Universidad Torcuato Di Tella

Maestría en Econometría

Master's Thesis: Causal Inference using STATA

Maria Cecilia Soto[1]

Thesis Tutor: Andrea Rotnitzky

September 2010

Abstract

This work has two main objectives: first to provide a short overview of available analytical methods that estimate Causal Effect measures when "association is not causation" and then to introduce a set of programs which estimate them.

The methods used are: Outcome Regression adjustment, Inverse Weighted probability, Double Robust bounded and Stratification by the propensity score.

In order to implement such methods we have developed five programs using STATA[2] software for both continuous and binary outcomes. When the outcome variable is binary the programs outputs estimators of the Average Treatment effect (ATE), the Causal Risk ratio (CCR) and the Causal Odd ratio (COR) while if the outcome variable is continuous it only outputs the ATE. In addition we constructed a special program (prop_score.ado) for the evaluation of the propensity score fit in order to use it in the propensity score stratification method.

These programs are: t_out_reg.ado, t_ipw.ado, t_prop_stat.ado, the dr_bounded.ado and the t_prop_score.ado.

Key words: Causal Inference, Outcome Regression adjustment, Inverse Weighted probability, Double Robust bounded, Propensity Score, Stratification by the propensity score.

[1] I am grateful to Andrea Rotnitzky for her excellent tutoring. She has been extremely generous with her time, knowledge and material. All remaining errors are my sole responsibility

[2] These programs were created in STATA 10

Introduction:

Sometimes, researchers are interested in answering questions such us: Is a drug beneficial for decreasing the number of deaths from a specific illness? Or does a public program designed to foster investment in R&D firms' expenditures have the expected result of the intervention? In these treatment evaluation questions the analyst is generally interested in the presence of an effect, its direction and magnitude.

Although the topic of treatment evaluation has a long development in the fields of biological and medical research, it has only recently become important in economic/social research. In the latter, most of the research in treatment impact evaluation concerns labor economics and tax policy applications. Despite the fact that the different programs created in order to estimate the Causal measures can be applied in Economic or social fields, the examples presented in this work are all applied to heath issues.

Before starting, it is important to understand the concept of counterfactual framework (Rubin, 1978). Let $\mathrm{Y_{0i}}$ (Y_{1i}) be the, possibly counterfactual, outcomes that would be observed if subject i took treatment 0 (treatment 1). Depending on the subjects' exposure to the treatment, only one possible outcome can be observed. For example if the subject i is treated, the outcome Y_{1i} will be known. But the analyst will also want to know what would have happened to this subject if it had not been treated. In other words, the analyst is also interested in knowing $\mathrm{Y_{0i}}$ (unobservable counterfactual). Because it is impossible to get both effects at the same time on an individual level, it is necessary to work in an aggregated way. In this work we will refer to the subjects included in the treatment group as treatment group while the untreated/control group will include all the subjects which have not received the treatment.

The first section will try to explain under which type of experiment and assumptions the analyst can aspire to know the Causal Treatment Effect while using the information of the treatment and untreated group. It briefly presents the difference between randomized studies and observational ones, and emphasizes the assumptions that the analyst has to assume in the latter type of studies in order to identify the Causal Effect measures.

The second section presents the Linder data set. This dataset will be used to illustrate how the different programs work.

The third section presents a theoretical framework for each method followed by the syntax of the program which estimates each method and gives an example using the Lindner Data set using that software. The results of each program will be shown using a binary outcome variable and a continuous one. Standard errors of the different Causal measures are obtained by the bootstrap method.

The last section presents a comparison of the different estimators of the contrasts according to the different methods.

Section 1

The following causal contrasts are output by our programs.

Average Treatment Effect (ATE):

$$\mathrm{ATE} = \mathrm{E}(\mathrm{Y}_1) - \mathrm{E}(\mathrm{Y}_0)$$

Causal Risk ratio (CRR):

$$\mathrm{CRR} = \mathrm{E}(\mathrm{Y}_1)/\,\mathrm{E}(\mathrm{Y}_0)$$

Causal Odd ratio (COR):

$$\mathrm{COR} = \frac{\mathrm{E}(\mathrm{Y}_1)}{1-\mathrm{E}(\mathrm{Y}_1)} / \frac{\mathrm{E}(\mathrm{Y}_0)}{1-\mathrm{E}(\mathrm{Y}_0)}$$

Randomized Studies

The Causal Effect contrasts (ATE, CRR and COR) are identified if data come from randomized studies. In these studies each subject from a sample is randomized to the treatment group with probability p and with probability 1 – p to the control group.

Because in randomized studies group membership is determined by a random mechanism, then it holds that:

$$Y_a \coprod A, for\ a = 0{,}1 \qquad \textbf{(1)}$$

where $\coprod$ stands for independence.

