

**A SAS macro to estimate Average Treatment Effects with Propensity Score
Matching**

Nicolas Moreau¹

<http://cemoi.univ-reunion.fr>

**Centre d'Economie et de Management de l'Océan Indien
Université de La Réunion**

October 2016

Abstract

This paper presents a SAS macro to estimate the Average Treatment Effect (ATE) and the Average Treatment Effect for the Treated (ATET) based on propensity score with nearest neighbor matching. The robust standard errors derived in Abadie and Imbens (2016) are computed.

JEL : C210

¹ E-mail: nicolas.moreau@univ-reunion.fr

Introduction

In this paper we present the SAS macro *ps_matching*. *ps_matching* estimates the average treatment effect (ATE) and the average treatment effect for the treated (ATET) based on propensity scores with nearest neighbor matching with replacement. In this respect, we draw heavily on Abadie and Imbens (2016). We refer the reader to this article for a clear and comprehensive presentation of the propensity score matching estimator.

Following these authors, *ps_matching* provides adjusted standard errors to account for the fact that the propensity score is estimated in a first step.

The source code is available at <http://cemoi.univ-reunion.fr>.

Syntax of *ps_matching*

The syntax is `%ps_matching(data=,y=,w=,x=,M=,Link=,L=,Lt=)`;

where *data* specifies the data set, *y* the outcome variable, *w* the binary variable treatment indicator, *x* the list of covariates to be used in the matching, and *M* the number of matches to be made per observation. *M* could be any integer between 1 and the minimum of the number of treated units and controls in the sample.

If there are ties and if different matched pairs (i,j) and (i,l) lead to the same distance $d_{ij} = d_{il}$, then the number of matches per unit is greater than *M*.

If *Link*=probit, a probit estimation of the propensity score is used instead of the default logit model when *Link* is not specified by the user.

L is a small positive integer (typically, $L=2$) needed to compute the variance estimator of ATE (see Abadie and Imbens, 2016), *Lt* being another small positive integer (typically, $Lt=1$) needed to compute the variance estimator of ATET (see again Abadie and Imbens, 2016).

Note that all variables in *y*, *x* and *w* must be numeric.

Results presentation and output data files

ATE and ATET are automatically computed.

The first two output tables summarize the model specification and estimation options. The third table provides summary statistics, such as the number of treated units, the number of controls matched to treated units, and so on. The fourth and final table shows the main results. The “Estimate” column reports the estimated ATE and ATET. The next column shows the corresponding robust standard errors. “z” corresponds to the z-statistics to test whether ATE and ATET are 0; these are computed as the estimated parameters divided by their corresponding standard error. The “P-value” column reports the p-values for the z-statistics for a two-sided test. The last two columns show the lower and upper bounds of the 95% confidence interval for the z-statistics.

To evaluate the validity of the common support assumption, a box-plot is provided. To assess the validity of the balancing hypothesis, *ps_matching* performs, for each quintile of the estimated propensity score, normalized differences (see Austin, 2009), box-plots and empirical distribution functions to compare the covariate distributions between treatment groups before and after

matching. All these different plots are supplied for continuous variables alone. For binary variables, contingency tables are provided.

In *ps_matching*, all covariates that are not binary are considered continuous.

Two temporary output data files are created, which need to be stored in a specific folder with a *libname* statement to become permanent.

Outdata1 includes an internal identification number for observation *i* created by the program that is based on the original sort order and called *_id_*. It also includes the outcome variable *y*, the covariates *x* and the treatment group indicator *w*. *cardJMi* specifies the number of matches for unit *i*. *count* is the number of times unit *i* is used as a match. *km_i* specifies the number of times unit *i* is used as a matched for any observation *j* of the opposite treatment group weighted by the total number of matches for the given observation *j*. *pscore_pred* is the estimated propensity score for observation *i*.

Outdata2 includes the list of indices for the *M* closest matches for unit *i*. *id* is the internal identification number for observation *i* and *idM* the corresponding identification number of *i*'s closest matches in the opposite treatment group. For each *id*, there is one row per match. For instance, if unit 3 is matched with units 5, 6, and 10, there are three rows in *outdata2* that correspond to *id* =3, the first with *idM*=5, the second with *idM*=6, and the last with *idM*=10. For each matched pair (*i,j*), the score distance (as an absolute value) between unit *i* and unit *j* of the opposite treatment group is stored so as unit *j*'s outcome value.

An example

We will illustrate the use of *ps_matching* by using data from a study of the effect of a mother's smoking status during pregnancy on infant birth weight as reported by Cattaneo (2010) and used in Stata Treatment-Effects Reference Manual, release 14.

bweight: infant birth weight
dummy_smoke= 1 if mother smoked during pregnancy, 0 otherwise
dummy_married= 1 if married, 0 otherwise
mage: mother's age
mage2: squared mother's age
medu: mother's education in years
dummy_fbaby=1 if mother's first birth, 0 otherwise.

The instructions `%ps_matching(data=cattaneo2,y=bweight,w=dummy_smoke,x=dummy_married mage mage2 medu dummy_fbaby,M=1,Link=,L=2,Lt=1);`

give the following results:

ESTIMATING AVERAGE TREATMENT EFFECTS with PROPENSITY SCORE MATCHING

ModelSpecification

Outcome variable: bweight

Binary treatment: dummy_smoke

Matching variables: dummy_married mage mage2 medu dummy_fbaby

EstimationOptions

Number of matches requested (M): 1
Number of matches requested in the same treatment group (L and Lt): 2 and 1

