This report is part of the RAND Corporation tool series. RAND tools may include models, databases, calculators, computer code, GIS mapping tools, practitioner guidelines, web applications, and various toolkits. All RAND tools undergo rigorous peer review to ensure both high data standards and appropriate methodology in keeping with RAND's commitment to quality and objectivity.

Propensity scores for multiple treatments: A tutorial for the `mnps` macro in the `twang` SAS macros

Dan McCaffrey, Lane Burgette, Beth Ann Griffin and Craig Martin *
RAND Corporation

June 16, 2016

1 Introduction

The Toolkit for Weighting and Analysis of Nonequivalent Groups, `twang`, was designed to make causal estimates in the binary treatment setting. The package was developed in the R statistical computing and graphics environment and ported to SAS through a family of macros available at `http://www.rand.org/statistics/twang/downloads.html`.

The `twang` package in R now also functions to handle more than two treatment conditions through the `mnps` function, which stands for multinomial propensity scores. McCaffrey et al. (2013) describe the methodology behind the `mnps` function; the purpose of this document is to describe the syntax and features related to the implementation of the `mnps` R function through macros in SAS.

At a high level, the `mnps` function decomposes the propensity score estimation into several applications of the `ps` function, which was designed for the standard dichotomous treatment setting and which can be accessed via the `%ps` macro in SAS. For this reason, users who are new to `twang` are encouraged to learn about the `ps` function and `%ps` macro before using the `mnps` function and macros. A tutorial describing the use of `twang` macros for comparing two treatments is found at `http://www.rand.org/statistics/twang/sas-tutorial.html` or in pdf at `http://www.rand.org/content/dam/rand/www/external/statistics/twang/tl136_tutorial.pdf`. The information in that tutorial, including directions on installing R and setting the macro parameters to interface with it, will not be repeated here.

2 An ATE example

To demonstrate how to implement an analysis using the `%mnps` macro in SAS, we use a random subset of the data described in McCaffrey et al. (2013). This truncated dataset is called `AOD.sas7dbat`. The macros and data can be downloaded from `http://www.rand.org/statistics/twang/downloads.html`. This example study includes three treatment groups, and the data include records for 200 youths in each treatment group of an alcohol and other drug treatment evaluation. To prepare for using the example data and the `twang` macros, we first set a `libname` to the folder that contains the example data set `aod.sas7bdat` and `%include` the file with the macros. For the tutorial we assume that the data are in a folder named "sasdata". Users will need to provide the correct folder (including the directory path as necessary). We also assume

*The development of this software and tutorial was funded by National Institute of Drug Abuse grants number 1R01DA015697 (PI: McCaffrey) and 1R01DA034065 (PIs: Griffin/McCaffrey).

that the downloaded macros are accessible through a call to the file "twang_mac_v3.1.0.sas" so that the current version of the macros is version 3.1.0. Users will need to use the appropriate path to the file where they have stored it. They should also be sure to use the most up-to-date version of the macros. Version 3.1.0 was current at the time this tutorial was written. Users SAS code for implementing this example would need to include the following lines before either the data or macros are referenced.[1]

```
libname sasin "sasdata";
%include "twang_mac_v3.1.0.sas";
```

For the `AOD` dataset, the variable `treat` contains the treatment indicators, which have possible values `community`, `metcbt5`, and `scy`. The other variables included in the dataset are:

- `suf12`: outcome variable, substance use frequency at 12 month follow-up
- `illact`: pretreatment covariate, illicit activities scale
- `crimjust`: pretreatment covariate, criminal justice involvement
- `subprob`: pretreatment covariate, substance use problem scale
- `subdep`: pretreatment covariate, substance use dependence scale
- `white`: pretreatment covariate, indicator for non-Hispanic white youth

In such an observational study, there are several quantities that one may be interested in estimating. The estimands that are most commonly of interest are the average treatment effect on the population (ATE) and the average treatment effect on the treated (ATT). The differences between these quantities are explained at length in McCaffrey et al. (2013), but in brief the ATE answers the question of how, on average, the outcome of interest would change if everyone in the population of interest had been assigned to a particular treatment relative to if they had all received another single treatment. The ATT answers the question of how the average outcome would change if everyone who received one particular treatment had instead received another treatment. We first demonstrate the use of `mnps` when ATE is the effect of interest and then turn to using the function to support estimation of ATT.

2.1 Estimating the weights

The following is an example of the command for running `%mnps` to obtain propensity score weights for three or more treatment groups:

```
%mnps(treatvar= treat,
      vars = illact crimjust subprob subdep white,
      dataset = sasin.AOD,
      estimand = ATE,
      stopmethod = es.mean ks.mean,
      ntrees = 3000,
      output_dataset=sasin.aodwgt,
      return_ps=TRUE,
      plotname=mnps_example_plot.pdf,
      Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
      objpath=C:\Users\uname\twang_mnps_example)
```

[1] Code used in this tutorial can be found in stand alone text file at `http://www.rand.org/statistics/twang/downloads.html/mnps_turorial_code.sas`.

The main arguments for the %mnps macro are:

- treatvar – This argument identifies the name of the treatment variable. In the AOD data set, the treatment variable is labeled "treat".

  The variable specified by treatvar must take on at least three values indicating three or more treatment groups. If your study involves only two treatment conditions then the %ps macro must be used instead of %mnps. As is typical in SAS the variable name specified in treatvar is not case sensitive but any references to the values of this variable are.

- vars – This argument specifies the pretreatment variables to be used for controlling for differences among the multiple treatment groups.

- dataset – This argument simply tells the macro the name of the dataset that contains the variables for the propensity score estimation.

- estimand – This argument specifies the type of causal effects to be estimated with the weights and it can either be "ATT" or "ATE".

In addition to these three arguments, there are a number of arguments related to fine-tuning the generalized boosted model GBM upon which the twang methods are built. In particular, GBMs used by the twang methodology rely on a composition of tree-based regression models that are built in an iterative fashion. As the iterations or number of regression trees added to the model increases, the model becomes more complex. However, at some point, more complex models typically result in worse balance on the pretreatment variables and therefore are less useful in a propensity score weighting context (Burgette, McCaffrey and Griffin, In Press). The ntrees argument controls the maximum number of iterations. The fitting algorithm will select the number of trees or iterations, less than or equal to the maximum set by ntrees, which provides the best balance according to the balance criteria specified by the user.

The balance criteria used to tune the propensity score model are specified in the stopmethod argument. As with the %ps macro, four stopping rule balance criteria are available for the %mnps macro. They are "es.mean", "es.max", "ks.mean", and "ks.max". The four stopping rules are defined by two components: a balance metric for each covariate and a rule for summarizing across covariates. A balance metric summarizes the difference between two univariate distributions of a single pretreatment variable (e.g., illicit activities scale). The stopping rules in twang use two balance metrics: absolute standardized mean difference (ASMD; also referred to as the absolute standardized bias or the effect size (ES)) and the Kolmogorov-Smirnov (KS) statistic. The stopping rules use two different rules for summarizing across covariates: the mean of the covariate balance metrics ("mean") or the maximum of the balance metrics ("max"). The first piece of the stopping rule name identifies the balance metric (ES or KS) and the second piece specifies the rule for summarizing across balance metrics. For instance, es.mean uses the effect size or ASMD and summarizes across variables with the mean and the ks.max uses the KS statistics to assess balances and summarizes using the maximum across variables. The other two stopping rules use the remaining two combinations of balance metrics and summary statistics. In this example, we chose to examine both es.mean and ks.mean.

The Rcmd argument specifies the R program executable file for running R. Details on specifying this parameter can be found in *Toolkit for Weighting and Analysis of Nonequivalent Groups: A tutorial for the twang SAS Macros*. The objpath argument specifies a directory for any output files created by the macro to be stored. It is not required but it can useful for capturing plots and intermediate files generated by the macro. The example uses a folder "C:\Users\uname\twang_mnps_example" which would need to be changed to name and path of the user's folder. Additional details on objpath argument are in *Toolkit for Weighting and Analysis of Nonequivalent Groups: A tutorial for the twang SAS Macros*.

After running the code above, the macro returns estimated propensity score weights along with diagnostic information. In the dataset of results, there is one propensity score weight per observation for each stopping rule specified by the `stopmethod` argument. The input data with the weights appended can be saved in specified temporary or permanent SAS dataset using the `output_dataset` argument. In this example, we set `output_dataset=sasin.aodwgt` to save the data in a permanent SAS dataset, aodwgt, in the same folder where input data were stored. We will use the weights later to estimate the treatment effect; saving the weights avoids the need to rerun the weight estimation routine at that time. The output data set contains the propensity score weights in variables named "es_mean_ATE" and "ks_mean_ATE" which are the weights for the corresponding stopping rules. The output data set also includes all the variables from the input dataset.

In addition, because the `return_ps` argument equals `TRUE`, the output dataset includes the estimated propensity scores. If `return_ps=FALSE`, the default, then the propensity scores are not returned. When the estimand is "ATE", as it is in this example, there is one set of propensity scores for each treatment level and each stopping rule. Hence, in this example, there are six propensity score variables: one for the each of the `es_mean` and `ks_mean` stopping rules for each of the three levels of treatment. The propensity score equals the probability that an individual with given values for a set of covariates is in the specified treatment group as opposed to any other treatment, when the estimand is "ATE". When the estimand is "ATT", unlike this example, the propensity score equals the probability that an individual with given values for a set of covariates is in the target treatment group rather than in the specified treatment group.[2] They are missing values for the other levels of treatment. The variable names for the propensity scores follow the same conventions as the names of the weights, but the propensity scores have the string "ps_" and a "treatment number" appended to the start. The treatment number equals the numeric alphabetical rankings of the group names or level of the treatment variable specified by the `treatvar`. In this example, the values are `community` = 1, `metcbt5` = 2, and `scy` = 3. Thus, for a study youth in the Community treatment group, the variables `ps_1_es_mean_ATE` and `ps_1_ks_mean_ATE` contain the estimated propensity scores for the `es_mean` and `ks_mean` stopping rules, respectively. The variables `ps_2_es_mean_ATE` and `ps_2_ks_mean_ATE` or `ps_3_es_mean_ATE` and `ps_3_ks_mean_ATE` contain the estimated propensity scores for youth in the MET/CBT-5 and SCY groups. A crosswalk from the treatment levels to the propensity score variables names is printed in the SAS log (see below). in estimating treatment effects.