Consequently, if

$$p = P(A = a) > 0\ for\ a = 0{,}1 \qquad \textbf{(2)}$$

then

$$\mathrm{E}(\mathrm{Y_a}) = \mathrm{E}(\mathrm{Y_a} \backslash \mathrm{A} = \mathrm{a})$$

In addition, if, as it is reasonable to assume in randomized studies,.

$$Y_a = Y\ when\ A = a, for\ a = 0{,}1 \qquad \textbf{(3)}$$

Then we conclude that $\mathrm{E}(\mathrm{Y_a} \backslash \mathrm{A} = \mathrm{a}) = \mathrm{E}(\mathrm{Y} \backslash \mathrm{A} = \mathrm{a})$

Thus, under (1) (2) and (3), the probability distribution of the counterfactuals $Y_a, a = 0,1$ can be written in terms of the distribution of the observed data (Y, A) and hence it is identified (Robins, 1986; see also, Lunceford and Davidian, 2004).

$$\mathrm{E}(\mathrm{Y_a}) = \mathrm{E}(\mathrm{Y_a} \backslash \mathrm{A} = \mathrm{a}) = \mathrm{E}(\mathrm{Y} \backslash \mathrm{A} = \mathrm{a}) \qquad \textbf{(4)}$$

Restriction (1) is known as the randomization assumption. Restriction (2) is known as the positivity assumption and Restriction (3) is known as the consistency assumption.

Observational Studies

In an observational study the assignment to the treatment or control group is not under the control of the investigator. Consequently, the randomization restriction (1) is not guaranteed to hold. However If L were a vector that includes all prognostic factors that are used to decide treatment group, the following restriction would hold.

$$Y_a \coprod A \backslash L \qquad for\ a = 0,1 \qquad \textbf{(5)}$$

Restriction (5) is often referred to as conditional randomization. In contrast to randomized studies, in observational studies restriction (5) is an assumption which is not guaranteed to hold, because it is never possible to know if indeed the vector L that one is able to measure contains all the prognosis factors used to select treatment group.

Under (5) and if the following holds:

$$P(A = a \backslash L) > 0 \quad for\ a = 0,1 \qquad \textbf{(6)}$$

it follows that:

$$E(Y_a \backslash L) = E(Y_a \backslash A = a, L)$$

If in addition consistency (3) holds, then we have:

$$E(Y_a \backslash L) = E(Y_a \backslash A = a, L) = E(Y \backslash A = a, L)$$

Restriction (6) is again referred to as the positivity assumption.

So, finally we have:

$$E(Y_a) = \int (E(Y \backslash A = a, L = l)\ \ f_L(l)\, dl$$

Therefore under (3) (5) and (6) it follows that $E(Y_a)$ is identified and so are ATE, CRR and COR.

Furthermore,

$$E(Y_a) = E\left[\frac{I(A=a)}{P(A=a\backslash L)}Y\right]$$

Because with $\pi_a(L) = P(A=a\backslash L)$ we have that by consistency

$$E\left(\frac{I(A=a)}{\pi_a(L)}Y\right) = E\left(\frac{I(A=a)}{\pi_a(L)}Y_a\right)$$

By applying double expectations

$$= E\left(E\left[\frac{I(A=a)}{\pi_a(L)}Y_a \middle| L, Y_a\right]\right)$$

$$= E\left(E[I(A=a)\backslash L, Y_a]\frac{1}{\pi_a(L)}Y_a\right)$$

And by conditional randomization

$$E[I(A=a)\backslash L, Y_a] = \pi_a(L)$$

Therefore

$$= E\left(\pi_a(L)\frac{1}{\pi_a(L)}Y_a\right)$$

$$= E(Y_a)$$

The quantity

$$\pi_a(L) = P(A=a\backslash L)$$

is called the propensity score for treatment a

See for example Hernan and Robins, 2006; also Lunceford and Davidian, 2004.

Section 2

In this section we present the Lindner data set used in the examples. With this data we will illustrate how the different programs work.

The Lindner data set is a subset of that analyzed in the study by Kereiakes, 2000, and publicity available in the R- data repository. It has information on percutaneous coronary interventions (PCI) on 996 patients by Ohio Heart Health Center operators at The Christ Hospital, in 1997. It is an observational study which has patient demographics and procedural data collected by the interventional physician. The Hospital charges were obtained from the McKesson/HBOC TrendStar decision support software system and mortality information was collected from follow-up telephone contact with patients and/or their families 6 months after PCI.

Our objective is to illustrate the estimation of the causal contrast (ATE, CRR and COR) of the effect of **Abciximab,** a platelet glycoprotein (GP) IIb/IIIa receptor blockade administrated during PCI, on all-cause 6-month mortality and of the contrast ATE for estimating the effect on the medical costs incurred (within 6 months of initial PCI) using different estimators of the causal contrasts.

The substantive questions are:

1) **Does Abciximab have a beneficial Causal Effect on decreasing the numbers of deaths?**
2) **Does Abciximab have a Causal Effect on medical cost?**

Basic variables:

- Exposure A → abcix is equal to 1 if the patient has received Abciximab during PCI and is equal to 0 in the contrary case.
- Outcome Y (To answer question 1) → dead is equal to 1 if the patient is dead and is equal to 0 in the contrary case.
- Outcome Y (To answer question 2) → cardbill = Cardiac related costs incurred within 6 months of patient's initial PCI, measured in 1998 US dollars.
- Baseline variables: Different characteristics of the patients:

  Stent → is equal to 1 if the patient has a coronary stent deployment and is equal to 0 in the contrary case.