SummaryStatistics

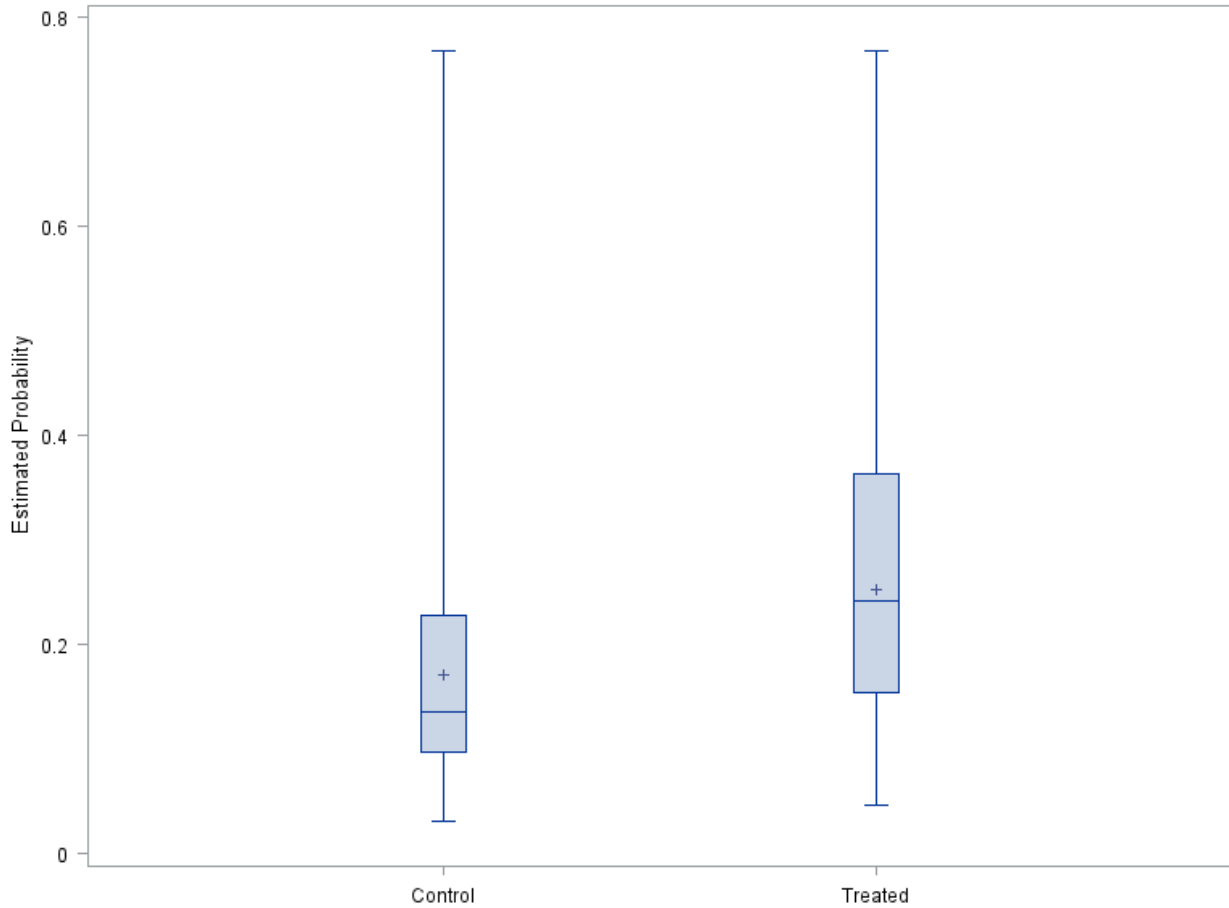
Number of observations: 4642
Number of control units: 3778
Number of treated units: 864
Number of treated units matched to controls: 813
Number of control units matched to treated: 2804
Number of times a treated unit is used as a match (MIN): 1
Number of times a treated unit is used as a match (MAX): 75
Number of times a control unit is used as a match (MIN): 1
Number of times a control unit is used as a match (MAX): 19

Estimation results with Abadie and Imbens (2016) robust standard errors

Results

	Estimate	Std.Error	z	P-value	L. bound 95% CI	U. bound 95% CI
Average Treatment Effect (ATE)	-210.952	33.047	-6.383	0	-275.724	-146.18
Average Treatment Effect for the Treated (ATET)	-237.391	26.625	-8.916	0	-289.576	-185.206

Common support: visual diagnostic



Normalized covariate mean differences between treated and controls

	Unmatched Samples	Matched samples (T)	Matched samples (C)
DUMMY_MARRIED	-0.596	-0.514	-0.579
MAGE	-0.3	-0.136	-0.344
MAGE2	-0.303	-0.13	-0.348
MEDU	-0.547	-0.472	-0.518
DUMMY_FBABY	-0.166	-0.145	-0.154

Note: 'Matched samples (T)' is for normalized mean differences between all sample treated and their matches, 'Matched samples (C)' for normalized mean differences between all sample controls and their matches.

**Comparing binary covariate distributions between treatment groups before and after matching
Interval 1 of the estimated propensity score**

		SampleType			
		All treated	All controls	Matched controls	Matched treated
dummy_married					
0	Distribution in %	.	0.12	.	.
1	Distribution in %	100.00	99.88	100.00	100.00
dummy_fbaby					
0	Distribution in %	26.09	30.81	26.09	23.69
1	Distribution in %	73.91	69.19	73.91	76.31

**Comparing binary covariate distributions between treatment groups before and after matching
Interval 2 of the estimated propensity score**

		SampleType			
		All treated	All controls	Matched controls	Matched treated
dummy_married					
0	Distribution in %	.	0.24	.	.
1	Distribution in %	100.00	99.76	100.00	100.00
dummy_fbaby					
0	Distribution in %	22.55	34.83	22.77	31.73
1	Distribution in %	77.45	65.17	77.23	68.27

**Comparing binary covariate distributions between treatment groups before and after matching
Interval 3 of the estimated propensity score**

		SampleType			
		All treated	All controls	Matched controls	Matched treated
dummy_married					
0	Distribution in %	3.33	2.27	2.82	1.11
1	Distribution in %	96.67	97.73	97.18	98.89
dummy_fbaby					
0	Distribution in %	86.67	94.53	89.44	96.35
1	Distribution in %	13.33	5.47	10.56	3.65

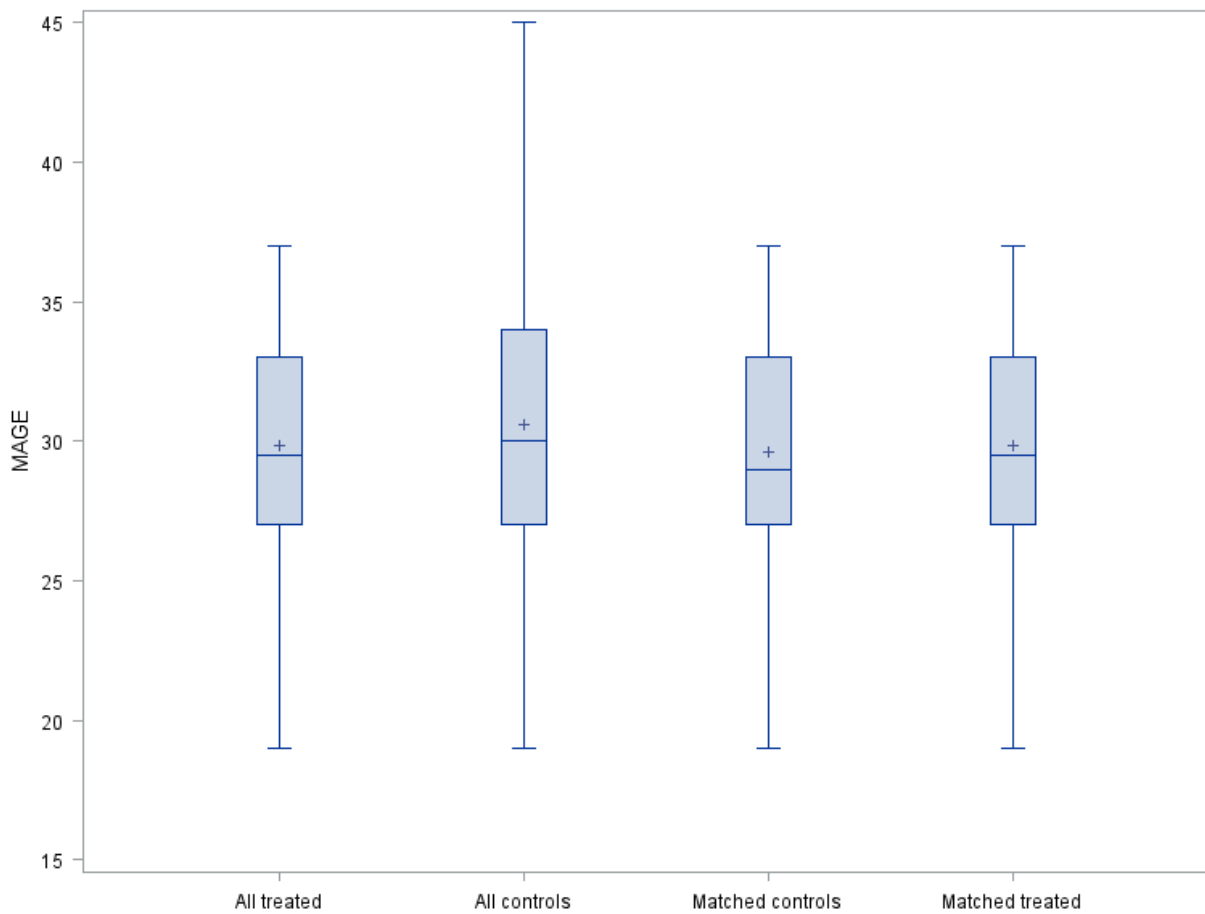
**Comparing binary covariate distributions between treatment groups before and after matching
Interval 4 of the estimated propensity score**