The final argument in our example call to `%mnps` is `plotname`. The `plotname` argument is optional. When specified, it must contain the name of the file where default plots will be stored. The full path to the file can be specified or if the `objpath` argument is specified just the file name can be specified and the file with be stored in the folder specified by `objpath`. When the `plotname` argument is specified, `%mnps` creates a set of default plots, as separate pages in a single file, save in under the specified name and location. Otherwise `%mnps` does not create plots.

[2] When the estimand is "ATT", then there are no propensity score variables for the target treatment level.

SAS Log: Crosswalk from Treatment Levels to Propensity Score Variable Names for the ATE Example

```
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     CROSSWALK FROM TREATMENT LEVELS TO PROPENSITY SCORE VARIABLES
  Tx Level                                  Propensity Score Variable
 community                                         ps_1_es_mean_ATE
 community                                         ps_1_ks_mean_ATE
 metcbt5                                           ps_2_es_mean_ATE
 metcbt5                                           ps_2_ks_mean_ATE
 scy                                               ps_3_es_mean_ATE
 scy                                               ps_3_ks_mean_ATE
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

An essential component of propensity score weighting is the assessment the quality of the weights. The `twang` macros provide both graphical and tabular displays to support that assessment. The default plots generated by specifying the `plotname` argument in `%mnps` are essential component of that assessment and the initial exploration of the balance. However, the default is to create 5 different diagnostic plots which may have multiple pages for some of the default graphics. The set of plots may become large when there are several treatment options and does not give users the full control over the plots that `twang` provides. The `%mnplot` macro provides greater control over the plots to be created. It can be used to generate specific types of plots in separate files and allows for using the various subsetting tools in `twang` that can be helpful for identifying where balance problems occur, when they occur.

Like the `%plot` macro used to create plots following a call to `%ps`, `%mnplot` macro must be called after `%mnps` has been called. It uses the `mnps` object created and stored by R. By default `%mnps` stores the `mnps` object, "mnps.RData", in the temporary SAS work folder. If `objpath` is specified the object is stored in the specified folder. User can change the name of the `mnps` object but in our examples we assume the default value, 'mnps.RData". The `inputobj` argument of `%mnplot` specifies the `mnps` object. The path to the file or just the file name can be specified. If just the file name is provided then the macro will look for the file in the path specified by the `objpath` argument. If that argument is not specified, the macro will look for the file specified by `inputobj` in the default working directory. To avoid confusion about the location of the `mnps` object, we suggest user specify the `objpath` argument when using `%mnps` and its related macros including `%mnplot` and `%mnbaltable`, described below, just as we did in the example and when using the `%ps` and `%plot` macros. We provide examples of the `%mnplot` macro below.

After running the `%mnps` macro, a useful first step in the assessment of the weights before using the them to estimate treatment effects is to make sure that we let the models run for a sufficiently large number of iterations in order to optimize the balance statistics of interest. We do this using the "convergence plots" created by `twang` to visually determine whether any of the balance measures of interest still appear to be decreasing at the maximum number of iterations specified by the `ntrees` argument, which we set to 3,000 for this example (10,000 iterations is the default). The convergence plots are one of the default plot types created by specifying `plotname` argument in `%mnps`. The following code is an example of using `%mnplot` to produce the convergence plot. This code produces just the convergence plots and saves them to a file:

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_1.pdf,
        plotformat=pdf,
        multipage=TRUE,
        plots=1,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

The argument `multipage` specifies that each panel created by the plot function be placed on a separate page. The comparison of multiple treatments often results in plot with multiple panels. By default the panels are placed on a single page and this can result in plots with problematic aspect ratios and which are difficult to read. Using `multipage = TRUE` avoids this problem.

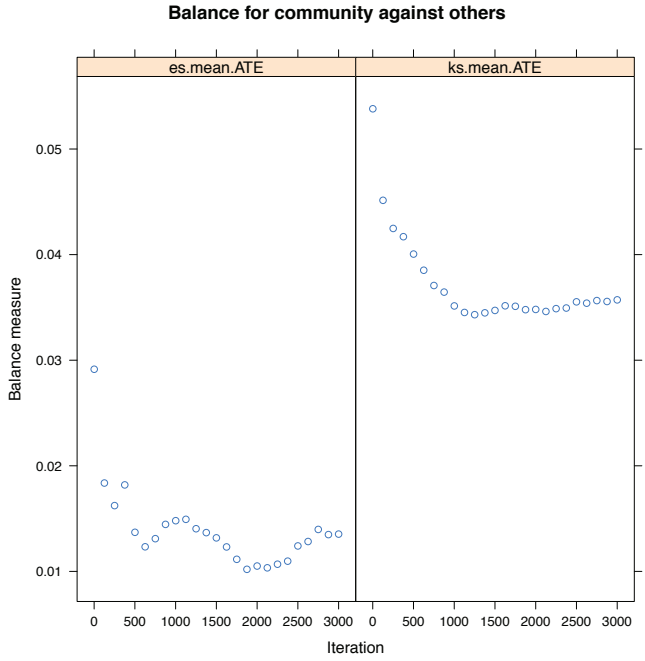

Figure 1: Example of an optimization plot for both stopping rules (es.man and ks.max) for estimating the propensity scores for comparing the Community treatment conditions to the combination of the other two to generate ATE weights for the AOD dataset.

As noted above, `mnps` estimates weights by repeated use of the `ps` function and comparing each treatment to the pooled sample of other treatments. Thus, there is one convergence plot corresponding to each of those fits. Each plot is then further divided into one panel for each stopping rule used in the estimation. Since we used the "es.mean" and "ks.mean" stopping rules there are two panels in each plot. Figures 1 to 3 show the output from the `%mnplot` code above. In these figures, it appears that each of the balance measures are optimized with substantially fewer than 3,000 iterations, so we do not have evidence that we should re-run the `%mnps` macro with a larger number of iterations or trees.

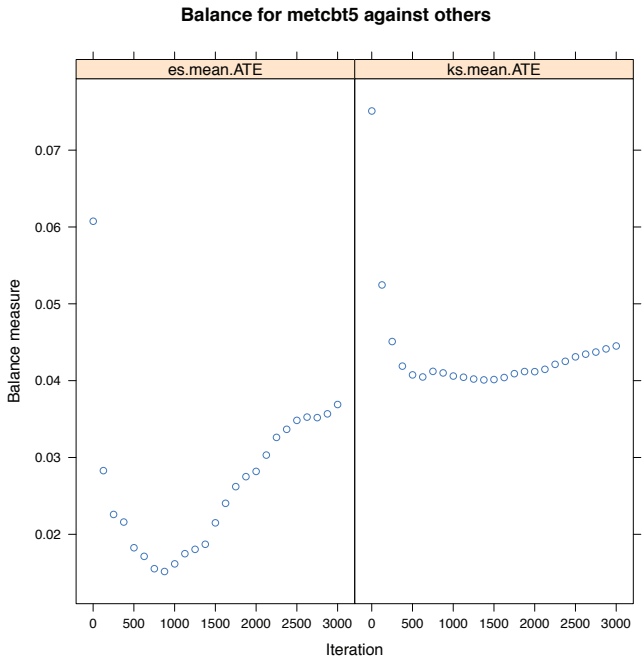

Figure 2: Example of an optimization plot for both stopping rules (es.man and ks.max) for estimating the propensity scores for comparing the MET/CBT-5 treatment conditions to the combination of the other two to generate ATE weights for the AOD dataset.

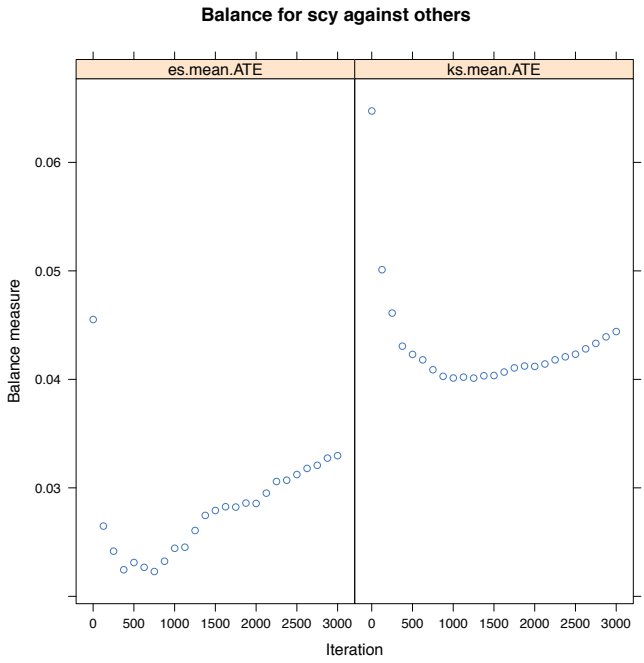

Figure 3: Example of an optimization plot for both stopping rules (es.man and ks.max) for estimating the propensity scores for comparing the SCY treatment conditions to the combination of the other two to generate ATE weights for the AOD dataset.

A useful second step is to check the key assumption in propensity score analyses that each experimental unit has a non-zero probability of receiving each treatment. The plausibility of this assumption may be assessed by examining the overlap of the empirical propensity score distributions. This diagnostic is available by setting the `plots` argument to "2" in the `%mnplot` macro. We use the `subset` option to specify which stopping rule we wish present in the plot.[3] The default panel layout for this plot results in readable figures so we do not use `multipage=TRUE`.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_2.pdf,
        plotformat=pdf,
        plots=2,
        subset=es.mean,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

The code above produces Figure 4. The overlap plot uses data from only one stopping rule, by default the one that comes first alphabetically. Hence, without specifying `subset=es.mean` the overlap for the weights from the model fit using that rule would be compared. To see the results for the "ks.mean" rule requires using `subset=ks.mean`.

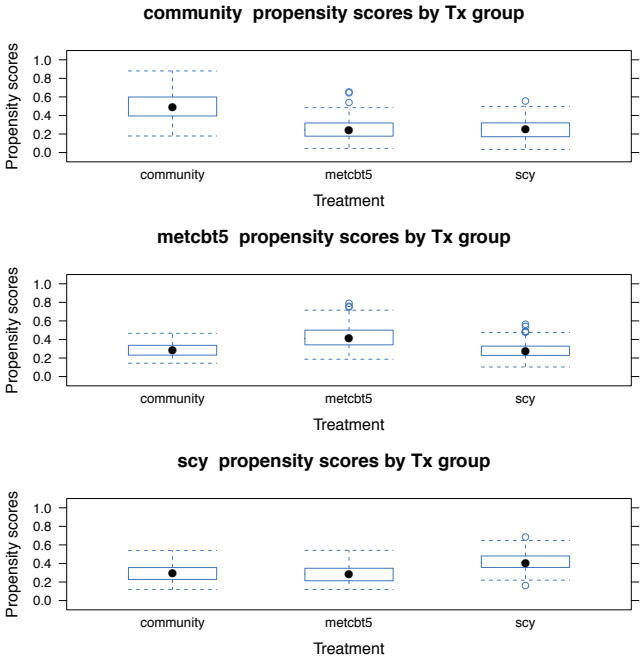

Figure 4: Example of an overlap boxplot for the es.mean stopping rule for estimating the propensity scores to generate ATE weights for the AOD dataset.

As shown in Figure 4, the overlap assumption generally seems to be met in our example, although there should be some concern that adolescents in the metcbt5 and scy conditions do

[3]The value for the `subset` argument can be a character variable with the name of the stopping rule, as was used in the example code, or a number corresponding to the stopping rule. Stopping rules are numbered by the alphabetical ordering among the rules specified in the `mnps` call.

not overlap well with the community group given the top most graphic. See McCaffrey et al. (2013) for more details on this issue.

2.2 Graphical assessments of balance