  Height → Height in centimeters.

  Female → is equal to 1 if the patient's gender is female and is equal to 0 in the contrary case.

  Diabetic → is equal to 1 if the patient has a Diabetes mellitus diagnosis and is equal to 0 in the contrary case.

  Acutemic → is equal to 1 if the patient had an Acute myocardial infarction within the previous 7 days and is equal to 0 in the contrary case.

Ejecfrac → Left ejection fraction, measured from 0 percent to 90 percent.

Ves1proc → Number of vessels involved in the patient's initial PCI procedure, measured from 0-5.

Abciximab was administered to 70.1 % of the total patients and the 6 month mortality rate was different between the treated and the untreated groups, 1.58% in the former and 5.03 % in the latter. As reported in Kereiakes, 2000, treated patients compared to untreated ones were less often diabetic and were more likely to have incurred a myocardial infarction within 30 days before PCI, had lower left ventricular ejection fractions, more coronary vessels undergoing angioplasty and were more likely to have a coronary stent deployed. Thus, overall, treated patients had a worse prognosis that untreated patients.

Section 3

Outcome regression adjustment:

Binary outcome:

Let $p_i = P(Y_i = 1 \backslash A_i, L_i)$ be the conditional probability of survival for the ith subject in the entire cohort, given its treatment status and its baseline covariates.

If the outcome Y_i is binary this method first fits a logistic regression model of Y_i on A_i and L_i. For example,

$$log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 A_i + \beta_2^T L_i$$

Then, the fitted values are computed

$$\hat{p}_{ai} = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 a + \hat{\beta}_2^T L_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 a + \hat{\beta}_2^T L_i}}$$

Finally, the outcome regression estimator of $P(Y_a = 1)$ is

$$\hat{e}_{a,R} = n^{-1} \sum_{i=1}^{n} \hat{p}_{ai}$$

When the assumed logistic regression model is correct $\hat{e}_{a,R}$ Is a consistent estimator of $P(Y_a = 1)$. Then a consistent estimator of the causal odds ratio is:

$$\widehat{COR} = \frac{\hat{e}_{1,R}/1 - \hat{e}_{1,R}}{\hat{e}_{0,R}/1 - \hat{e}_{0,R}}$$

Likewise, a consistent estimator of the Average Treatment effect is:

$$\widehat{ATE} = \hat{e}_{1,R} - \hat{e}_{0,R}$$

And, a consistent estimator of Causal Risk ratio (CCR) is:

$$\widehat{CCR} = \hat{e}_{1,R}/\hat{e}_{0,R}$$

See for example Lunceford and Davidian, 2004; and D' Agostino, 1998.

If the outcome variable is continuous, we may fit a lineal regression model such as:

$$Y_i = \beta_0 + \beta_1 A_i + \beta_2^T L_i + error_i$$

then, the causal average in treatment $E(Y_i)$ is estimated by:

$$\hat{e}_{a,R} = n^{-1} \sum_{i=1}^{n} \hat{Y}_{ai}$$

where $\hat{Y}_{ai}$ is the fitted value for a subject i with covariates L_i if A_i is set to a, in our example

$$\hat{Y}_{ai} = \hat{\beta}_0 \quad + \hat{\beta}_1 A_i + \hat{\beta}_2^T L_{ai}$$

The regression model of our example does not include interactions with the treatment. Then, $\widehat{ATE} = \hat{e}_{1,R} - \hat{e}_{0,R}$ is equal to $\hat{\beta}_1$

1) The syntax of the command is as follows:

```
t_out_reg varlist,  treatvar(string) [dummies(varlist)]
[inter1(varlist)] [inter2(varlist)] [inter3(varlist)]
[inter4(varlist)] [inter5(varlist)] [inter6(varlist)]
[inter7(varlist)] [inter8(varlist)] [inter9(varlist)]
[inter10(varlist)] [binary] [bootstrap]
```

Description: The t_out_reg.ado.

The first variable of the varlist is the outcome variable (Y), and the next variables are the independent variables (L). The variable treatvar is the treatment variable (A). If the treatment variable is binary and the objective is to compute CRR and COR, then the binary option must be entered. The bootstrap option outputs estimates of the standard error of the different causal contrasts estimates by using the bootstrap method. All STATA bootstrap command options function in the t_out_reg.a program. (e.g. = reps()).

As this program was created in STATA 10, the option dummies were added to allow for dummies variables in the model. This is useful if there are categorical variables and it is desired to include a dummy variable for each category as independent variables. Likewise, option inter* allows the formation of interaction variables by including in each inter option the two variables to be multiplied.

Output of the t_out_reg.ado

The output reports the causal contrasts, its bootstrap SE estimator, the bootstrap CI based on the normal approx. and the bootstrap z value.

Example:

If the binary option is not chosen, only the effect ATE will be reported.