		SampleType			
		All treated	All controls	Matched controls	Matched treated
dummy_married					
0	Distribution in %	50.44	47.89	51.39	48.20
1	Distribution in %	49.56	52.11	48.61	51.80
dummy_fbaby					
0	Distribution in %	49.56	53.48	48.61	50.65
1	Distribution in %	50.44	46.52	51.39	49.35

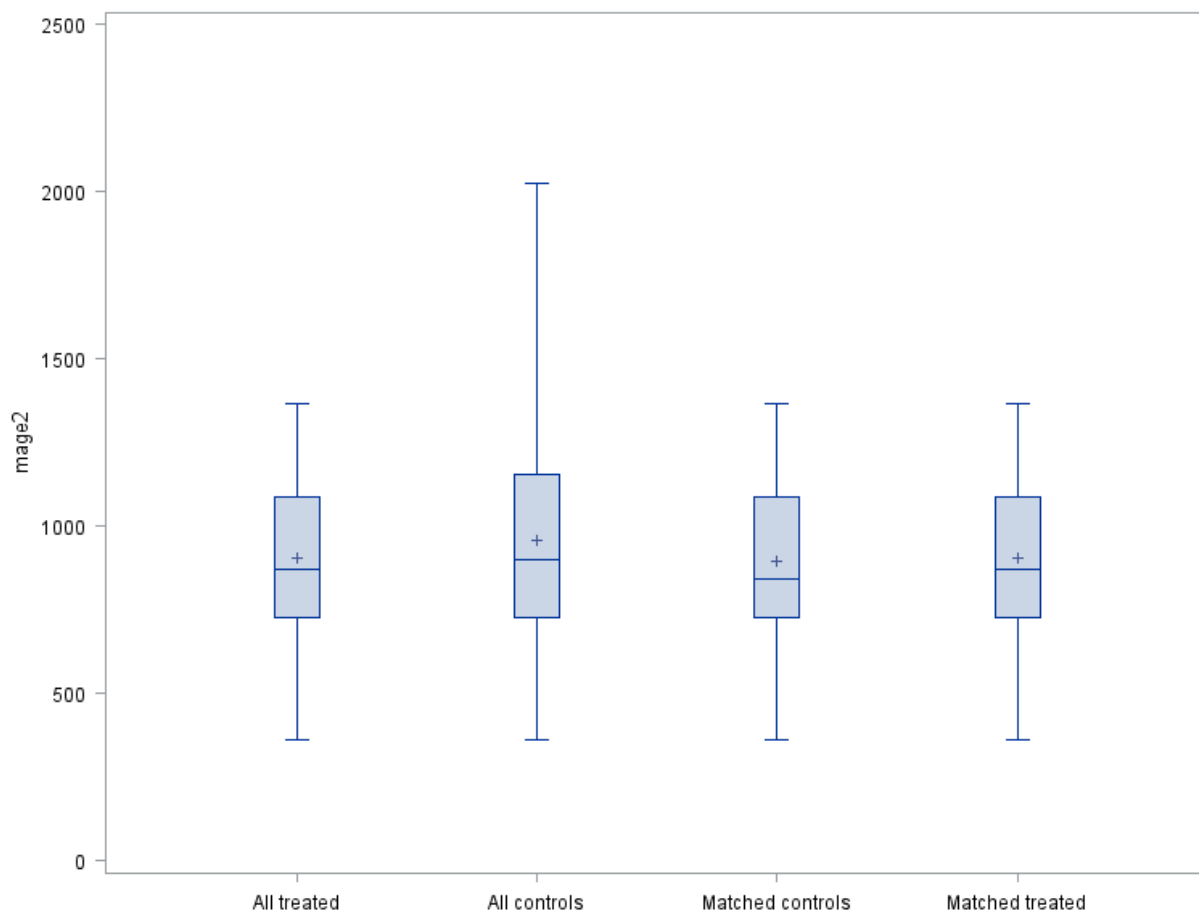
**Comparing binary covariate distributions between treatment groups before and after matching
Interval 5 of the estimated propensity score**

		SampleType			
		All treated	All controls	Matched controls	Matched treated
dummy_married	0 Distribution in %	99.11	96.11	99.68	99.59
	1 Distribution in %	0.89	3.89	0.32	0.41
dummy_fbaby	0 Distribution in %	78.40	68.70	77.60	68.09
	1 Distribution in %	21.60	31.30	22.40	31.91

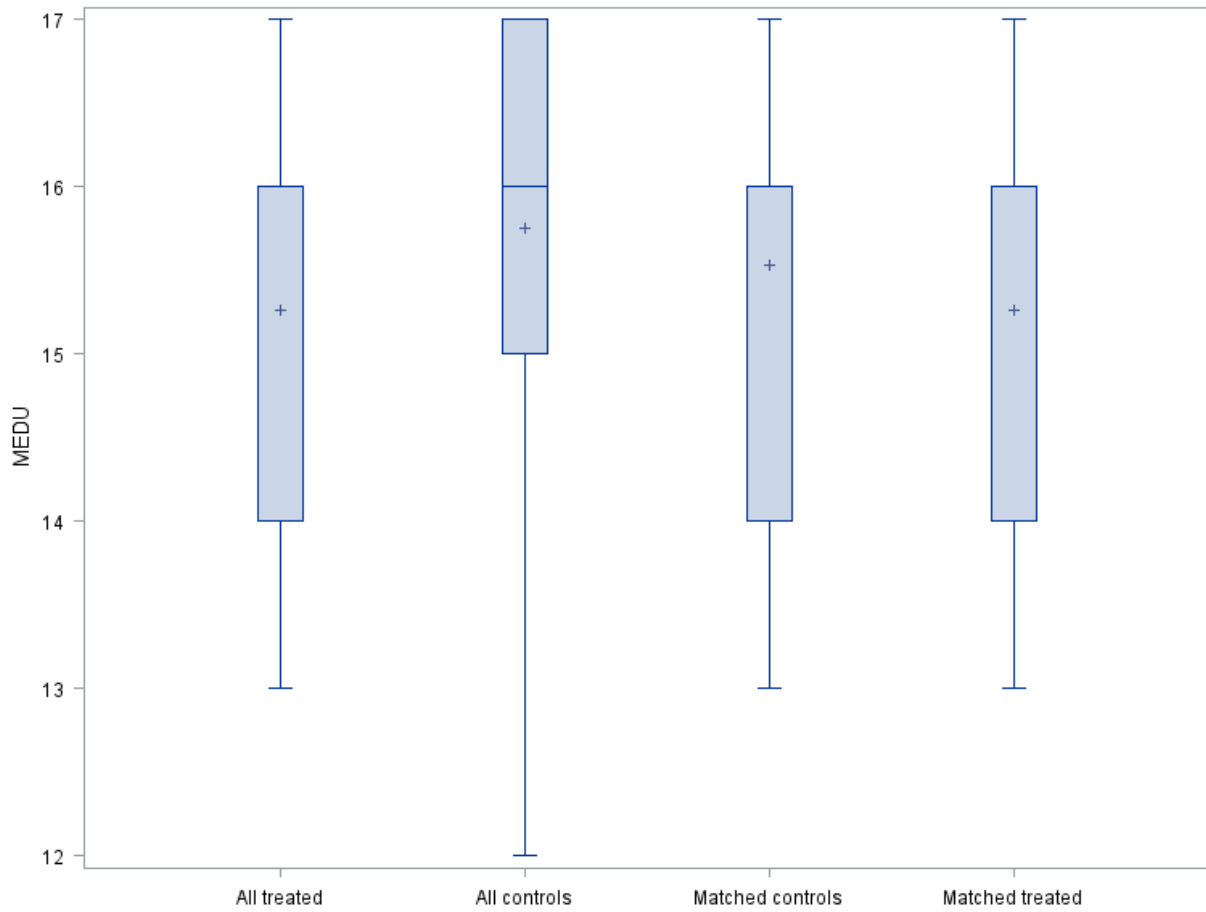
Checking balance through box-plots, interval 1 of the estimated propensity score distribution