As with the `%ps` and `%plot` macros for the binary treatment setting, `%mnps` and `%mnplot` also can generate plots to display information on commonly-used balance statistics. Checking balance is an essential part of any propensity score analysis and must be done thoroughly prior to moving into outcome analyses. The `twang` SAS macros provide the user with two ways to assess balance: graphical displays or tabular displays. Here we discuss how to create graphical displays of balance. Graphical displays can be produced by setting the `plots` argument for `%mnplot` equal to "3", "4", or "5". In particular, when the `plots` argument is set equal to "3", `%mnplot` provides comparisons of the absolute standardized mean differences (ASMD) between the treatment groups on the pretreatment covariates, before and after weighting. When the `plots` argument is set equal to "4", the display is of t-test and chi-squared statistic p-values from comparing the two groups before and after weighting and setting the argument to "5" generates the corresponding p-value plots for tests of the KS statistics. However, whereas there is a single plot for these balance diagnostics in the binary treatment setting, in the multiple treatment case, one can either examine a plot for each of the pairwise comparisons (e.g., Community versus the others, MET/CBT-5 versus the others, or SCY versus the others), or summarize the balance statistics, in some way, across the treatment conditions. As a default, the `%mnplot` macro returns the maximum of the pairwise balance statistics across treatment groups for each of the covariates:

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_3.pdf,
        plotformat=pdf,
        plots=3,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

The code above produces Figure 5. As shown in that figure, after propensity score weighting, the maximum ASMD decreases for all pretreatment covariates. The statistically significant difference (before taking the maximum across treatment groups) is indicated by the solid circle. One may see the balance plots for the individual fits by setting the `pairwisemax` argument to "FALSE" as in the following code.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_3b.pdf,
        plotformat=pdf,
        plots=3,
        pairwisemax=FALSE,
        figurerows=3,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

The additional `figurerows` argument instructs the function to spread the plots over three rows, as shown in Figure 6. By default the plots would be arranged in a single row rather than a column. This produces an unreadable figure. We note here that red lines represent pretreatment covariates for which the pairwise ASMDs increase after weighting.

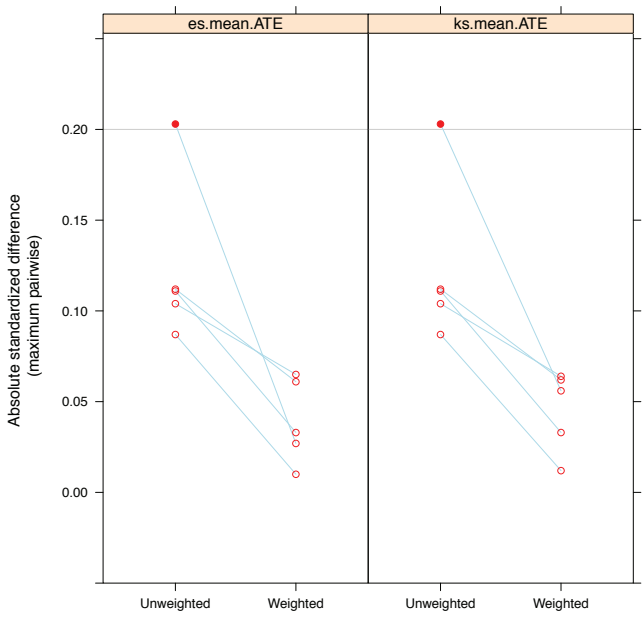

Figure 5: Example of a standardized effect size plot for estimating the propensity scores to generate ATE weights for the AOD dataset.

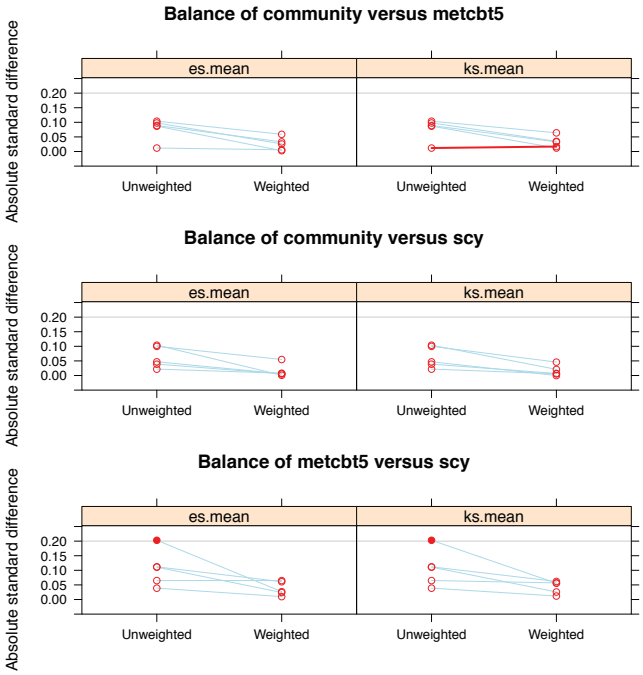

Figure 6: Example of a standardize effect size plot for each pairwise comparison of treatments for estimating the propensity scores to generate ATE weights for the AOD dataset.

Setting the `plots` argument equal to "4" displays t-test or chi-squared statistic pairwise minimum p-values for differences between each of the individual treatment groups and observations in all other treatment groups. The following command produces the results shown in Figure 7. As seen in that figure, the pairwise minimum p-values all increase after propensity score weighting.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_4.pdf,
        plotformat=pdf,
        plots=4,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

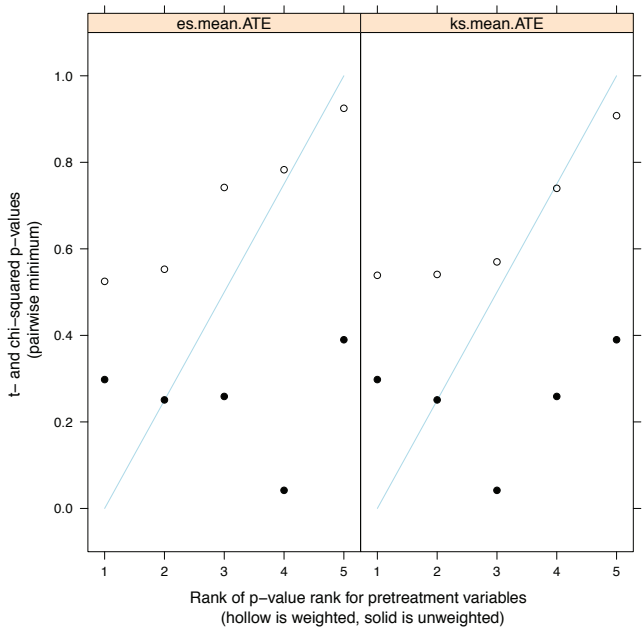

Figure 7: Example of a p-value plot for t-tests and Chi-square tests for estimating the propensity scores to generate ATE weights for the AOD dataset.

Some of the figures include many frames, which can result in figures that are difficult to read. There are three arguments to the `%mnplot` to control the placement of multiple panels across pages in the graphics file. First, the `treatments` argument can be used to specify only comparisons that involve a specific treatment level or, in the ATE case, only comparisons between two specified treatment levels. Similarly, the `singleplot` argument can be used to plot only a single page from a call that will generate multiple pages of plots or multiple frames on a single page. For example, `singleplot = 2` would display only the second page of those produced by the plot command (see figure below). Finally, as described previously, specifying `multipage = TRUE` prints in succession the frames generated by the plot function.

The following code and corresponding figure (Figure 8) demonstrate using these arguments to plot the p-values for the KS tests when comparing the treatment levels of community and scy. By specifying `pairwisemax=FALSE`, each pairwise comparison of treatment will be plotted and by

specifying `multipage=TRUE`, each comparison will be on a separate page. The comparisons are ordered alphabetically so the first page is community versus metcbt5, the second is community versus scy, and the third is metcbt5 versus scy. By setting `singleplot=2`, we select the p-values for the tests from the comparison of community with scy.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_5.pdf,
        plotformat=pdf,
        plots=5,
        pairwisemax=FALSE,
        multipage=TRUE,
        singleplot=2,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

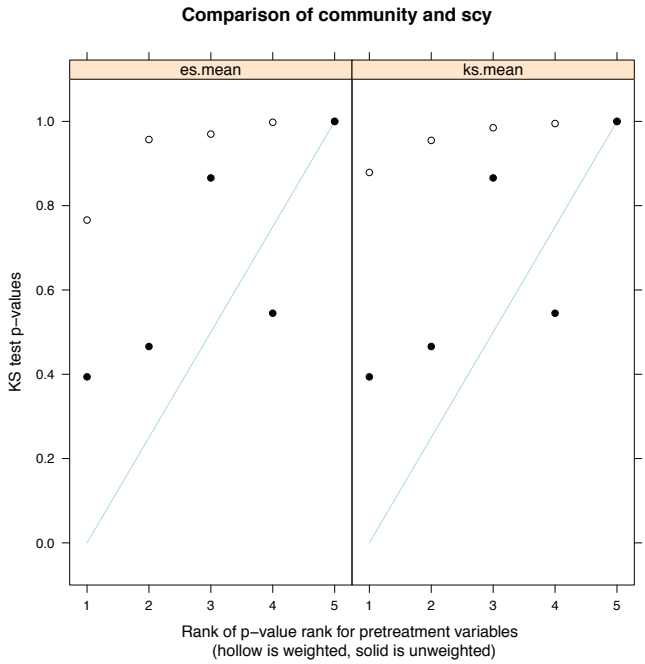

Figure 8: Example of a p-value plot KS tests for estimating the propensity scores to generate ATE weights for the AOD dataset, using control arguments to plot only the tests for the comparison of community to scy.

2.3 Tabular assessments of balance

There are two primary ways to obtain tabular assessments of balance via the `twang` macros in SAS. First, the `%mnps` macro returns a number of useful tabular summaries of the balance as part of its default output when running the macro. Additionally, the user can make use of the `%mnbalance` macro to obtain more fine-tuned and simplistic summaries of balance.

In terms of the output produced by default by the `%mnps` macro, the first table is the Summary Table which presents the maximum values of the balance statistics and minimum values

of p-values across all covariates and all pairwise comparisons of the treatments (e.g., in our case study Community versus MET/CBT-5; Community versus SCY; MET/CBT-5 versus SCY). There is one line in the table for the comparisons prior to weighting and one line for weighting with the weights generate by the model selected by each stopping rule. This table generated by our example `%mnps` code follows. In our example we use the "es.mean" and "ks.mean" stopping rules so the summary table has three rows. The summary allow us to quickly see how the maximum ASMDs have gotten smaller and minimum p-values have gotten larger, as desired, after propensity score weighting.

SAS Output: Summary Table for the ATE Example

Summary table
10:03 Monday, December 22, 2014 1
Summary of pairwise comparisons

Obs	max_std_eff_sz	min_p	max_ks	min_ks_pval	stop_method
1	0.2026644577	0.0416156159	0.13	0.0680192046	unw
2	0.0652929809	0.5252523487	0.0666198546	0.7663958881	es.mean
3	0.0644845458	0.5394742624	0.0645909263	0.798529274	ks.mean

The Summary Table also prints a second page of sample sizes and effective sample sizes. This table has one row per treatment group and one column for the original sample size ("n") and then one column for the effective sample size for the weights generated using each stopping rule. As shown in the output from our `%mnps` call, this summary table allows for a quick check of whether the weights are highly variable and could potentially yield a very imprecise treatment effect estimate. In this case the effective sample sizes are all close to the actual sample size because the groups are fairly well balanced even before weighting.

SAS Output: Summary Table of Sample Size and Effective Sample Sizes for the ATE Example

Summary table
10:03 Monday, December 22, 2014 2
Sample sizes and effective sample sizes

Obs	treatment	n	ESS_es_mean	ESS_ks_mean
1	community	200	184.51236554	187.47130269
2	metcbt5	200	186.18737054	183.39867453
3	scy	200	189.5017217	185.7008875

The macro also produces a table of balance statistics for each covariate for each pairwise comparison of treatments. This table from our `%mnps` example follows. This table contains balance information for unweighted comparisons and weighted comparisons using the weights generated under each specified stopping rule. There is one record per covariate, per pairwise comparison, per weight (including the unweighted case). The printed table has breaks between each stop method. Each record contains:

- `tmt1` – the name of the first treatment group in the pairwise comparison; names are sorted alphabetically

- `tmt2` – the name of the second treatment group in the pairwise comparison
- `var` – the name of the covariate being assessed
- `mean1` – the covariate mean for the first treatment group
- `mean2` – the covariate mean for the second treatment group
- `pop_sd` – the pooled within sample standard deviation from all treatment groups
- `std_eff_sz` – the ASMD or absolute effect size equal to the absolute value of the difference in the group means divided by the `pop_sd`
- `p` – the p-value of the t-test (continuous variables) or the Chi-squared test (categorical variables)
- `ks` – the KS statistic for comparing the covariate distribution for the two groups
- `ks_pval` – the approximate p-value for testing the KS statistic
- `stop_method` – the stop method used for generating the weights or "unw" for the unweighted comparison.

SAS Output: The Default Balance Table for the ATE Example