Command line: Continuous outcome

- `t_out_reg cardbill stent acutemi ejecfrac ///`
  `p ves1proc, treatvar(abcix) bootstrap reps(100)`

```
 ATE- Outcome Regression Adjustment
Bootstrapped standard errors
```

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z
1185.489	853.206	-496.996	2867.974	1.389

Otherwise, the ATE, COR and CRR effects will be reported.

Command line: Binary outcome

- `t_out_reg death stent acutemi ejecfrac ves1proc,///`
  `treatvar(abcix) bootstrap reps(200) binary`

ATE - Outcome Regression Adjustment - binary
Bootstrapped standard errors

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z
-0.056	0.019	-0.094	-0.018	-2.884

CRR - Outcome Regression Adjustment -binary
Bootstrapped standard errors

CRR	boots.Std.Err.	[95% Conf.	Interval(N)]	z
0.207	0.110	-0.011	0.425	1.875

COR - Outcome Regression Adjustment -binary
Bootstrapped standard errors

COR	boots.Std.Err.	[95% Conf.	Interval(N)]	z
0.195	0.092	0.014	0.376	10.110

Inverse probability weighting

As we have seen under (3) (5) and (6):

$$\boxed{E(Y_a) = E\left(\frac{I(A=a)}{\pi_a(L)}Y\right)}$$

Where

$$I(A=a) = \begin{cases} 1\ if\ A=a \\ 0\ if\ A\neq a \end{cases}$$

In observational studies, $\pi_a(L)$ is unknown and must be estimated. However when L is high dimensional, it is not possible to separately estimate the propensity score $\pi_a(l)$ for each possible value l of L. Instead a model, such as:

$$log\left\{\frac{\pi_1(L_i)}{1-\pi_1(L_i)}\right\} = \alpha_0 + \alpha_1^T L_i$$

is assumed, and $(\alpha_{0,}\alpha_{1,})$ is estimated by maximum likelihood.

Then $\pi_{a,i}(l) = \pi_a(l_i)$ is estimated with:

$$\hat{\pi}_{1,i} = \frac{e^{\hat{\alpha}_0 + \hat{\alpha}_0^T L_i}}{1 + e^{\hat{\alpha}_0 + \hat{\alpha}_0^T L_i}}$$

$$\hat{\pi}_{0,i} = 1 - \hat{\pi}_{1,i}$$

To compute the IPW estimator of $E(Y_a)$ the outcomes of those that took treatment A=a are averaged but weighted by $\frac{1}{\hat{\pi}_{a,i}}$, giving

$$\hat{e}_{IPW,1} = \frac{\sum_{i=1}^{n} \frac{A_i}{\hat{\pi}_{1i}} Y_i}{\sum_{i=1}^{n} \frac{A_i}{\hat{\pi}_{1i}}}$$

$$\hat{e}_{IPW,0} = \frac{\sum_{i=1}^{n} \frac{1 - A_i}{\hat{\pi}_{0i}} Y_i}{\sum_{i=1}^{n} \frac{1 - A_i}{\hat{\pi}_{0i}}}$$

2) The syntax of the command is as follows:

```
t_ipw ,outcome(string) treatvar(string) pvars(varlist) [binary]
[bootstrap]
```

Description: t_ipw.ado

The variable treatvar() is the treatment variable (A) and the outcome() must contain the outcome variable (Y). The pvar varlist contains the L variables for the propensity score model. The bootstrap option outputs estimates of the standard error of the different causal contrasts estimates by using the bootstrap method. All STATA bootstrap command options function in the t_out_reg.a program. (e.g. = reps()).

Output of the t_ipw.ado

Example

Command line: Continuous outcome

```
 t_ipw,   outcome(cardbill)   pvars(stent      acutemi   ejecfrac   ///
ves1proc p_inter*) treatvar( abcix) bootstrap rep(200)
```

ATE - IPW - Linear
Bootstrapped standard errors

ATE	Boots. sdt Err.	[95% Conf.	Interval(N)]	z
714.966	**1049.707**	**-1355.010**	**2784.943**	**0.681**

Command line: Binary outcome

- `t_ipw, outcome(death) pvars(stent acutemi ejecfrac /// ves1proc p_inter*) treatvar( abcix) binary bootstrap rep(200)`

beta- IPW
Bootstrapped standard errors

odds_rat	Boots.sdt Err.	[95% Conf.	Interval(N)]	z
0.171	**0.169**	**-0.162**	**0.504**	**1.011**

<u>Stratification by the propensity score:</u>

This method (Rosenbaum and Rubin, 1983, 1984) first computes the estimated propensity scores $\hat{\pi}_{1i}$ and then generates five strata according to the quintiles $\hat{q}_j, j = 0,1,2,3,5\ of \hat{\pi}_{1i}$ from the entire sample (both groups). Within each stratum the sample mean of Y_i is computed for those treated with treatment a. Finally $E(Y_a)$ is calculated with the average of the five sample means obtained in the last step. That is:

$$\hat{e}_{a,PS,strat} = \frac{1}{5}\sum_{j=1}^{5}\left\{\frac{1}{n_{a,j}}\sum_{i\ treated\ with\ a\ in\ strata\ j} Y_i\right\}$$

