18 | .1898455 .3799389 0.50 0.617 -.554821 .9345119
munc20 | .0677735 .3644694 0.19 0.852 -.6465733 .7821203
munc22 | .4294785 .3523061 1.22 0.223 -.2610287 1.119986
munc23 | .2787817 .337271 0.83 0.408 -.3822572 .9398207
munc25 | .9587129 .3657202 2.62 0.009 .2419145 1.675511
munc26 | .9092476 .4915283 1.85 0.064 -.0541302 1.872625
munc27 | -.0566235 .5041997 -0.11 0.911 -1.044837 .9315897
munc28 | -.4718881 .3039634 -1.55 0.121 -1.067645 .1238692
munc31 | -.1784544 .4029291 -0.44 0.658 -.968181 .6112722
munc32 | .1323968 .471577 0.28 0.779 -.7918771 1.056671
munc33 | .692131 .4978782 1.39 0.164 -.2836923 1.667954
_cons | -2.72643 .5185444 -5.26 0.000 -3.742758 -1.710101

```

```

-----
. /*STEP 2: We obtain the propensity score by PREDICTING THE PROBABILITY OF PARTICIPATION*/
.
. predict pscore if e(sample)
(option pr assumed; Pr(program))
(59 missing values generated)

.
. /*In this example, we will use the kernel algorithm for doing the matching*/
.
. psmatch2 program, pscore(pscore) kernel common bwidth(.08)

***** NOTE: psmatch2 has to be downloaded from the internet and installed in your STATA copy.

.
.
. /*BALANCED SAMPLE TESTS*/

.
. pstest agec2 agec3 agec4 kinship2 kinship3 agehead language school2 school3 pregnancy
oportunidades procampo oproductive osocial loctime voting children12 child05 pers517 pers60
HHsize hhwk2 hhwk3 discap NSE, support(_support) treated(program) summary both

```