```
                Balance table: unw                                  3
                                        10:03 Monday, December 22, 2014
```

Obs	tmt1	tmt2	var	mean1	mean2	pop_sd
1	community	metcbt5	illact	0.097	0.007	1.014
2	community	metcbt5	crimjust	-0.065	0.037	1.041
3	community	metcbt5	subprob	-0.06	0.026	0.985
4	community	metcbt5	subdep	0.046	0.058	1.031
5	community	metcbt5	white	0.16	0.2	0.383
6	community	scy	illact	0.097	0.12	1.014
7	community	scy	crimjust	-0.065	-0.174	1.041
8	community	scy	subprob	-0.06	-0.013	0.985
9	community	scy	subdep	0.046	-0.058	1.031
10	community	scy	white	0.16	0.175	0.383
11	metcbt5	scy	illact	0.007	0.12	1.014
12	metcbt5	scy	crimjust	0.037	-0.174	1.041
13	metcbt5	scy	subprob	0.026	-0.013	0.985
14	metcbt5	scy	subdep	0.058	-0.058	1.031
15	metcbt5	scy	white	0.2	0.175	0.383

Obs	std_eff_sz	p	ks	ks_pval	stop_method
1	0.089	0.385	0.1	0.27	unw
2	0.098	0.328	0.105	0.221	unw
3	0.087	0.39	0.09	0.394	unw
4	0.012	0.91	0.055	0.924	unw

5	0.104	0.298	0.04	0.997	unw
6	0.022	0.823	0.06	0.866	unw
7	0.104	0.295	0.08	0.545	unw
8	0.047	0.631	0.09	0.394	unw
9	0.1	0.312	0.085	0.466	unw
10	0.039	0.688	0.015	1	unw
11	0.111	0.259	0.11	0.178	unw
12	0.203	0.042	0.13	0.068	unw
13	0.039	0.696	0.065	0.793	unw
14	0.112	0.251	0.09	0.394	unw
15	0.065	0.523	0.025	1	unw

Balance table: ks.mean
10:03 Monday, December 22, 2014

Obs	tmt1	tmt2	var	mean1	mean2	pop_sd
31	community	metcbt5	illact	0.083	0.05	1.014
32	community	metcbt5	crimjust	-0.084	-0.048	1.041
33	community	metcbt5	subprob	-0.001	-0.012	0.985
34	community	metcbt5	subdep	0.007	0.024	1.031
35	community	metcbt5	white	0.169	0.194	0.383
36	community	scy	illact	0.083	0.077	1.014
37	community	scy	crimjust	-0.084	-0.106	1.041
38	community	scy	subprob	-0.001	-0.001	0.985
39	community	scy	subdep	0.007	-0.04	1.031
40	community	scy	white	0.169	0.172	0.383
41	metcbt5	scy	illact	0.05	0.077	1.014
42	metcbt5	scy	crimjust	-0.048	-0.106	1.041
43	metcbt5	scy	subprob	-0.012	-0.001	0.985
44	metcbt5	scy	subdep	0.024	-0.04	1.031
45	metcbt5	scy	white	0.194	0.172	0.383

Obs	std_eff_sz	p	ks	ks_pval	stop_method
31	0.033	0.74	0.062	0.839	ks.mean
32	0.035	0.723	0.052	0.95	ks.mean
33	0.012	0.908	0.053	0.934	ks.mean
34	0.017	0.873	0.049	0.966	ks.mean
35	0.064	0.539	0.025	1	ks.mean
36	0.006	0.95	0.045	0.985	ks.mean
37	0.021	0.83	0.041	0.995	ks.mean
38	0	1	0.058	0.879	ks.mean
39	0.046	0.65	0.051	0.955	ks.mean
40	0.007	0.946	0.003	1	ks.mean
41	0.026	0.797	0.062	0.832	ks.mean
42	0.056	0.57	0.064	0.809	ks.mean
43	0.012	0.911	0.035	0.999	ks.mean

44	0.062	0.541	0.065	0.799	ks.mean
45	0.057	0.58	0.022	1	ks.mean

Balance table: es.mean
10:03 Monday, December 22, 2014

Obs	tmt1	tmt2	var	mean1	mean2	pop_sd
16	community	metcbt5	illact	0.085	0.052	1.014
17	community	metcbt5	crimjust	-0.092	-0.065	1.041
18	community	metcbt5	subprob	-0.013	-0.016	0.985
19	community	metcbt5	subdep	0.015	0.021	1.031
20	community	metcbt5	white	0.173	0.195	0.383
21	community	scy	illact	0.085	0.077	1.014
22	community	scy	crimjust	-0.092	-0.093	1.041
23	community	scy	subprob	-0.013	-0.007	0.985
24	community	scy	subdep	0.015	-0.042	1.031
25	community	scy	white	0.173	0.17	0.383
26	metcbt5	scy	illact	0.052	0.077	1.014
27	metcbt5	scy	crimjust	-0.065	-0.093	1.041
28	metcbt5	scy	subprob	-0.016	-0.007	0.985
29	metcbt5	scy	subdep	0.021	-0.042	1.031
30	metcbt5	scy	white	0.195	0.17	0.383

Obs	std_eff_sz	p	ks	ks_pval	stop_method
16	0.033	0.742	0.057	0.896	es.mean
17	0.026	0.793	0.054	0.931	es.mean
18	0.003	0.974	0.062	0.831	es.mean
19	0.006	0.958	0.05	0.965	es.mean
20	0.059	0.582	0.023	1	es.mean
21	0.008	0.937	0.048	0.97	es.mean
22	0.001	0.989	0.037	0.998	es.mean
23	0.006	0.949	0.067	0.766	es.mean
24	0.055	0.582	0.051	0.957	es.mean
25	0.006	0.95	0.002	1	es.mean
26	0.024	0.812	0.065	0.793	es.mean
27	0.027	0.783	0.057	0.896	es.mean
28	0.01	0.925	0.036	0.999	es.mean
29	0.061	0.553	0.065	0.794	es.mean
30	0.065	0.525	0.025	1	es.mean

For propensity score analyses with multiple treatments, the balance table information returned can be quite overwhelming and, with many covariates, sorting through that information can be challenging. More parsimonious versions of the summaries are available using the optional `collapseto` argument of the `%mnps`. The following example shows a call to `%mnps` with the `collapseto` argument set to "covariate". Other options for the `collapseto` argument are "none" or "pair", which return the default output, and "stop.method" which produces a single summary of balance across all covariates and all pairwise group comparisons for each stop method

yielding the same information as the Summary Table. Setting `collapseto = covariate` gives the maximum of the ASMD and the minimum of the p-value across all pairwise comparisons for each pretreatment covariate and stopping rule combination. There is one printed page per stopping rule or unweighted, as shown in the output that follow the example macro call.