Where $n_{a,j}$ is the number of subjects treated with a in the j^{th} stratum

The theoretical justification for this estimator is based on the observation that if conditional randomization (5) holds, then:

$$Y_a \coprod A \backslash \pi_1(L) \qquad for\ a = 0,1 \qquad \textbf{(7)}$$

So that the propensity score is a scalar "covariate" sufficient to adjust for confounding. The method further assumes that:

$$Y_a \coprod A \backslash \pi_1^* \qquad for\ a = 0,1 \qquad \textbf{(8)}$$

where π_1^* is a categorical variable taking 5 levels indicating the quantiles of the distribution of $\pi_1(\mathrm{L})$. However, under (7), (8) holds only when within levels of π_1^* there is no residual confounding. To evaluate whether (8) holds (approximately) it is advisable to examine the degree of balance of $\pi_1(\mathrm{L})$ within each stratum, as well as for each covariate L. The function t_prop_score.ado was written to examine this issue.

This program presents different outputs like the region of overlap, two sample t tests comparing the treatment and control means of each of the covariates L, the Kolmogorov Smirnov test for the propensity score by quintiles and different graphics to help examine if (8) is approximately true.

A lack of overlap in the support of the distribution of the estimated propensity scores $\pi_1(L_i)$ in the two treatment groups is a problem because it suggests failure of the positivity assumption (6). One possible solution is to restrict the analysis data set to the subsample in which overlap is achieved. Generally, we achieved this by restricting the dataset to:

- Those untreated subjects with propensity scores for treatment lower than those of any treated subject.
- Those treated subjects with propensity scores for treatment higher than those of any untreated subject.

In such case, inference about causal effect is relevant to the subpopulation in which propensity scores overlap.

3) The syntax of the t_prop_score.ado is:

```
T_prop_score varlist [if] [in], name_ps(string) treatvar(string)
[probit] [dummies(varlist)] [inter1(varlist)] [inter2(varlist)]
[inter3(varlist)] [inter4(varlist)] [inter5(varlist)]
[inter6(varlist)] [inter7(varlist)] [inter8(varlist)]
[inter9(varlist)] [inter10(varlist)] [detail]
```

Description: T_prop_score.ado

The varlist must include the L variables. The name_ps is the name of the variable that will store the fitted values of the propensity score for A=1. The treatvar is the name of the treatment variable. The default model for the estimation of the propensity score is a logistic regression, the option probit must be selected if a probit model is preferred for the estimation. The options dummies and inter* allow for dummies and interactions as in the out_reg.ado command explained above. If the option detail is selected output will displayed to assist the analyst to decide if the propensity score is balanced within each stratum.

Output of the t_prop_score.ado

Command line:

- `t_prop_score stent acutemi ejecfrac ves1proc, /// name_ps(f) inter1(acutemi ejecfrac)inter2(acutemi ves1proc) /// treatvar (abcix) detail.`

- **Propensity Score estimation**

```
Logistic regression                               Number of obs   =        996
                                                  LR chi2(6)      =      87.80
                                                  Prob > chi2     =     0.0000
Log likelihood = -563.84274                       Pseudo R2       =     0.0722
```

abcix	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
stent	.5783757	.1502637	3.85	0.000	.2838643	.872887
acutemi	5.32626	1.929894	2.76	0.006	1.543738	9.108782
ejecfrac	-.0105431	.0075204	-1.40	0.161	-.0252829	.0041967
ves1proc	.8141389	.1454633	5.60	0.000	.529036	1.099242
p_intereje~c	-.0624801	.035369	-1.77	0.077	-.1318019	.0068418
p_interves~c	-.7794379	.4420002	-1.76	0.078	-1.645742	.0868667
_cons	-.1818307	.4441394	-0.41	0.682	-1.052328	.6886665

The region of overlap is [.3182774, .94462087]

abcix	overlap 0	1	Total
control	0	298	298
treat	27	671	698
Total	27	969	996

- **Distribution of control and treatment group by quintile**

	abcix		Total
	control	treat	
Quintile 1	92	102	194
Quintile 2	86	146	232
Quintile 3	51	109	160
Quintile 4	47	149	196
Quintile 5	22	165	187
Total	298	671	969

- **Two-sample t test with equal variances**

The estimator of the difference is diff = mean(0) - mean(1)