Checking balance through box-plots, interval 1 of the estimated propensity score distribution

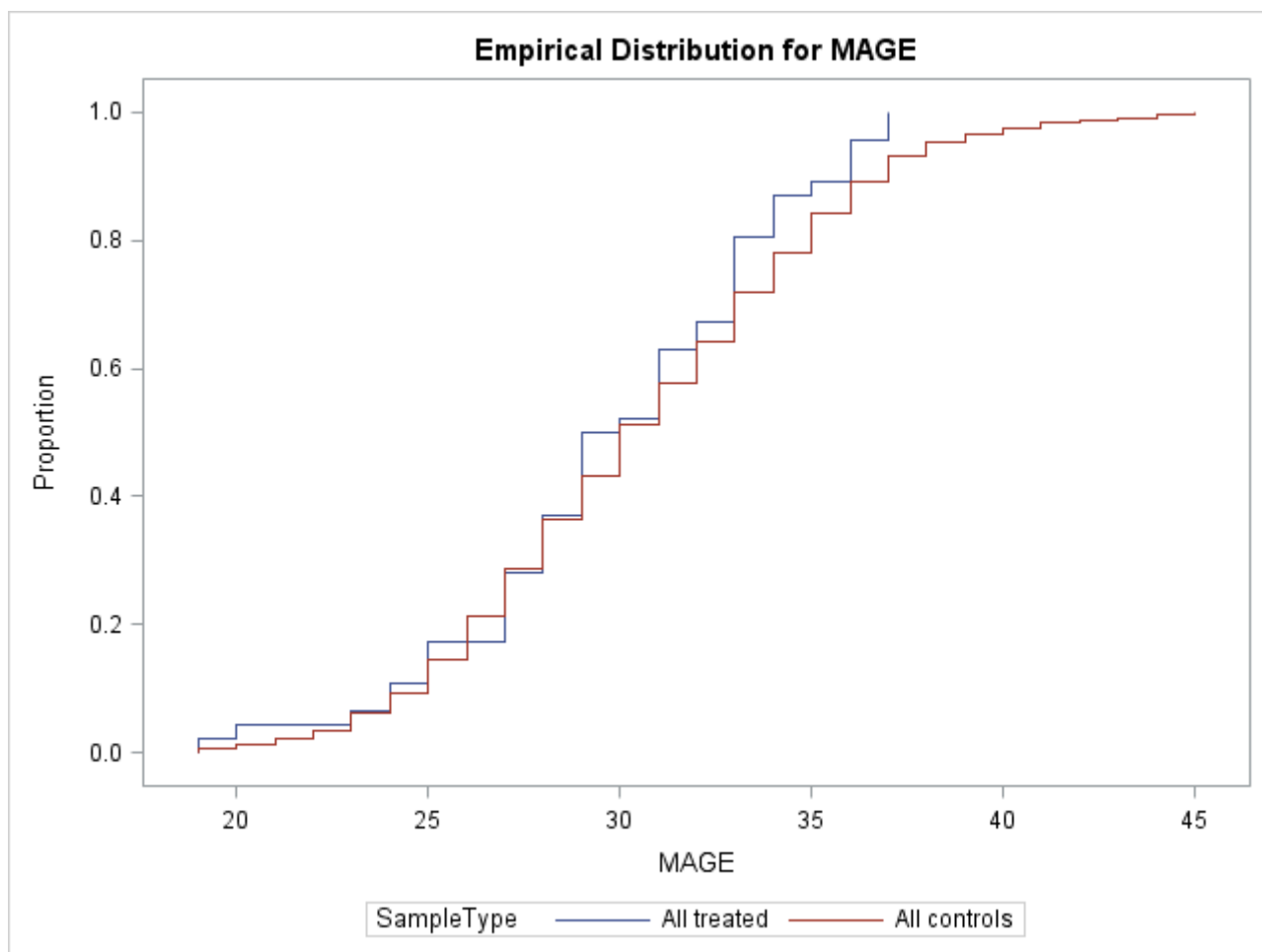


Checking balance through box-plots, interval 1 of the estimated propensity score distribution