```
%mnps(treatvar= treat,
      vars = illact crimjust subprob subdep white,
      dataset = sasin.AOD,
      estimand = ATE,
      stopmethod = es.mean ks.mean,
      ntrees = 3000,
      collapseto=covariate,
      Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
      objpath=C:\Users\uname\twang_mnps_example)
```

SAS Output: The Collapsed Balance Table for the ATE Example

```
                Balance table: unw                          1
                                10:48 Tuesday, December 23, 2014

              max_std_                                    stop_
Obs  var        eff_sz      min_p     max_ks   min_ks_pval  method

 1  illact        0.111      0.259      0.11      0.178     unw
 2  crimjust      0.203      0.042      0.13      0.068     unw
 3  subprob       0.087      0.39       0.09      0.394     unw
 4  subdep        0.112      0.251      0.09      0.394     unw
 5  white         0.104      0.298      0.04      0.997     unw
                    Balance table: ks.mean                  2
                                10:48 Tuesday, December 23, 2014

              max_std_                                    stop_
Obs  var        eff_sz      min_p     max_ks   min_ks_pval  method

 11  illact       0.033      0.74      0.062      0.832    ks.mean
 12  crimjust     0.056      0.57      0.064      0.809    ks.mean
 13  subprob      0.012      0.908     0.058      0.879    ks.mean
 14  subdep       0.062      0.541     0.065      0.799    ks.mean
 15  white        0.064      0.539     0.025      1        ks.mean
                    Balance table: es.mean                  3
                                10:48 Tuesday, December 23, 2014

              max_std_                                    stop_
Obs  var        eff_sz      min_p     max_ks   min_ks_pval  method

 6  illact        0.033      0.742      0.065      0.793     es.mean
 7  crimjust      0.027      0.783      0.057      0.896     es.mean
 8  subprob       0.01       0.925      0.067      0.766     es.mean
 9  subdep        0.061      0.553      0.065      0.794     es.mean
 10  white        0.065      0.525      0.025      1         es.mean
```

As shown, for each pretreatment variable, the maximum ASMD has decreased and the minimum p-values have increased after applying weights that arise from either stopmethod.

The %mnbaltable macro provides another way to produce more concise and targeted balance tables. It can also be used to produce a full balance table with balance information for all combinations of the covariates, pairwise comparisons, and stopping methods. The macro must be run after %mnps and like %mnplot uses the mnps object created by R and saved in the default working folder or the folder specified the objpath argument. The following code demonstrates the use of %mnbaltable to create the same table as previous example of the %mnps macro with the collapseto argument. The call assumes the mnps object is saved in the default file. The %mnbaltable macro can be run any time after %mnps is run provided the mnps object is saved.

```
%mnbaltable(inputobj=mnps.RData,
            collapseto=covariate,
            Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
            objpath=C:\Users\uname\twang_mnps_example)
```

The %mnbaltable macro also provides options other than collapseto to control the information presented in the balance tables. These other options are not available in %mnps. Rather than collapsing the values of the table as described above, these other options allow for subsetting the balance table output. The arguments subset_var and subset_stop_method instruct the macro to include only the results for the indicated covariates or stop method. The subset_treat argument instructs the macro to return only the pairwise comparisons including the specified treatment or, if two treatment levels are indicated, the comparisons of those two treatments. Note that subset_treat may not be used when collapseto is specified as "stop.method" or "covariate". However, it may be used in combination with the subset_var and subset_stop_method arguments to generate tables restricted to specific pairwise comparisons of different pairs of treatments. Also, the values of the treatment variable specified for subset_treat are case sensitive. The following code demonstrates subset_treat and subset_var, used in combination in the example.

```
%mnbaltable(inputobj=mnps.RData,
            subset_treat=community metcbt5,
            subset_var=white illact crimjust,
            Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
            objpath=C:\Users\uname\twang_mnps_example)
```

SAS Output: The Balance Table for the ATE Example restricted to a subset of treatments and covariates

```
                        Balance table: unw                              4
                                      11:48 Tuesday, December 23, 2014


Obs  tmt1       tmt2       var          mean1      mean2     pop_sd

 1  community   metcbt5   illact        0.097      0.007      1.014
 2  community   metcbt5   crimjust     -0.065      0.037      1.041
 3  community   metcbt5   white         0.16       0.2        0.383

                                                      stop_
Obs   std_eff_sz      p        ks      ks_pval      method

 1      0.089      0.385      0.1       0.27       unw
```

```
 2      0.098      0.328      0.105    0.221   unw
 3      0.104      0.298      0.04     0.997   unw

                    Balance table: ks.mean                        5
                                  11:48 Tuesday, December 23, 2014


Obs  tmt1       tmt2      var         mean1      mean2     pop_sd

 7  community   metcbt5   illact       0.083      0.05      1.014
 8  community   metcbt5   crimjust    -0.084     -0.048     1.041
 9  community   metcbt5   white        0.169      0.194     0.383

                                                    stop_
Obs   std_eff_sz      p       ks      ks_pval   method

 7      0.033      0.74      0.062     0.839    ks.mean
 8      0.035      0.723     0.052     0.95     ks.mean
 9      0.064      0.539     0.025     1        ks.mean

                      Balance table: es.mean                      6
                                  11:48 Tuesday, December 23, 2014


Obs  tmt1      tmt2     var            mean1      mean2     pop_sd

 4  community   metcbt5   illact       0.085      0.052      1.014
 5  community   metcbt5   crimjust    -0.092     -0.065      1.041
 6  community   metcbt5   white        0.173      0.195      0.383

                                                    stop_
Obs   std_eff_sz        p       ks      ks_pval   method

 4      0.033      0.742      0.057      0.896   es.mean
 5      0.026      0.793      0.054      0.931   es.mean
 6      0.059      0.582      0.023      1       es.mean
```

This code demonstrates `subset_stop_method`. It is combined with the use of the `collapseto` argument to yield a aggregated summaries of balance for each of the five covariates for the "es.mean" stopping rule only.

```
%mnbaltable(inputobj=mnps.RData,
            collapseto=covariate,
            subset_stop_method=es.mean,
            Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
            objpath=C:\Users\uname\twang_mnps_example)
```

SAS Output: The Balance Table for the ATE Example collapsed and restricted to one stopping rule

```
                          Balance table: es.mean                    7
                                    11:48 Tuesday, December 23, 2014

                max_std_                                          stop_
Obs  var          eff_sz      min_p     max_ks   min_ks_pval    method

 1   illact        0.033      0.742      0.065      0.793      es.mean
 2   crimjust      0.027      0.783      0.057      0.896      es.mean
 3   subprob       0.01       0.925      0.067      0.766      es.mean
 4   subdep        0.061      0.553      0.065      0.794      es.mean
 5   white         0.065      0.525      0.025      1          es.mean
```

One of the primary uses of the balance table is to identify covariates that have large imbalances across groups for one or more pairwise comparisons. With multiple treatments, variables, and stopping rules, finding such covariates can be difficult. The `%mnbaltable` macro includes arguments to make this job easier by subsetting the table on the basis of the values of ES and KS and their related p-values. The arguments `es_cutoff` and `ks_cutoff` result in tables that include only records where the statistics exceed the given cutoff values and the arguments `p_cutoff`, and `ks_p_cutoff` include only rows where the p-values are less than the given cutoff values. For example `p_cutoff = 0.1` would print only rows with p-values less than or equal to 10%, and `es_cutoff = 0.2` includes only rows with ES values greater than or equal to 0.2 in absolute value. The following example demonstrates how to use these subsetting arguments.

```
%mnbaltable(inputobj=mnps.RData,
            es_cutoff=0.10,
            Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
            objpath=C:\Users\uname\twang_mnps_example)
```

SAS Output: The Balance Table for the ATE Example restricted to large ES

```
                  Balance table: unw  12:13 Tuesday, December 23, 2014  8

Obs  tmt1        tmt2      var          mean1      mean2     pop_sd

 1   community   metcbt5   white         0.16       0.2      0.383
 2   community   scy       crimjust     -0.065     -0.174    1.041
 3   community   scy       subdep        0.046     -0.058    1.031
 4   metcbt5     scy       illact        0.007      0.12     1.014
 5   metcbt5     scy       crimjust      0.037     -0.174    1.041
 6   metcbt5     scy       subdep        0.058     -0.058    1.031

                                                    stop_
Obs   std_eff_sz      p       ks       ks_pval     method

 1      0.104      0.298      0.04     0.997      unw
 2      0.104      0.295      0.08     0.545      unw
 3      0.1        0.312      0.085    0.466      unw
 4      0.111      0.259      0.11     0.178      unw
 5      0.203      0.042      0.13     0.068      unw
 6      0.112      0.251      0.09     0.394      unw
```

After examining the graphical and tabular diagnostics provided by `twang`, we can analyze the outcome variable using the propensity score weights generated by the `mnps` function. Although two stop methods were specified initially ("es.mean" and "ks.mean"), at this point we have to commit to a single set of weights. From the `%mnbaltable` call above, we see that the balance properties are very similar for the two stopping rules, and from the Summary Table, we see that the effective sample sizes (ESS) are similar as well. Hence, we expect the two stop methods to give similar results; we choose to analyze the data with the `es.mean` weights.

2.4 Estimating treatment effects

As in McCaffrey et al. (2013) we consider estimating treatment effects on suf12, the substance frequency scale which measures frequency of substance use during the past 90 days prior to the 12-month follow-up visits for individuals in the study. In order for the weights to be used properly, it is necessary to use a procedure that correctly interprets weights as probability weights. These are PROC SURVEYREG, PROC SURVEYFREQ, PROC SURVEYLOGISTIC, PROC SURVEYMEANS and PROC SURVEYPHREG. Other procedures that allow for weights do not treat them as probability weights, which can result in biased estimates or standard errors.

We can estimate the average treatment effects using PROC SURVEYREG.