Where Ho: diff = 0

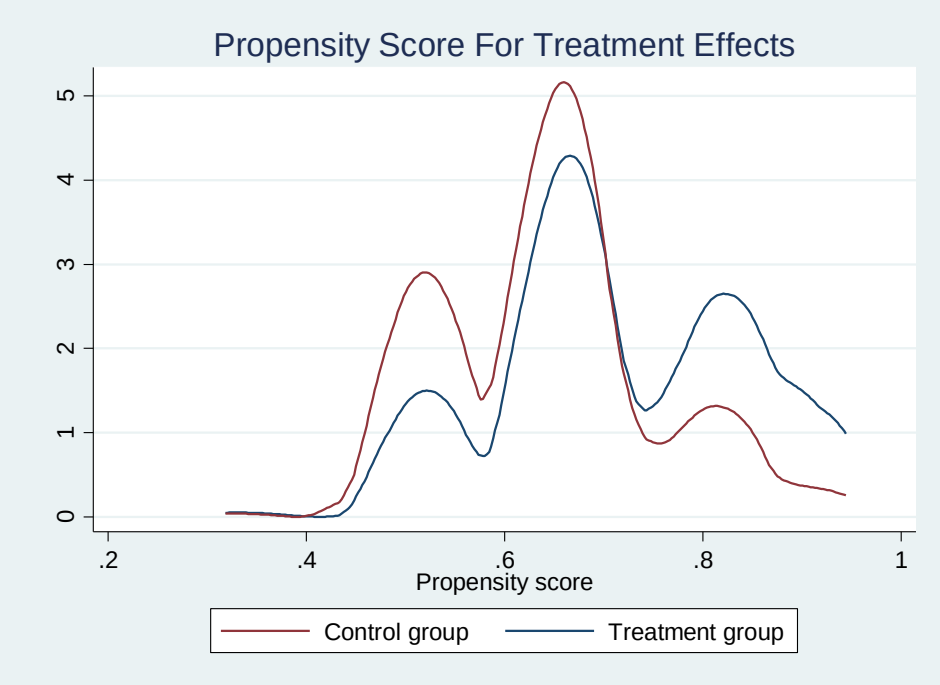
	Est_dif(Q1)	PV(Q1)	Est_dif(Q2)	PV(Q2)	Est_dif(Q3)	PV(Q3)	Est_dif(Q4)	PV(Q4)	Est_dif(Q5)	PV(Q5)
Pscore	0.00281828	0.5945	-0.00310073	0.025748	-0.00189853	0.3352	-0.01078185	0.1515	-0.0073672	0.4454
stent	-0.00873828	0.6245	-0.00955718	0.591001	0.02662349	0.6819	-0.08539197	0.2555	-0.0575758	0.5690
acutemi	0	.	0.01162791	0.193221	-0.03543803	0.3102	0.04041125	0.4169	-0.0666667	0.5592
ejecfrac	-0.17391304	0.8975	1.2871934	0.117239	-0.74833603	0.6094	-2.2681708	0.2584	-0.3606061	0.8574
ves1proc	0.0185422	0.3667	0	.	-0.02914193	0.6377	-0.13237184	0.0720	0.19090909	0.3227

- **Kolmogorov-Smirnov test for the propensity score, by quintile and for the whole sample:**

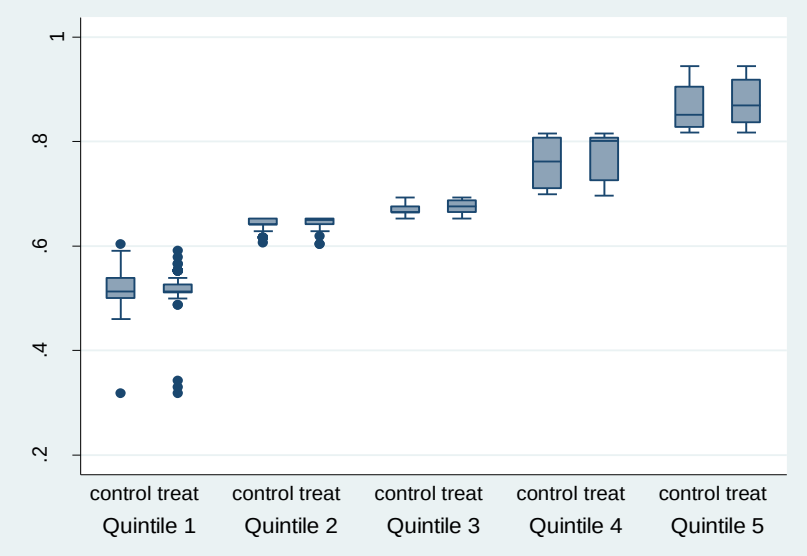
Where H0 = the distribution of the variable is equal between both groups

```
e(ksmirnov)[6,1]
            Kolmogorov-Smirnov(PV)
     Total                1.309e-11
Quintile_1                .76990825
Quintile_2                 .0173207
Quintile_3                 .9125785
Quintile_4                .39440343
Quintile_5                .65361195
```

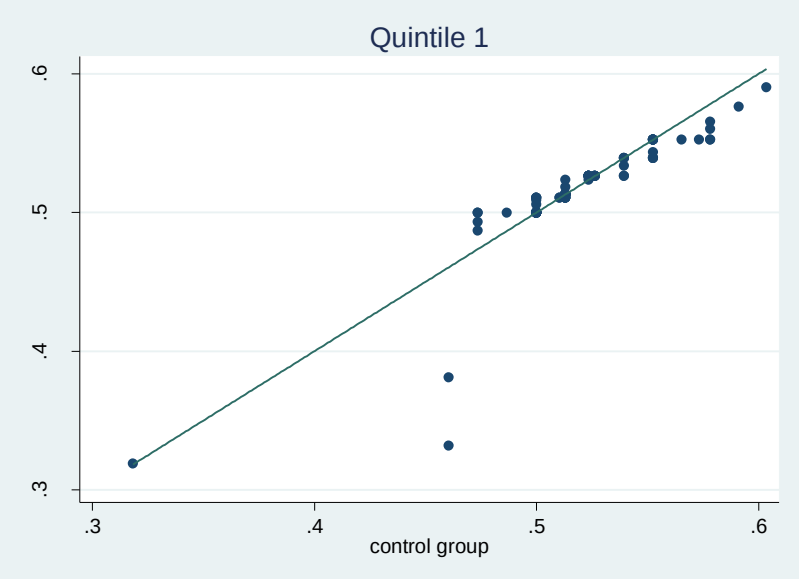
- **Propensity score Kernel density function for treatment effects:**



Box plots by quintile of the estimated propensity score.