```

-----
Variable      Sample |      Mean          %reduct |      t-test
              | Treated Control  %bias  |bias| |      t  p>|t|
-----+-----+-----+-----+-----+-----
agec2  Unmatched | .31518 .24851  14.8   |      2.66  0.008
              Matched | .30116 .28548   3.5   |      0.55  0.580
agec3  Unmatched | .19472 .17857   4.1   |      0.74  0.459
              Matched | .20656 .25528 -12.5  -201.7 |     -1.86  0.063

```

agec4	Unmatched	.24587	.17411	17.7		3.17	0.002
	Matched	.23938	.20441	8.6	51.3	1.35	0.176
kinship2	Unmatched	.77558	.73661	9.1		1.62	0.106
	Matched	.76641	.75763	2.0	77.5	0.33	0.740
kinship3	Unmatched	.05116	.06101	-4.3		-0.76	0.446
	Matched	.05019	.05051	-0.1	96.8	-0.02	0.981
agehead	Unmatched	45.941	42.348	23.0		4.10	0.000
	Matched	45.745	45.451	1.9	91.8	0.31	0.755
language	Unmatched	.70462	.69792	1.5		0.26	0.794
	Matched	.69498	.66276	7.0	-380.7	1.11	0.267
school2	Unmatched	.52475	.54464	-4.0		-0.71	0.477
	Matched	.51931	.48026	7.8	-96.3	1.26	0.209
school3	Unmatched	.16337	.18006	-4.4		-0.79	0.430
	Matched	.17181	.18792	-4.3	3.5	-0.67	0.500
pregnancy	Unmatched	.09241	.13839	-14.4		-2.56	0.011
	Matched	.0888	.09103	-0.7	95.2	-0.13	0.900
oportunida~s	Unmatched	.82673	.54464	63.7		11.30	0.000
	Matched	.79923	.80261	-0.8	98.8	-0.14	0.892
procampo	Unmatched	.30033	.1994	23.5		4.20	0.000
	Matched	.28958	.28005	2.2	90.6	0.34	0.734
oproductive	Unmatched	.60561	.08185	132.1		23.90	0.000
	Matched	.53861	.49795	10.3	92.2	1.31	0.191
osocial	Unmatched	.27888	.18155	23.3		4.17	0.000
	Matched	.2722	.32514	-12.7	45.6	-1.86	0.063
loctime	Unmatched	32.426	28.213	26.1		4.66	0.000
	Matched	31.851	32.361	-3.2	87.9	-0.52	0.607
voting	Unmatched	.84653	.74107	26.3		4.67	0.000
	Matched	.8417	.84908	-1.8	93.0	-0.33	0.743
children12	Unmatched	.60726	.66369	-11.7		-2.10	0.036
	Matched	.59653	.59186	1.0	91.7	0.15	0.879
child05	Unmatched	.35974	.49702	-28.0		-4.99	0.000
	Matched	.37645	.31682	12.2	56.6	2.02	0.044
pers517	Unmatched	.72442	.66071	13.8		2.46	0.014
	Matched	.72008	.72696	-1.5	89.2	-0.25	0.805
pers60	Unmatched	.25083	.2128	9.0		1.61	0.107
	Matched	.24517	.22572	4.6	48.8	0.74	0.461
HHsize	Unmatched	4.8548	4.7202	6.3		1.12	0.264
	Matched	4.861	4.7738	4.1	35.2	0.64	0.519
hhwk2	Unmatched	.67492	.70833	-7.2		-1.29	0.196
	Matched	.66216	.65684	1.2	84.1	0.18	0.857
hhwk3	Unmatched	.18482	.15476	8.0		1.43	0.152
	Matched	.19112	.15376	9.9	-24.3	1.59	0.112
discap	Unmatched	.06106	.03869	10.3		1.85	0.065
	Matched	.05405	.05227	0.8	92.0	0.13	0.898
NSE	Unmatched	6.9059	6.2813	18.3		3.27	0.001
	Matched	7.1409	6.79	10.3	43.8	1.64	0.100

Summary of the distribution of the abs(bias)

BEFORE MATCHING

Percentiles		Smallest		
1%	1.463554	1.463554		
5%	3.985352	3.985352		
10%	4.142122	4.142122	Obs	25
25%	7.2347	4.281651	Sum of Wgt.	25
50%	13.82887		Mean	20.19649
		Largest	Std. Dev.	26.60591
75%	23.25818	26.2716		
90%	27.99257	27.99257	Variance	707.8744
95%	63.72905	63.72905	Skewness	3.272916
99%	132.0814	132.0814	Kurtosis	13.93032

AFTER MATCHING

Percentiles		Smallest		
1%	.1380121	.1380121		
5%	.6989866	.6989866		
10%	.764469	.764469	Obs	25
25%	1.493773	.8191938	Sum of Wgt.	25
50%	3.492205		Mean	4.995201
		Largest	Std. Dev.	4.259551
75%	8.613121	10.28414		
90%	12.15864	12.15864	Variance	18.14377
95%	12.49661	12.49661	Skewness	.6079733
99%	12.65049	12.65049	Kurtosis	1.863513

Sample	Pseudo R2	LR chi2	p>chi2
Unmatched	0.301	532.89	0.000
Matched	0.022	31.91	0.161

```
. /*We see that matching significantly decreased unbalance in the samples*/
```

```
. /*THIS IS THE ACTUAL ESTIMATION OF THE ATT*/
```

```
. psmatch2 program, out(farming) pscore(pscore) kernel common bwidth(.08)
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
farming	Unmatched	.52640264	.345425868	.180976773	.027705123	6.53
	ATT	.552123552	.405480544	.146643008	.046162109	3.18

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	0	634	634
Treated	88	518	606
Total	88	1,152	1,240

```
. keep if _support==1  
(149 observations deleted)
```

```
. /*WE NEED TO DO A BOOTSTRAP TO OBTAIN CORRECT STANDARD ERRORS*/
```

```
. bootstrap r(att), reps(150): psmatch2 program, kernel pscore(pscore) out(farming)
bwidth(.08)
```

```
(running psmatch2 on estimation sample)
```

```
Note: S.E. does not take into account that the propensity score is estimated.
```

```
Bootstrap replications (150)
```

```
-----+----- 1 -----+----- 2 -----+----- 3 -----+----- 4 -----+----- 5
..... 50
..... 100
..... 150
```

```
Bootstrap results                               Number of obs    =      1152
                                                Replications    =       150
```

```
command: psmatch2 program, kernel pscore(pscore) out(farming) bwidth(.08)
         _bs_1: r(att)
```

```
-----+-----
      |      Observed      Bootstrap      Normal-based
      |      Coef.      Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
     _bs_1 |      .146643      .0528458      2.77   0.006      .0430671      .2502189
-----+-----
```

```
*** In conclusion, our ATT estimate of the impact of the program is an increase of 14
percentage points. This is down from 19 percentage points (our "naive" estimate).
```