```
proc surveyreg data=sasin.aodwgt;
   class treat;
   model suf12 = treat / solution;
   weight es_mean_ATE;
   estimate "community vs. metcbt5"
            treat 1 -1 0;
run;
```

SAS Output: PROC SURVEYREG estimates of treatment effects for the ATE example

```
                              The SAS System
                                13:26 Tuesday, December 23, 2014   1

                           The SURVEYREG Procedure

                 Regression Analysis for Dependent Variable suf12

                                Data Summary

                    Number of Observations           600
                    Sum of Weights                1428.0
                    Weighted Mean of suf12       -0.02468
                    Weighted Sum of suf12       -35.23690


                                Fit Statistics

                         R-square          0.003616
                         Root MSE            1.0027
                         Denominator DF         599


                          Class Level Information
```

```
Class
Variable      Levels    Values

treat              3    community metcbt5 scy
```

Tests of Model Effects

Effect	Num DF	F Value	Pr > F
Model	2	1.00	0.3691
Intercept	1	0.43	0.5132
treat	2	1.00	0.3691

NOTE: The denominator degrees of freedom for the F tests is 599.

Estimated Regression Coefficients

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	-0.0344831	0.07401063	-0.47	0.6414
treat community	-0.0646436	0.10014853	-0.65	0.5189
treat metcbt5	0.0839397	0.10950728	0.77	0.4437
treat scy	0.0000000	0.00000000	.	.

```
NOTE: The denominator degrees of freedom for the t tests is 599.
      Matrix X'WX is singular and a generalized inverse was used to solve the
      normal equations.
      Estimates are not unique.
```


The SURVEYREG Procedure

Regression Analysis for Dependent Variable suf12

Estimate

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
community vs. metcbt5	-0.1486	0.1052	599	-1.41	0.1583

By default, PROC SURVEYREG includes dummy variables for community and metcbt5, scy is the holdout group (the holdout is the group with the label that comes last alphabetically). Consequently, the estimated effect for metcbt5 equals the weighted mean for the metcbt5 sample less the weighted mean for the scy sample, where both means are weighted to match the overall sample. Similarly, the effect for community equals the difference in the weighted means for the

community and scy samples. The coefficients estimate the causal effects of Community vs. SCY and MET/CBT-5 vs. SCY, respectively, assuming there are no unobserved confounders. We use the ESTIMATE statement to estimated the effect Community vs. MET/CBT-5.

Using this small subset of the data, we are unable to detect differences in the treatment group means. In the context of this application, the signs of the estimates correspond to higher substance use frequency for youths exposed to MET/CBT-5 or SCY and lower use for youth exposed to Community relative to SCY. The estimate statement is estimating the average treatment effect of Community relative to MET/CBT-5 for all the youths in the population. Youth exposed to Community have lower use but the estimate is not statistically significant.

By default PROC SURVEYREG uses an $\sqrt{(n-1)/(n-k)}$ adjustment to the standard error, where n is the sample size and p is the number of regressors. The corresponding procedure in R, the `svyglm` function of the `survey` package, does not use this adjustment,[4] so analyses with SAS and R will yield different standard errors. For moderate sample sizes the differences are very small. To remove the adjustment the VADJUST option can be added to the MODEL statement as shown in the following code. Setting this option in the MODEL statement also controls whether or not an adjustment is used in the tests specified by the ESTIMATE statement. We do not include the output of the following code. The results are nearly identical to those above except the standard errors differ in fourth or higher decimal places.

```
proc surveyreg data=sasin.aodwgt;
   class treat;
   model suf12 = treat / solution vadjust=none;
   weight es_mean_ATE;
   estimate "community vs. metcbt5"
            treat 1 -1 0;
run;
```

3 An ATT example

3.1 Estimating the weights

It is also possible to explore treatment effects on the treated (ATTs) using the `%mnps` macro. A key difference in the multiple treatment setting is that we must be clear as to which treatment condition "the treated" refers to. This is done through the `treatatt` argument. Here, we define the treatment group of interest to be the community group; thus, we are trying to draw inferences about the relative effectiveness of the three treatment groups for individuals like those who were enrolled in the community program. Since we are not interested in the propensity scores we do not specify the `return_ps` paramater; rather we use the default value of `FALSE`.

```
%mnps(treatvar = treat,
      vars = illact crimjust subprob subdep white,
      dataset = sasin.AOD,
      estimand = ATT,
      treatATT = community,
      stopmethod = es.mean ks.mean,
      ntrees = 3000,
      output_dataset=sasin.aodattwgt,
      Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
      objpath=C:\Users\uname\twang_mnps_example)
```

[4] We used `survey` package version 3.29-5

3.2 Graphical assessments of balance

The same basic graphical descriptions are available as in the ATE case, though it is important to note that these comparisons all assess balance relative to the "treatment" group rather than by comparing balance for all possibly pairwise treatment group comparisons as is done with ATE. Specifying the `plotname` argument will generate the full set of default plots. Alternatively the `%mnplot` can be used to create specific plots, as it was for ATE case. The following code produces the graphics shown in Figures 9 and 10.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_att_1.pdf,
        plotformat=pdf,
        multipage=TRUE,
        plots=1,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

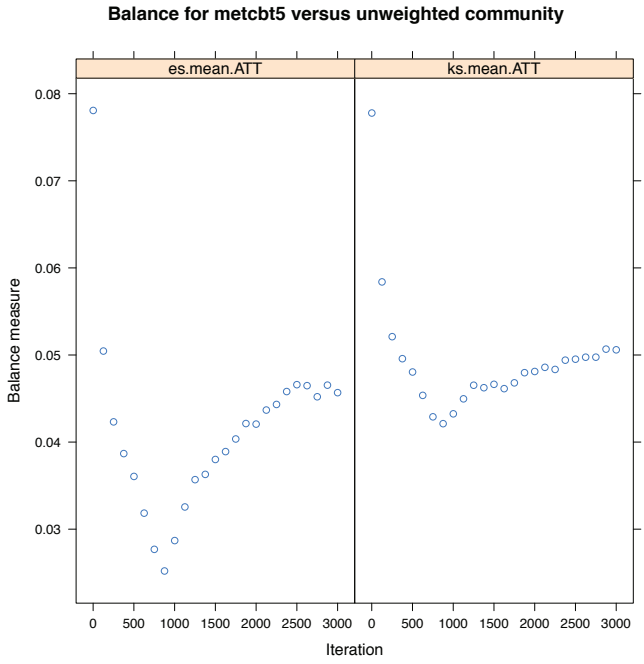

Figure 9: Example of an optimization plot for both stopping rules (es.man and ks.max) for estimating the propensity scores for comparing the MET/CBT-5 condition to the Community condition to generate ATT weights for the AOD dataset for a target population of those who received community treatment.

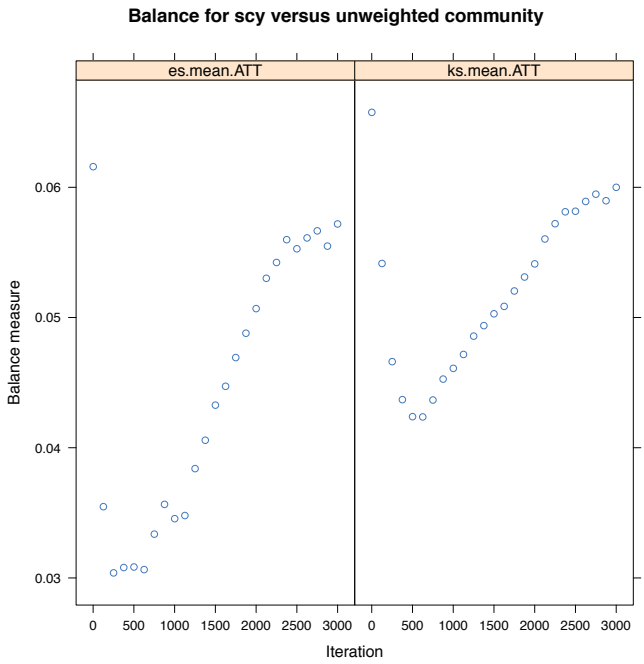

Figure 10: Example of an optimization plot for both stopping rules (es.man and ks.max) for estimating the propensity scores for comparing the SCY condition to the Community condition to generate ATT weights for the AOD dataset for a target population of those who received community treatment.

When the estimand is "ATT" there is one propensity score model fit for comparing each of the other treatments to the treatment specified by the `treatatt` argument. In this case, the target treatment is "community" so there is one model for comparing "metcbt5" to "community" and another for comparing "scy5" to "community". Consequently there is one optimization plot for the GBM model to compare "metcbt5" to "community" and another for comparing "scy" to "community". Similarly, we can look at the balance for each of the pairwise comparisons (here, SCY versus Community and MET/CBT5 versus Community) using the effect size plots (setting the `plots` argument to "3" or "es"). The following code produces Figure 11.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_att_3.pdf,
        plotformat=pdf,
        multipage=TRUE,
        plots=3,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

By default a call to `%mnplot` with the `plots` argument equal to "3", as with the previous code, generates a plot of the maximum standardized effect across both comparisons (SCY versus Community and MET/CBT5 versus Community) for each covariate. This is useful for determining if balance is satisfactory or if there are problems but it is not as useful for assessing the implications of balance problems if any exist. To probe the balance in more detail, as with

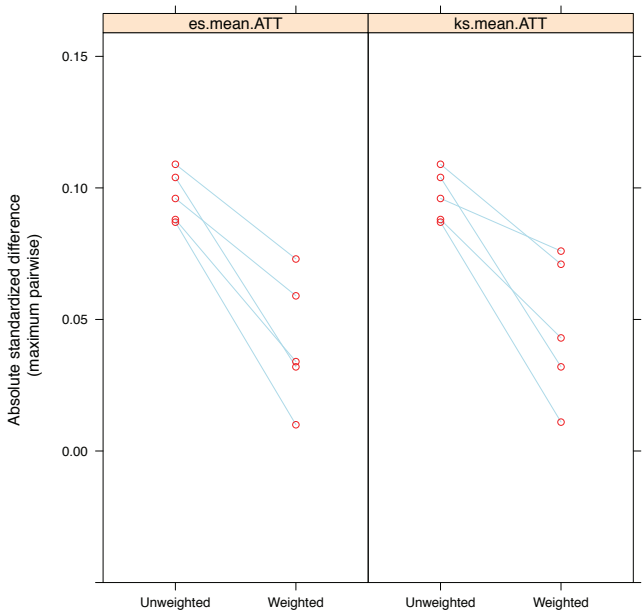

Figure 11: Example of an effect size plot for both stopping rules (es.man and ks.max) for comparing the MET/CBT-5 or the SCY condition to the Community condition to generate ATT weights for the AOD dataset for a target population of those who received community treatment. Plot of the maximum effect size across both comparisons for each covariate.

ATE, separate plots for each pairwise comparison can be created by specifying `pairwisemax` as "FALSE". We also set `multipage=T` so that each plot will be on a separate page. The following code produces Figures 12 and 13.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_att_3.pdf,
        plotformat=pdf,
        multipage=TRUE,
        plots=3,
        pairwisemax=FALSE,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

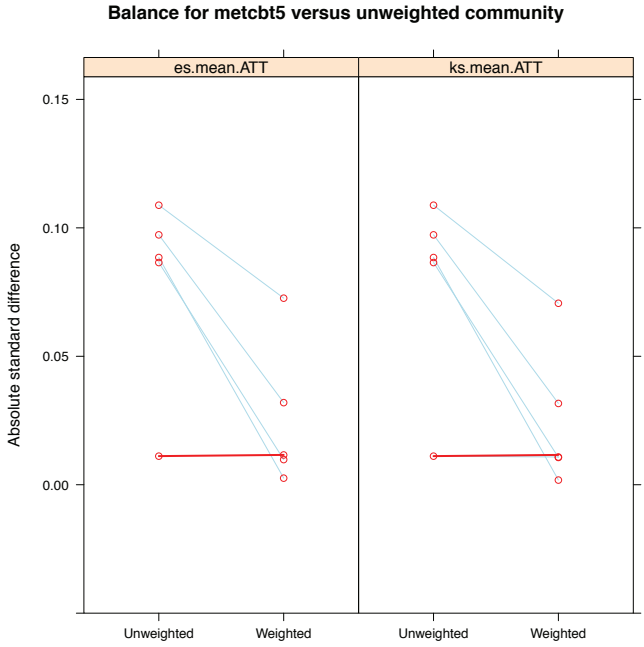

Figure 12: Example of an effect size plot for both stopping rules (es.man and ks.max) for comparing the MET/CBT-5 condition to the Community condition to generate ATT weights for the AOD dataset for a target population of those who received community treatment.

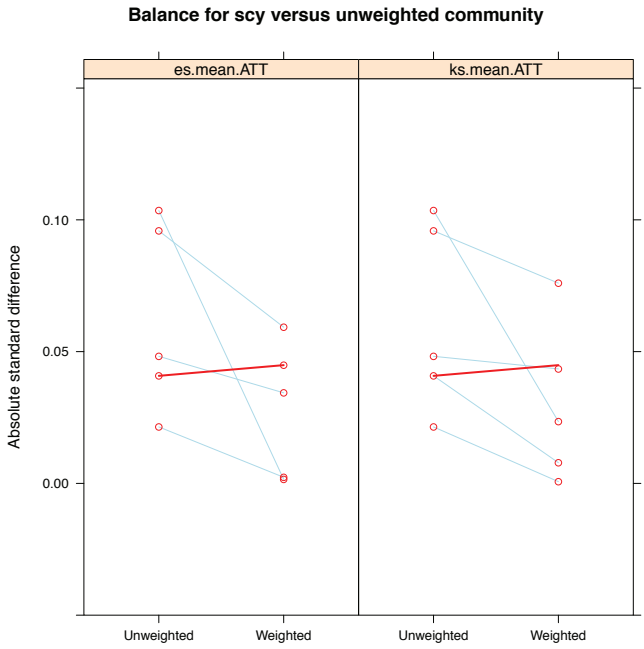

Figure 13: Example of an effect size plot for both stopping rules (es.man and ks.max) for comparing the SCY condition to the Community condition to generate ATT weights for the AOD dataset for a target population of those who received community treatment.

The p-value plots can also be useful as part of the assessment of balance. The minimum p-value across comparisons is the default and, like the effect size plot, the plots for each separate pairwise comparison can be created by setting the `pairwisemax` argument to "FALSE". The code below produces Figure 14. We include only the summary plot with minimum p-value for each covariate. In this example, with very small samples and well balanced groups, there are no statistically significant differences between the "metcbt5" or the "scy" samples and the "community" sample.