- **Finally t_prop_score.ado presents a graphic such us the following one for each quintile**:



The following program computes the estimator of causal contrast by the propensity score stratification method.

4) The syntax of the t_prop_strat.ado is:

```
t_prop_strat, outcome(string) treatvar(string) pvars(varlist)
[bootstrap]
```

Description: t_prop_strat.ado

The variable treatvar() is the treatment variable (A) and the outcome() must contains the outcome variable (Y). The pvar varlist contains the L variables for the propensity score. The bootstrap option outputs estimates of the standard error of the different causal contrasts estimates by using the bootstrap method. All STATA bootstrap command options function in the t_out_reg.a program. (e.g. = reps()).

In this program it is not necessary to declare if the outcome variable is binary. The program realizes this by itself and returns the three effect contrasts (ATE, CRR and COR) if the outcome is binary while if the outcome variable is continuous only the ATE effect will be reported.

Output of the t_prop_strat.ado

Example:

Command line: Continuous outcome

- `t_prop_strat , outcome( cardbill) treatvar( abcix) //// pvars( stentves1proc ejecfrac acutemi p_interejecfrac /// p_interves1proc) bootstrap r(200)`

ATE- Stratification
Bootstrapped standard errors

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z
1140.773	880.266	-595.072	2876.619	1.296

Command line: Binary outcome

- `t_prop_strat , outcome(death) treatvar( abcix) /// pvars( stent ves1proc ejecfrac acutemi p_interejecfrac /// p_interves1proc) bootstrap r(200)`

ATE - Stratification- binary
Bootstrapped standard errors

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z
-0.048	0.017	-0.081	-0.015	-2.837

CRR - Stratification -binary
Bootstrapped standard errors

CRR	boots.Std.Err.	[95% Conf.	Interval(N)]	z
0.209	0.105	0.002	0.417	1.989

COR - Stratification -binary
Bootstrapped standard errors

COR	boots.Std.Err.	[95% Conf.	Interval(N)]	z
0.199	0.118	-0.033	0.431	1.695

Double robust bounded:

This method (Robins and Rotnitzky, 2001; see also Lunceford and Davidian, 2004) is a technique that requires two models specifications: The same regression outcome model as for the outcome regression adjustment and the same propensity score model for IPW estimation. As we saw above, the former relies on a regression model for the outcome Y given A and L while the latter relies on a logistic regression model for the relationship between the propensity score and L.

The distinctive characteristic of the double Robust bounded method is that it will give a consistent estimator of the causal contrast if at least one of the models is right but does not require that both models be correct nor that the analyst know which of the models is correct.

Recalling the outcome regression adjusted estimator, the **first step** of the Double Robust Bounded method is illustrated by the following example.

Let $p_i = P(Y_i = 1 \backslash A_i, L_i)$ be the conditional probability that $Y_i = 1$, given treatment status and baseline covariates.

We assume a model for p_i ,for example,

$$log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 A_i + \beta_2^T L_i$$

Then, we compute the fitted values

$$\hat{p}_{a,i} = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 a + \hat{\beta}_2^T L_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 a + \hat{\beta}_2^T L_i}}$$

The outcome regression estimator of $P(Y_a = 1)$ is

$$\hat{e}_{a,R} = n^{-1} \sum_{i=1}^{n} \hat{p}_{a,i}$$

The **second step** of is to compute augmentation term. This is:

$$\hat{d}_a = \frac{\sum_{all\ subjects\ i\ with\ A_i = a} \frac{1}{\hat{\pi}_{a,i}} (Y_i - \hat{p}_i)}{\sum_{all\ subjects\ i\ with\ A_i = a} \frac{1}{\hat{\pi}_{a,i}}}$$

Finally, the **Double Robust bounded** estimator is computed by adding the outcome regression estimator and augmentation term.

$$\underbrace{\hat{e}_{a,DR}}_{DR\ estimator} = \underbrace{\hat{e}_{a,R}}_{Out\ Reg\ estimator\ (step1)} + \underbrace{\hat{d}_a}_{Augmentation\ term\ (step\ 2)}$$

The same Causal measures (ATE, COR, CRR) presented for the outcome regression method are estimated in this method too.