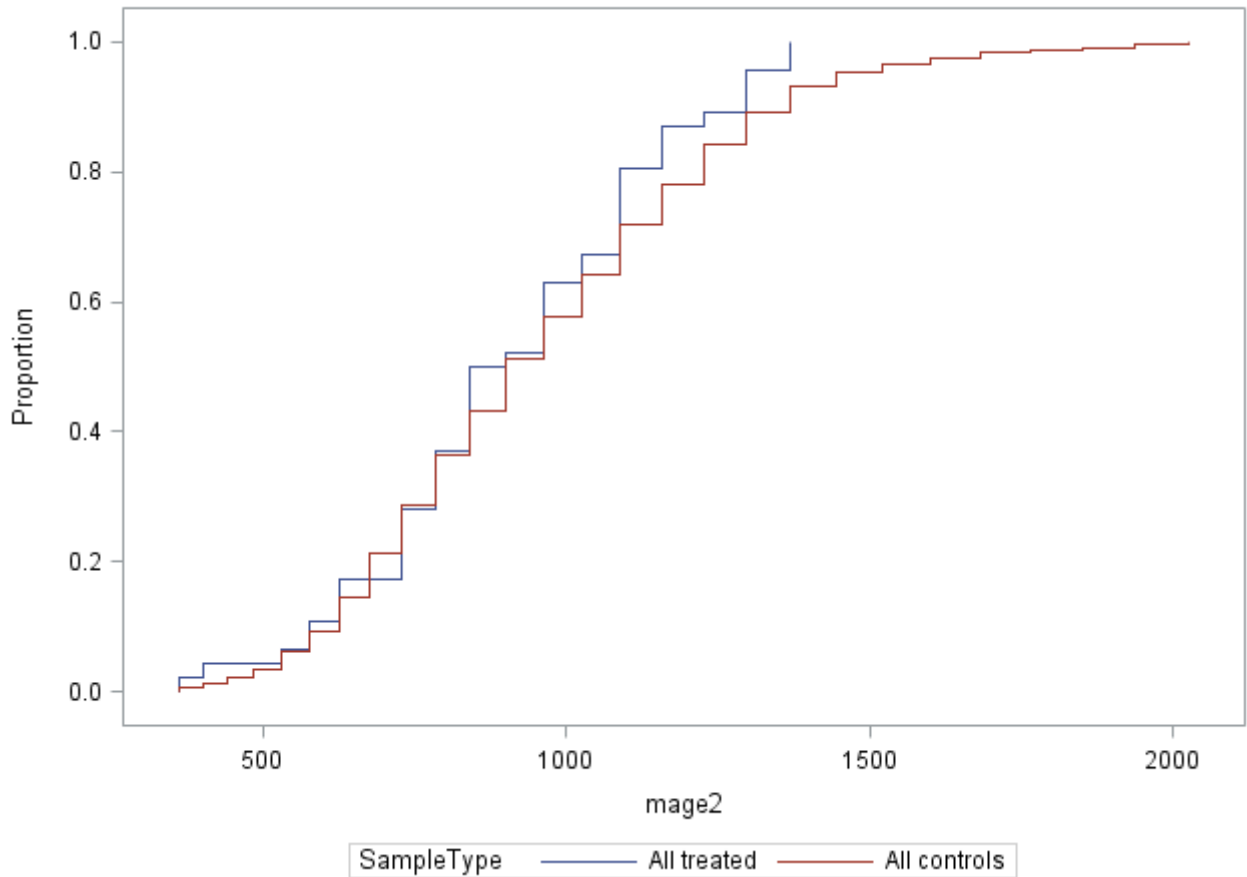


Checking balance through empirical distribution functions, interval 1 of the estimated propensity score distribution

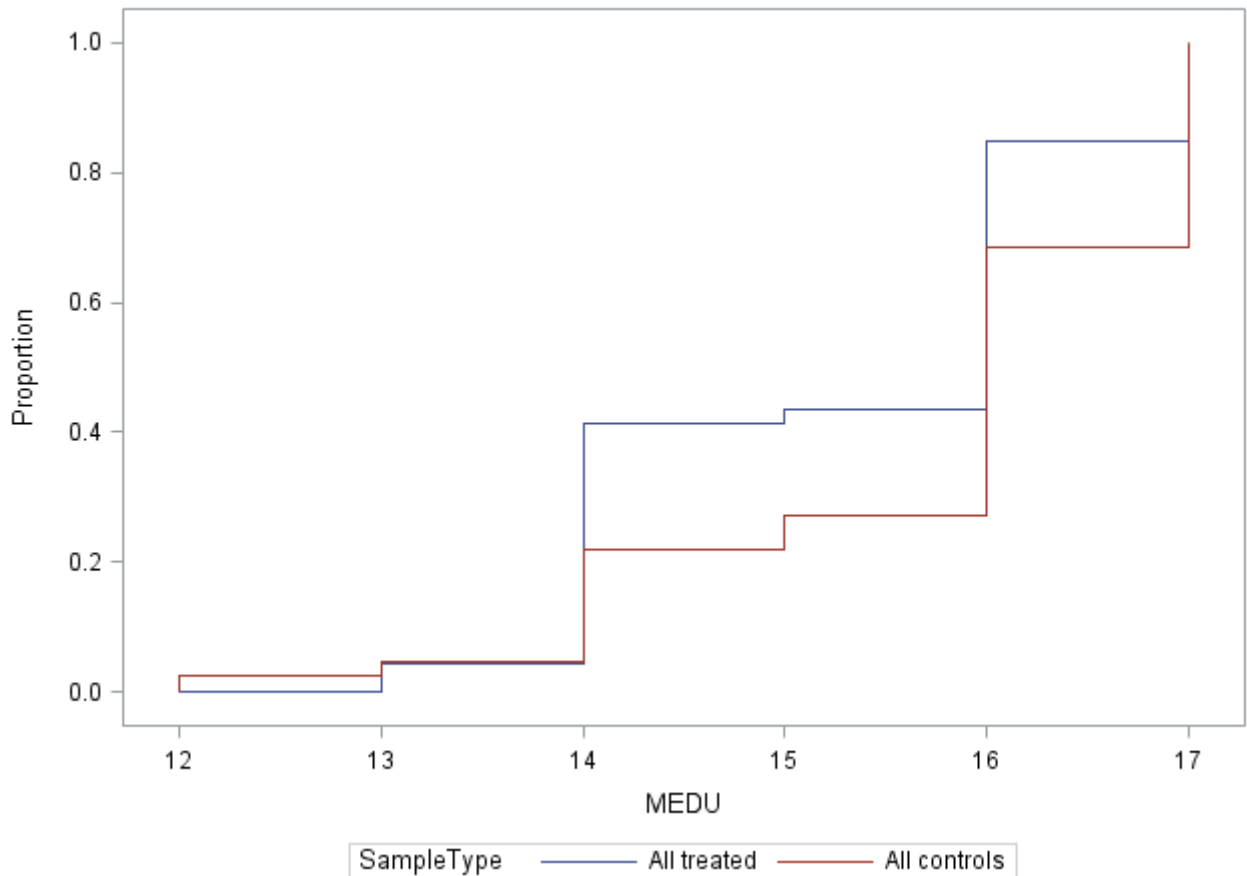
Original sample of treated vs original sample of controls