```
%mnplot(inputobj=mnps.RData,
        plotname=mnps_example_plot_att_4.pdf,
        plotformat=pdf,
        multipage=TRUE,
        plots=t,
        Rcmd=C:\Program Files\R\R-3.0.2\bin\x64\R.exe,
        objpath=C:\Users\uname\twang_mnps_example)
```

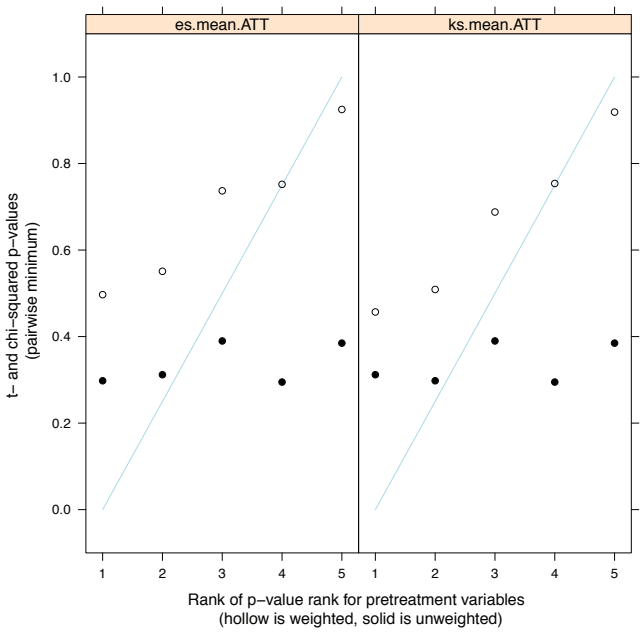

Figure 14: Example of a p-value plot for t-tests for both stopping rules (es.man and ks.max) for comparing the MET/CBT-5 and the SCY condition to the Community condition to generate ATT weights for the AOD dataset for a target population of those who received community treatment. Plot of the minimum p-value across both comparisons for each covariate.

3.3 Tabular assessments of balance

The `%mnps` macro prints default balance tables with the `estimand` equal to "ATT" like it did for "ATE". However, for ATT, it only reports pairwise comparisons that include the `treatATT` category. Only one summary table is created. It contains summary information for each comparison of a treatment to the target condition and for each stopping rule or unweighted. There

is one line in the table for each combination of the alternative treatment. Each record includes: `comp_treat`, which specifies the comparison treatment condition being weighted to match the target group, the `row.names`, which specifies the stopping rule used in calculating the weights or "unw" for unweighted, `n_treat` and `n_ctrl` for the sample sizes of the target population and the comparison group, respectively, `ess_treat` and `ess_ctrl`, which equal the effective sample sizes for each group, `max_es` and `mean_es`, the maximum and average absolute standardized effect sizes across the covariates, `max_ks`, `max_ks_p`, and `mean_ks`, the maximum KS statistic across covariates, the p-value testing this (if requested), and the average KS statistics across covariates, and `iter`, which equals the number of iterations used in the GBM model chosen by the stopping rule of the record. In this example, both the MET/CBT-5 and SCY condition samples are very similar to the Community condition sample prior to any weighting and consequently the balance is excellent after weighting.

SAS Output: Summary table for the ATT example

Summary table 09:10 Saturday, December 27, 2014 1

Summary of observations receiving each other treatment
weighted to match the observations receiving treatment community

Obs	comp_treat	row_name	n_treat	n_ctrl	ess_treat	ess_ctrl	max_es
1	metcbt5	unw	200	200	200	200	0.1088358309
2	metcbt5	es.mean.ATT	200	200	200	166.13229538	0.0726431842
3	metcbt5	ks.mean.ATT	200	200	200	165.52239545	0.070663584
4	scy	unw	200	200	200	200	0.1035526049
5	scy	es.mean.ATT	200	200	200	187.48228534	0.0592797363
6	scy	ks.mean.ATT	200	200	200	170.9089336	0.0759831448

Obs	mean_es	max_ks	max_ks_p	mean_ks	iter
1	0.0784516489	0.105	.	0.078	.
2	0.0257233329	0.0512612081	.	0.04182177	831
3	0.025121302	0.0509657576	.	0.0416737864	854
4	0.061971686	0.09	.	0.066	.
5	0.028475052	0.0691803042	.	0.0497268537	185
6	0.0302771643	0.0735519443	.	0.0420596509	576

The macro also prints the balance table for the individual covariates. The table includes the same statistics as the balance table for the `%ps` macro (`tx_mn`, `tx_sd`,`ct_mn`,`ct_sd`,`std_eff_sz`, `stat`,`p`,`ks`,`ks_pval` equal to the target treatment group mean and standard deviation, the comparison condition mean and standard deviation, the standardized effect size, the t-statistic testing the mean differences between groups and its associated p-value, the KS statistic and its p-value). It also includes the name of the covariate (`var`), the treatment group variable value for the comparison group (`control`) and the stop method (`stop_method`).

SAS Output: Balance table for individual covariates for the ATT example

Balance table: unw 09:10 Saturday, December 27, 2014 2

Obs	var	tx_mn	tx_sd	ct_mn	ct_sd	std_eff_sz
1	illact	0.097	1.045	0.007	1.035	0.087
2	crimjust	-0.065	1.05	0.037	1.038	-0.097
3	subprob	-0.06	0.965	0.026	1.019	-0.088
4	subdep	0.046	1.079	0.058	1.047	-0.011
5	white	0.16	0.368	0.2	0.401	-0.109
6	illact	0.097	1.045	0.12	0.963	-0.021
7	crimjust	-0.065	1.05	-0.174	1.028	0.104
8	subprob	-0.06	0.965	-0.013	0.972	-0.048
9	subdep	0.046	1.079	-0.058	0.964	0.096
10	white	0.16	0.368	0.175	0.381	-0.041

Obs	stat	p	ks	ks_pval	control	stop_ method
1	0.87	0.385	0.1	0.27	metcbt5	unw
2	-0.98	0.328	0.105	0.221	metcbt5	unw
3	-0.861	0.39	0.09	0.394	metcbt5	unw
4	-0.113	0.91	0.055	0.924	metcbt5	unw
5	-1.041	0.298	0.04	0.997	metcbt5	unw
6	-0.223	0.823	0.06	0.866	scy	unw
7	1.048	0.295	0.08	0.545	scy	unw
8	-0.481	0.631	0.09	0.394	scy	unw
9	1.012	0.312	0.085	0.466	scy	unw
10	-0.401	0.688	0.015	1	scy	unw

Balance table: ks.mean 3
09:10 Saturday, December 27, 2014

Obs	var	tx_mn	tx_sd	ct_mn	ct_sd	std_eff_sz
21	illact	0.097	1.045	0.086	1.023	0.011
22	crimjust	-0.065	1.05	-0.032	0.997	-0.032
23	subprob	-0.06	0.965	-0.062	0.988	0.002
24	subdep	0.046	1.079	0.057	1.048	-0.011
25	white	0.16	0.368	0.186	0.39	-0.071
26	illact	0.097	1.045	0.098	1.036	-0.001
27	crimjust	-0.065	1.05	-0.041	0.973	-0.023
28	subprob	-0.06	0.965	-0.018	0.979	-0.043
29	subdep	0.046	1.079	-0.036	0.994	0.076
30	white	0.16	0.368	0.163	0.37	-0.008

Obs	stat	p	ks	ks_pval	control	stop_ method