5) The syntax of the command is as follows:

```
t_dr_bounded [if] [in] [, pvars(varlist) ovars(varlist)
treatvar(varname) outcome(varname) Family(string) Link(string)]
[bootstrap]
```

Description: t_dr_bounded.ado

The variable treatvar() is the treatment variable (A) and the outcome() must contain the outcome variable (Y). The pvar varlist contains the L variables for the propensity score (step 1), and the ovar varlist includes the independent variables of the outcome model (step 2). Generalized linear models are used to estimate the outcome model. The family and the link option must be selected. Three family options are available: Gaussian (default), binomial and Poisson. The link function has different alternatives in the list (linkname); the default link function is the canonical link that each family specified. For example the link function for the Gaussian family is the identity one. The bootstrap option outputs estimates of the standard error of the different causal contrasts estimates by using the bootstrap method. All STATA bootstrap command options function in the `t_dr_bounded.ado` program. (e.g. = reps()).

Example:

Command line: Continuous outcome

- ```
  t_dr_bounded,  treatvar(abcix) outcome(cardbill) ///
  ovars( stent  acutemi ejecfrac ves1proc) ///
  pvars(stent acutemi ejecfrac ves1proc p_inte*) bootstrap rep(200)
  ```

```
ATE - Doubly Robust bounded Estimator
Bootstrapped standard errors
```

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z
818.475	1007.705	-1168.677	2805.626	0.812
```

Command line: Binary outcome

- `t_dr_bounded, treatvar(abcix) outcome(death) ovars(stent /// acutemi ejecfrac ves1proc) pvars(stent acutemi ejecfrac /// ves1proc p_inte*) family(bin) link(logit) bootstrap rep(200)`

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z
-0.060	0.024	-0.106	-0.013	-2.544

CRR - Doubly Robust bounded Estimator
Bootstrapped standard errors

CRR	boots.Std.Err.	[95% Conf.	Interval(N)]	z
0.196	0.098	0.001	0.391	1.995

COR - Doubly Robust bounded Estimator
Bootstrapped standard errors

COR	boots.Std.Err.	[95% Conf.	Interval(N)]	z
0.184	0.096	-0.007	0.375	7.834

Summary of results:

The following tables summarize the results for the four methods described in this article.

Outcome variable: Carbill (Y is continuous)

Contrast measure **ATE**

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
1185.489	853.206	-496.996	2867.974	1.389	**Outcome Regression Adjustment**
ATE	Boots. sdt Err.	[95% Conf.	Interval(N)]	z	
714.966	1049.707	-1355.010	2784.943	0.681	**Inverse Probability Weighting**
ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
1140.773	880.266	-595.072	2876.619	1.296	**Stratification by the Propensity Score**
ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
818.475	1007.705	-1168.677	2805.626	0.812	**Double Robust Bounded**

Outcome variable: Death (Y is dicothomic)

Contrast measure **COR**

COR	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
0.195	0.092	0.014	0.376	10.110	**Outcome Regression Adjustment**
odds_rat	Boots.sdt Err.	[95% Conf.	Interval(N)]		
0.171	0.169	-0.162	0.504	1.011	**Inverse Probability Weighting**
COR	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
0.199	0.118	-0.033	0.431	1.695	**Stratification by the Propensity Score**
COR	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
0.184	0.096	-0.007	0.375	7.834	**Double Robust Bounded**

Contrast measure **ATE**

ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
-0.056	0.019	-0.094	-0.018	-2.884	**Outcome Regression Adjustment**
ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
-0.048	0.017	-0.081	-0.015	-2.837	**Stratification by the Propensity Score**
ATE	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
-0.060	0.024	-0.106	-0.013	-2.544	**Double Robust Bounded**

Contrast measure **CRR**

CRR	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
0.207	0.110	-0.011	0.425	1.875	**Outcome Regression Adjustment**
CRR	boots.Std.Err.	[95% Conf.	Interval(N)]	z	
0.209	0.105	0.002	0.417	1.989	**Stratification by the Propensity Score**
CRR	boots.Std.Err.	[95% Conf.	Interval(N)]		
0.196	0.098	0.001	0.391	1.995	**Double Robust Bounded**

References:

D' Agostino RB. (1998). Propensity score methods for bias reduction in the comparison of treatment to a non-randomized control group. *Statistics in Medicine* 17: 2265-2281

Lunceford, JK. and Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine* 23, 2937-2960.

Hernan, M. and Robins, J. (2006). Estimating causal effects from epidemiologic data. *J. Epidemiol. Community Health* 60: 578-586.

Kereiakes DJ., Obenchain RL. ,Barber BL. et al (2000) . Abciximab provides cost effective survival advantagein high volume interventional practice. *Am Heart J* 140, 603-610.

Robins, JM. (1986). A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7:1393-1512.

Robins, J. and Rotnitzky, A. (2001). Comment on "Inference for semiparametric models: some questions and an answer" by Bickel and Kwon. *Statistica Sinica* 11:920-36.

Rosenbaum, PR. and Rubin, DB. (1983). The Central role of the propensity score in observational studies for causal effects. *Biometrica* 70: 41-55.

Rosenbaum, PR. and Rubin, DB. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*. 79: 526-524

Rubin, DB. (1978). "Bayesian inference for causal effects: The Role of Randomization", *Annals of Statistics*, 6:34-58.