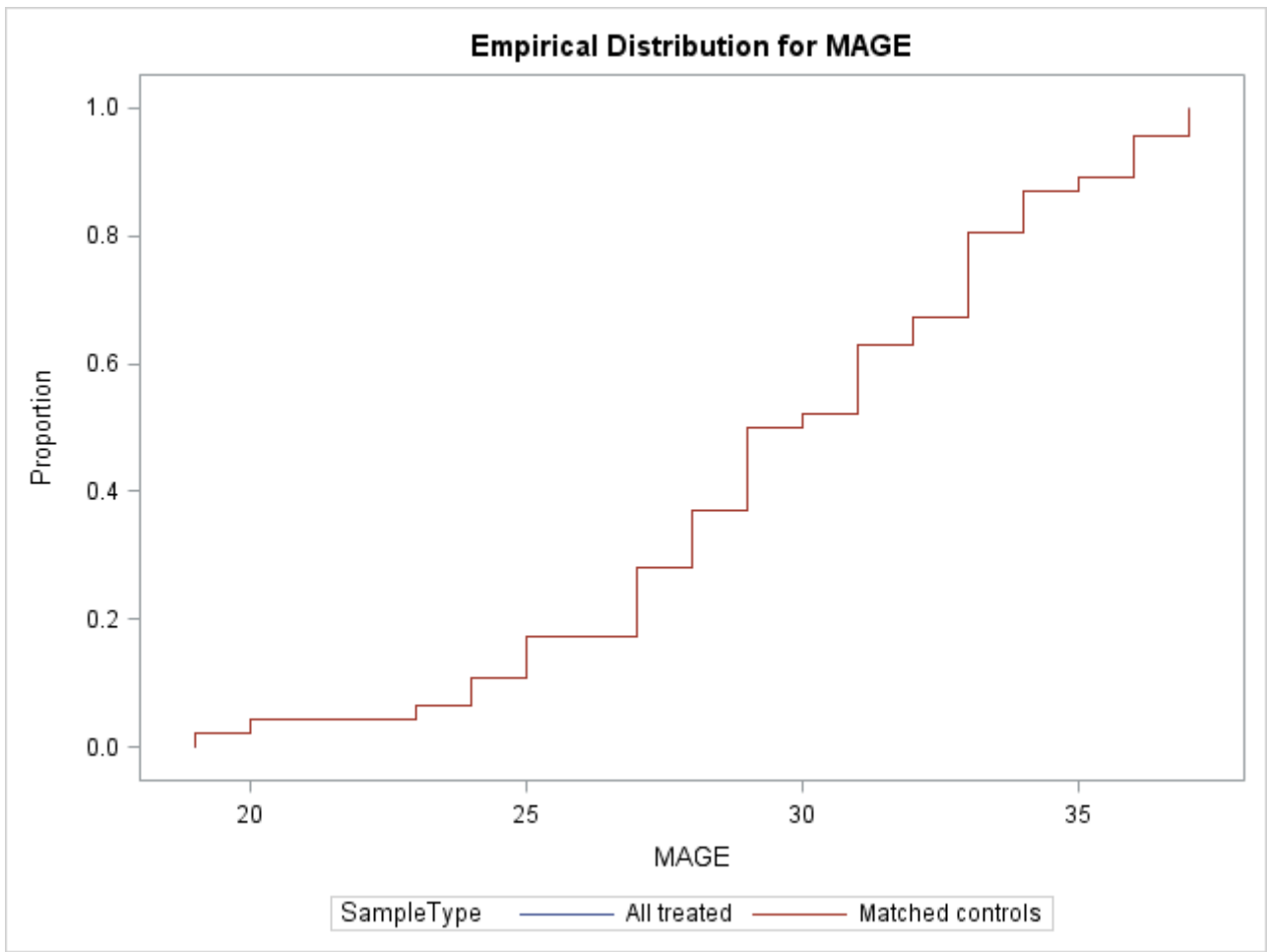
Empirical Distribution for mage2



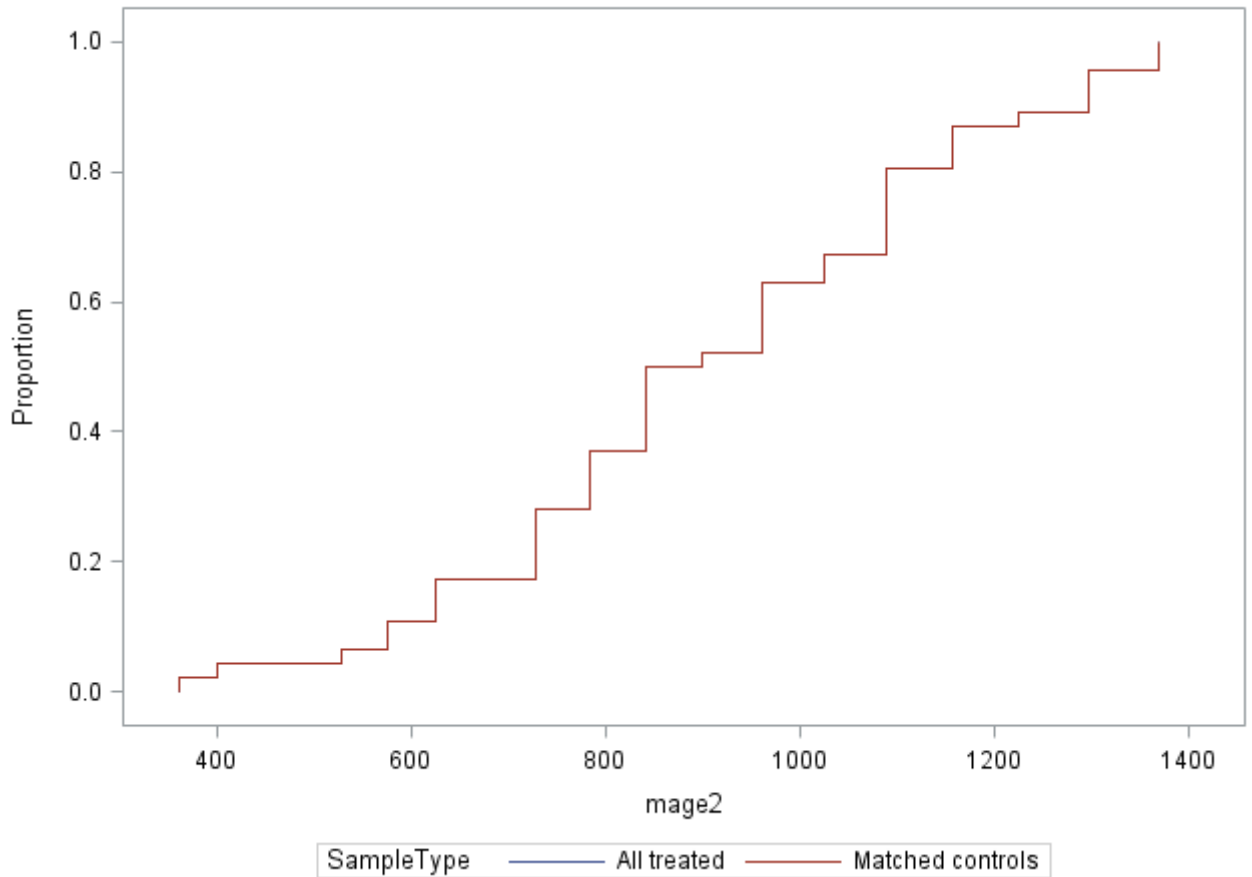
Empirical Distribution for MEDU



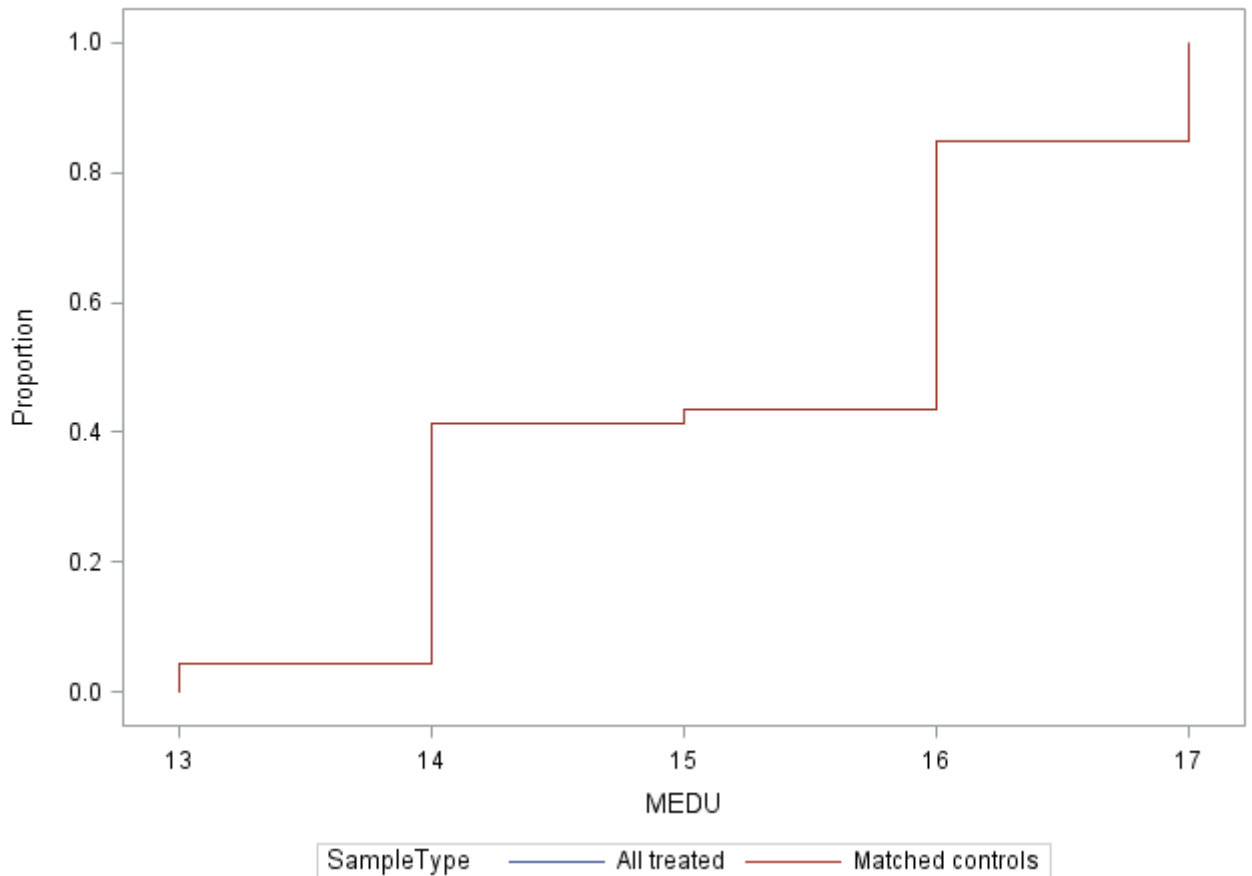
Checking balance through empirical distribution functions, interval 1 of the estimated propensity score distribution
Original sample of treated vs final sample of matched controls