```
 21      0.102      0.919      0.042      0.995  metcbt5   ks.mean
 22     -0.313      0.754      0.051      0.959  metcbt5   ks.mean
 23      0.018      0.986      0.039      0.997  metcbt5   ks.mean
 24     -0.104      0.917      0.05       0.963  metcbt5   ks.mean
 25     -0.662      0.509      0.026      1      metcbt5   ks.mean
 26     -0.006      0.995      0.05       0.96   scy       ks.mean
 27     -0.235      0.814      0.039      0.998  scy       ks.mean
 28     -0.402      0.688      0.045      0.987  scy       ks.mean
 29      0.744      0.457      0.074      0.664  scy       ks.mean
 30     -0.077      0.939      0.003      1      scy       ks.mean
```

```
                      Balance table: es.mean                  4
                              09:10 Saturday, December 27, 2014
```

```
Obs  var          tx_mn      tx_sd      ct_mn      ct_sd   std_eff_sz

 11  illact        0.097      1.045      0.087      1.024      0.01
 12  crimjust     -0.065      1.05      -0.032      0.998     -0.032
 13  subprob      -0.06       0.965     -0.062      0.989      0.003
 14  subdep        0.046      1.079      0.058      1.049     -0.012
 15  white         0.16       0.368      0.187      0.391     -0.073
 16  illact        0.097      1.045      0.1        1.005     -0.002
 17  crimjust     -0.065      1.05      -0.064      0.995     -0.002
 18  subprob      -0.06       0.965     -0.027      0.967     -0.034
 19  subdep        0.046      1.079     -0.018      0.993      0.059
 20  white         0.16       0.368      0.176      0.382     -0.045
```

```
                                          stop_
Obs      stat        p          ks      ks_pval  control   method

 11      0.094      0.925      0.041      0.995  metcbt5   es.mean
 12     -0.317      0.752      0.051      0.957  metcbt5   es.mean
 13      0.025      0.98       0.039      0.998  metcbt5   es.mean
 14     -0.112      0.911      0.051      0.959  metcbt5   es.mean
 15      -0.68      0.497      0.027      1      metcbt5   es.mean
 16     -0.023      0.982      0.056      0.902  scy       es.mean
 17     -0.016      0.988      0.052      0.941  scy       es.mean
 18     -0.336      0.737      0.055      0.904  scy       es.mean
 19      0.596      0.551      0.069      0.707  scy       es.mean
 20     -0.433      0.665      0.016      1      scy       es.mean
```

As with the ATE condition, the `%mnbaltable` macro allows for printing balance tables that summarize across groups using the `collapseto` argument or the subset arguments (`subset_var`, `subset_treat`, `subset_stop_methods`) or restrict to covariates with that are not balanced using the `es_cutoff`, `p_cutoff`, `ks_cutoff`, and `ks_p_cutoff`.

3.4 Estimating treatment effects

The effects of interest are comparison of each of the treatments to Community care. We can estimate that effect in a single linear regression model with dummy indicator variables for the MET/CBT-5 and SCY samples or model for `suf12` with `treat` as a class variable regressor and "treat = community" as the holdout. However, by default SAS codes the last level of a class variable as the holdout condition. Consequently, we code our own dummy variables for MET/CBT-5 and SCY and use those in the model for `suf12` fit using PROC SURVEYREG just like we did for the ATE example.

```
data sasin.aodattwgt;
   set sasin.aodattwgt;
   community = treat = "community";
   metcbt5 = treat = "metcbt5";
   scy = treat = "scy";
run;

proc surveyreg data=sasin.aodattwgt;
   model suf12 = metcbt5 scy / solution;
   weight es_mean_ATT;
run;
```

SAS Output: ATT effect estimate for MET/CBT-5 or SCY versus Community

```
                ATT: MET/CBT-5 or SCY vs. Community                    2
                                      06:49 Friday, January 9, 2015

                      The SURVEYREG Procedure

              Regression Analysis for Dependent Variable suf12

                          Data Summary

                Number of Observations           600
                Sum of Weights             534.83326
                Weighted Mean of suf12      -0.02048
                Weighted Sum of suf12      -10.95464


                          Fit Statistics

                R-square          0.006640
                Root MSE            0.9889
                Denominator DF         599


                    Tests of Model Effects

           Effect       Num DF    F Value    Pr > F

           Model             2       1.85    0.1574
           Intercept         1       2.70    0.1009
```

```
metcbt5           1       3.71    0.0547
scy               1       0.66    0.4158

NOTE: The denominator degrees of freedom for the F tests is 599.


                    Estimated Regression Coefficients

                                Standard
  Parameter      Estimate         Error    t Value    Pr > |t|

  Intercept    -0.1050526    0.06393962      -1.64      0.1009
  metcbt5       0.2007108    0.10426163       1.93      0.0547
  scy           0.0807567    0.09917475       0.81      0.4158

NOTE: The denominator degrees of freedom for the t tests is 599.
```

Note in this case that the estimated treatment effect of community on those exposed to the community treatment is slightly stronger than in the ATE case (high numbers are bad for the outcome variable). Although not statistically significant, such differences are compatible with the notion that the youths who actually received the community treatment responded more favorably to it than the "average" youth would have (where the average is taken across the whole collection of youths enrolled in the study).

The discussion in McCaffrey et al. (2013) may be useful for determining whether the ATE or ATT is of greater interest in a particular application.

4 Conclusion

Often, more than two treatments are available to study participants. If the study is not randomized, analysts may be interested in using a propensity score approach. Previously, few tools existed to aide the analysis of such data, perhaps tempting analysts to ignore all but two of the treatment conditions. We hope that this extension to the `twang` package will encourage more appropriate analyses of observational data with more than two treatment conditions.

Acknowledgements

The random subset of data was supported by the Center for Substance Abuse Treatment (CSAT), Substance Abuse and Mental Health Services Administration (SAMHSA) contract # 270-07-0191 using data provided by the following grantees: Adolescent Treatment Model (Study: ATM: CSAT/SAMHSA contracts # 270-98-7047, # 270-97-7011, #2 77-00-6500, # 270-2003-00006 and grantees: TI-11894, TI-11874, TI-11892), the Effective Adolescent Treatment (Study: EAT; CSAT/SAMHSA contract # 270-2003-00006 and grantees: TI-15413, TI-15433, TI-15447, TI-15461, TI-15467, TI-15475, TI-15478, TI-15479, TI-15481, TI-15483, TI-15486, TI-15511, TI15514, TI-15545, TI-15562, TI-15670, TI-15671, TI-15672, TI-15674, TI-15678, TI-15682, TI-15686, TI-15415, TI-15421, TI-15438, TI-15446, TI-15458, TI-15466, TI-15469, TI-15485, TI-15489, TI-15524, TI-15527, TI-15577, TI-15584, TI-15586, TI-15677), and the Strengthening Communities-Youth (Study: SCY; CSAT/SAMHSA contracts # 277-00-6500, # 270-2003-00006 and grantees: TI-13305, TI-13308, TI-13313, TI-13322, TI-13323, TI-13344, TI-13345, TI-13354). The authors thank these grantees and their participants for agreeing to share their data to support the development of the `mnps` functionality.

References

[1] Burgette, L.F., D.F. McCaffrey, B.A. Griffin (forthcoming). "Propensity score estimation with boosted regression." In W. Pan and H. Bai (Eds.) *Propensity Score Analysis: Fundamentals, Developments and Extensions*. New York: Guilford Publications, Inc.

[2] McCaffrey, D.F., B.A. Griffin, D. Almirall, M.E. Slaughter, R. Ramchand, and L.F. Burgette (2013). "A tutorial on propensity score estimation for multiple treatments using generalized boosted models." *Statistics in Medicine*, *32*(19), 3388–3414.

[3] Ridgeway, G., D. McCaffrey, B.A. Griffin, and L. Burgette (2014). "twang: Toolkit for weighting and analysis of non-equivalent groups." Available at http://cran.r-project.org/web/packages/twang/vignettes/twang.pdf.