Empirical Distribution for mage2

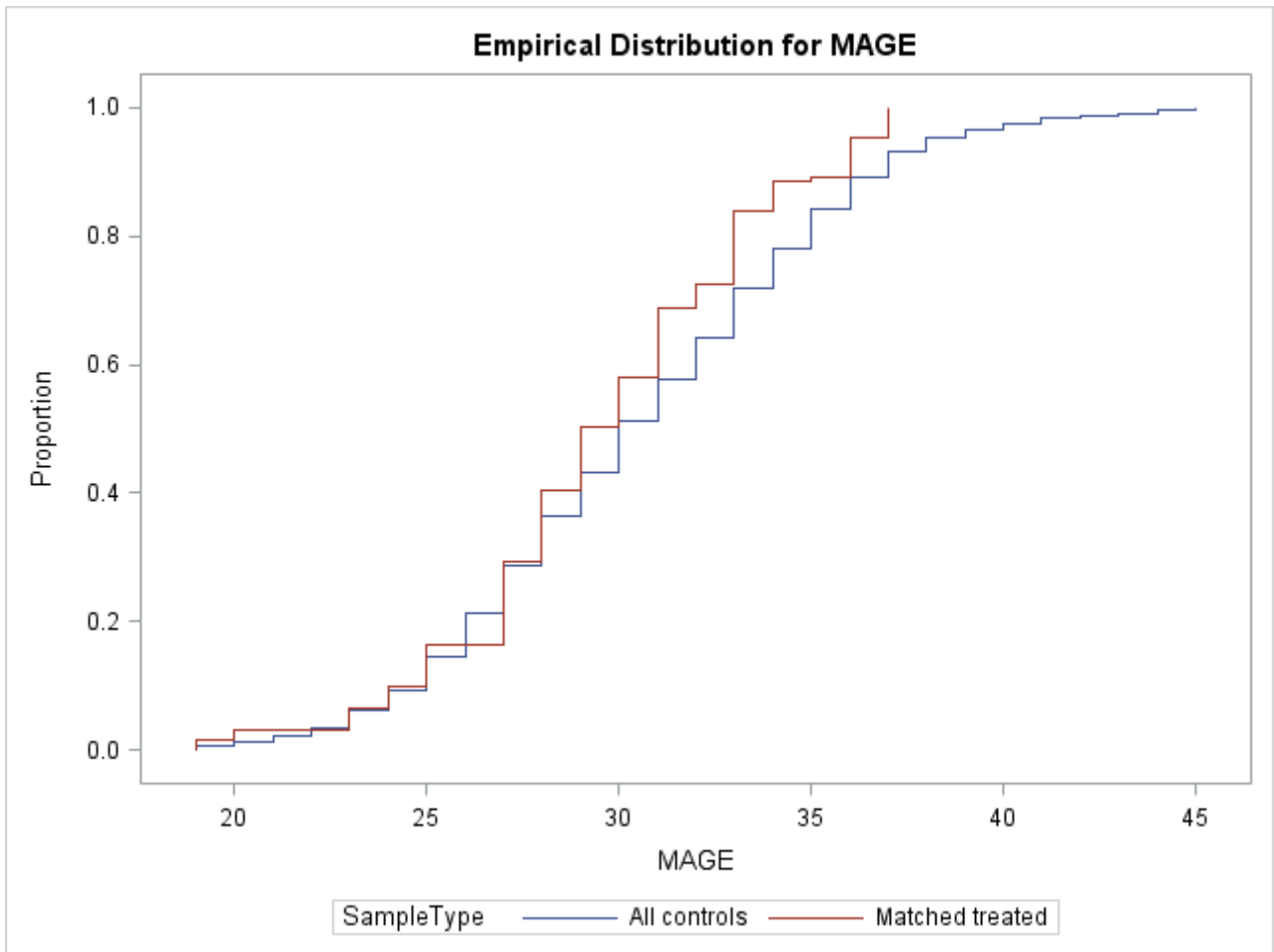


Empirical Distribution for MEDU

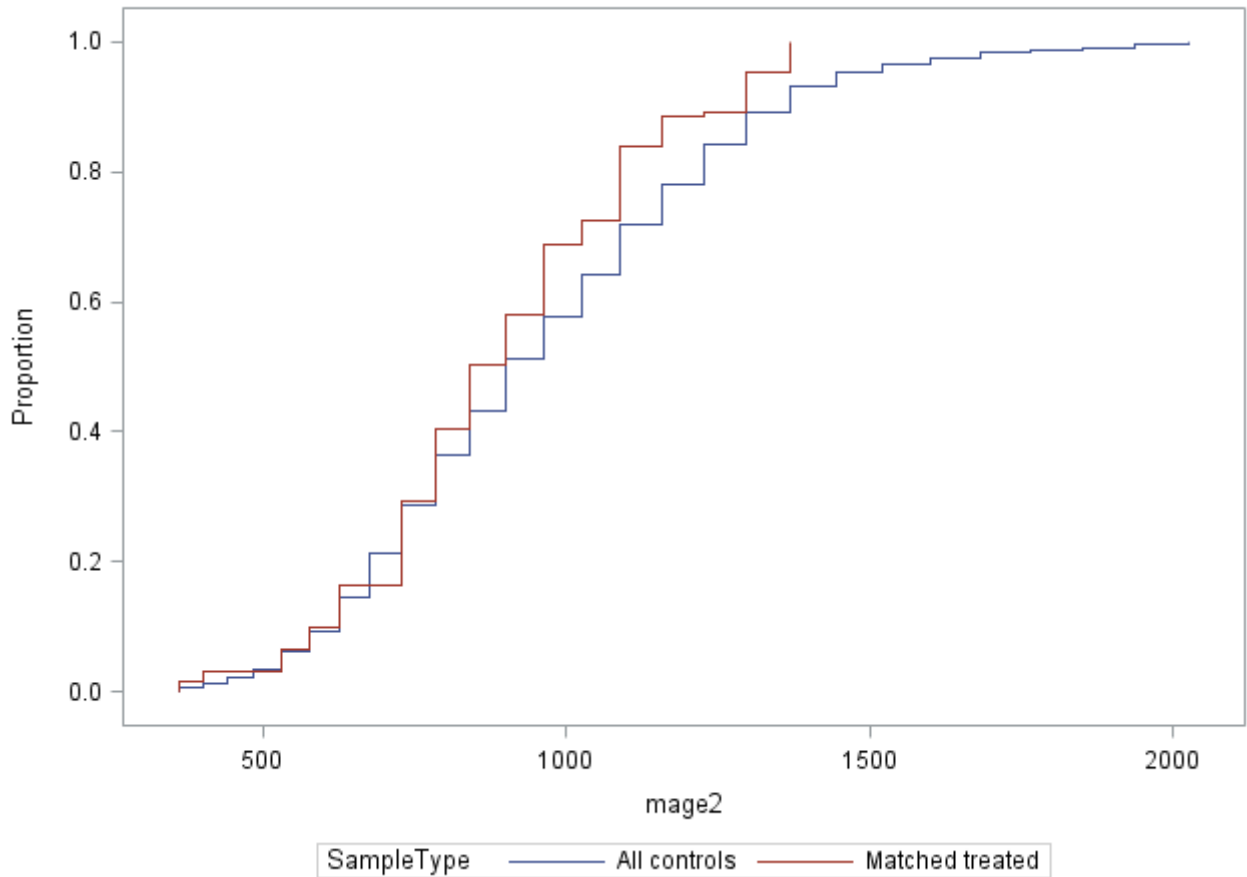


Checking balance through empirical distribution functions, 1th interval of the estimated propensity score distribution

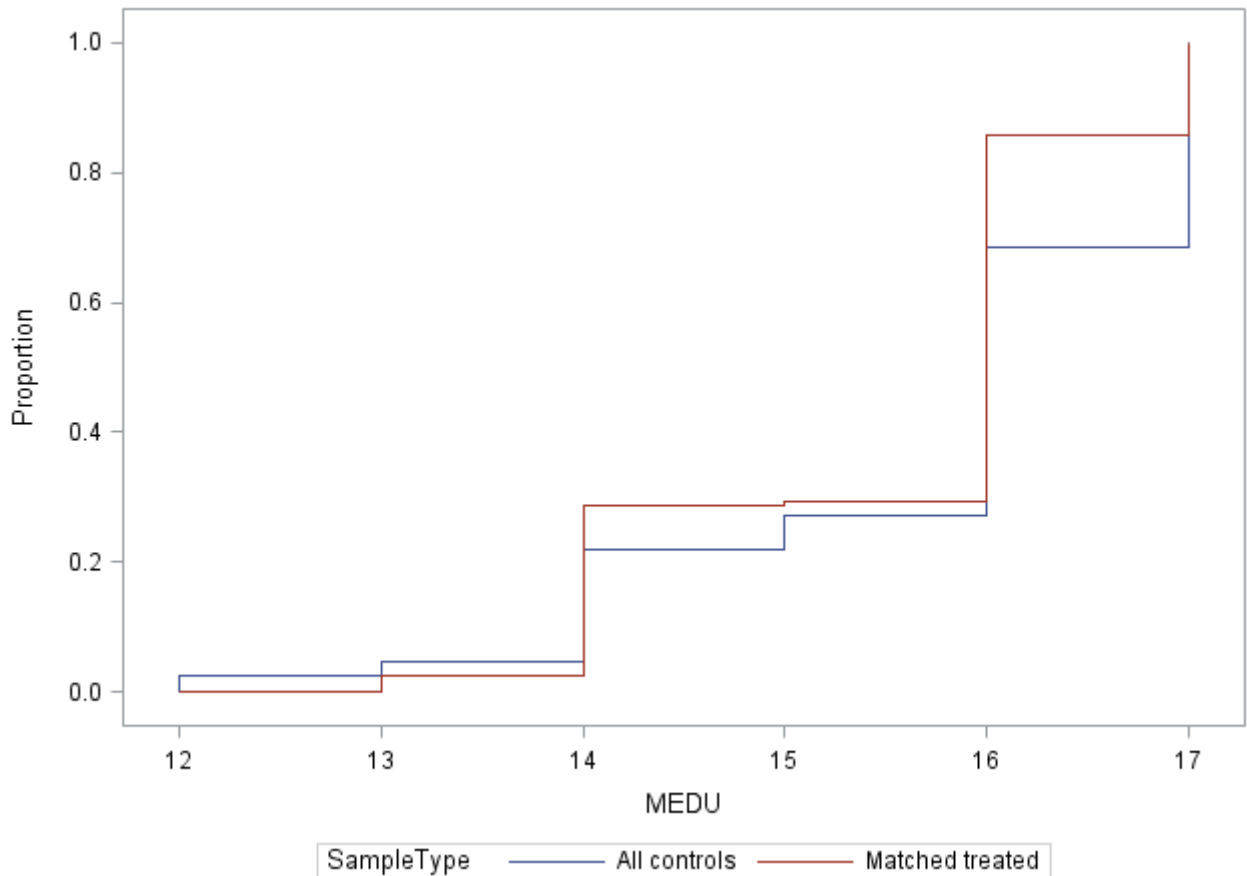
Original sample of controls vs final sample of matched treated



Empirical Distribution for mage2



Empirical Distribution for MEDU



and so on for intervals 2 to 5 of the estimated propensity score distribution.

References

Abadie A., Drukker D., Leber Herr J. and Imbens G. W. (2004), "Implementing matching estimators for average treatment effects in Stata", *The Stata Journal*, vol. 4, n°3: 290-311.

Abadie A, and Imbens G. W. (2016), "Matching on the estimated propensity score", *Econometrica*, Vol. 84, n°2: 781-807.

Austin P. C. (2009), "Balance diagnostics for comparing the distribution of baseline covariates between treatment group in propensity-score matched samples", *Statistics in Medicine*, vol. 28: 3083–3107.

Cattaneo M. D. (2010), "Efficient semiparametric estimation of multi-valued treatment effects under ignorability", *Journal of Econometrics*, Vol. 155: 138-154.

Stata Treatment-Effects Reference Manual, A Stata Press Publication, Release 14.