Web-based Supporting Materials for A Tutorial on Propensity Score Estimation for Multiple Treatments Using Generalized Boosted Models by Daniel F. McCaffrey, Beth Ann Griffin, Daniel Almirall, Mary Ellen Slaughter, Rajeev Ramchand, and Lane Burgette

1. Supplemental Tables

Table 1. Means for Treatment Groups (Unweighted and ATT Weighted)

	Community		ATT-Comm		MET/CBT		ATT-MET/CBT		SCY		ATT-SCY	
	Unweighted		Wgtd Means		Unweighted		Wgtd Means		Unweighted		Wgtd Means	
Covariate	Mean	SD	CBT	SCY	Mean	SD	Comm	SCY	Mean	SD	Comm	CBT
Demographics (Race/Ethnicity)												
% non-Hisp White	0.67	0.47	0.63	0.54*	0.51	0.5	0.59	0.48	0.35	0.48	0.47*	0.38
% Afr Amer (nonH)	0.14	0.34	0.09	0.18	0.08	0.27	0.11	0.09	0.32	0.47	0.18*	0.26
% Hispanic	0.10	0.3	0.17*	0.14	0.24	0.43	0.19	0.25	0.20	0.4	0.20	0.21
% Other	0.09	0.29	0.11	0.13	0.17	0.37	0.11	0.17	0.14	0.35	0.14	0.15
Substance Use												
In Recovery	0.20	0.4	0.23	0.15	0.25	0.43	0.26	0.25	0.18	0.39	0.18	0.20
Substance Freq	0.14	0.15	0.14	0.16	0.11	0.12	0.11	0.11	0.15	0.15	0.15	0.14
Subs Prob (month)	3.11	3.59	3.23	3.49	2.71	3.36	2.85	2.66	3.36	3.84	3.31	3.15
Subs Prob (year)	7.93	4.45	7.04*	7.91	6.30	4.27	6.50	6.14	7.09	4.49	7.52	6.82
Subs Dep (month)	1.02	1.62	1.11	1.21	0.89	1.56	0.78	0.89	1.2	1.83	1.08	1.08
Subs Dep (year)	3.10	2.33	2.67	3.02	2.33	2.22	2.42	2.23	2.69	2.34	3	2.57
Days Drunk/High	13.78	20.98	15.98	17.88	11.42	20.92	10.94	10.85	18.44	26.68	16.95	16.7
Criminal Activities												
Crime Violence	8.84	5.62	8.2	8.79	6.03	5.01	6.13	6.18	7.83	6.02	7.49	7.25
Illegal Act	0.22	0.19	0.2	0.21	0.09	0.11	0.11*	0.09	0.12	0.14	0.16*	0.11
Crim Just Sys Invlmt	0.52	0.47	0.40*	0.43	0.31	0.43	0.35	0.32	0.42	0.45	0.43	0.41
Days Instit.	13.60	22.7	10.29	11.81	4.44	15.43	2.64	4.23	10.93	19.94	7.4	9.06
Mental Health Functioning												
Intl Mental Distress	7.72	8.23	8.89	9.15	7.82	8.51	7.16	7.97	9.33	9.12	8.74	8.70
Problem Orient	0.99	1.71	0.92	1.06	0.69	1.46	0.61	0.67	1	1.74	1.03	0.85
Emotional Problems	0.28	0.19	0.26	0.27	0.21	0.18	0.23	0.21	0.24	0.2	0.26	0.23
Behavioral Complxty	12.26	7.4	12.03	12.61	9.81	7.78	10.56	9.5	10.89	8.43	11.59	10.78
Prev. Mntl Hlth Trt	0.45	0.50	0.47	0.44	0.36	0.48	0.36	0.39	0.41	0.49	0.41	0.39
Mental Health Trt	0.01	0.07	0.01	0.01	0.01	0.05	0.01	0.01	0.01	0.05	0.01	0.01
Environmental Risk												
Social Risk	13.84	5.36	12.8	13.65	12.31	4.24	13.51*	12.23	13.39	4.68	13.98	13.03
Crime Environ	0.24	0.36	0.12*	0.15*	0.05	0.18	0.05	0.05	0.13	0.23	0.12	0.11
Environ Risk	31.43	11.69	34.55*	34.43*	35.47	8.79	34.15	35.19	37.23	9.17	34.95*	36.84
Living house/apart	0.85	0.36	0.84	0.85	0.88	0.32	0.94	0.88	0.82	0.39	0.88	0.83
Living Jail	0.11	0.31	0.05	0.07	0.05	0.23	0.03	0.04	0.06	0.23	0.06	0.06
Other Living Sit.	0.04	0.2	0.11*	0.08	0.06	0.24	0.03	0.07	0.13	0.33	0.06	0.11
Training Activity	0.55	0.29	0.62*	0.57	0.66	0.23	0.67	0.66	0.57	0.28	0.59	0.59
N or ESS (Weighted)	444		379.96	199.17	2459		42.19	768.92	1351		87.09	1108.47

Table 2. Unweighted and parametric and nonparametric weighted maximum absolute standardized mean differences for each pretreatment covariate in our case study data when interest lies in estimating pairwise ATEs.

		Multinomial Logistic			
Covariate	Unweighted	Raw Weights	Trimmed Weights	Dropping Extreme Weights	GBM
Demographics (Race/Ethnicity)					
Percent non-Hispanic White	0.39	0.8	0.09	0.25	0.22
Percent African American (nonH)	0.43	2.01	0.04	0.02	0.12
Percent Hispanic	0.28	0.46	0.01	0.22	0.04
Percent Other	0.17	0.41	0.16	0.11	0.15
Substance Use					
In Recovery	0.11	0.5	0.11	0.12	0.09
Substance Frequency Scale	0.21	0.13	0.32	0.24	0.07
Subs Prob Scale (past month)	0.13	1.88	0.17	0.03	0.03
Subs Prob Scale (past year)	0.26	1.56	0.35	0.18	0.14
Subs Dep Scale (past month)	0.13	2.02	0.16	0.05	0.05
Subs Dep Scale (past year)	0.24	1.63	0.36	0.15	0.13
Days Drunk/High (past 90 days)	0.21	0.27	0.2	0.08	0.04
Criminal Activities					
Crime Violence Scale	0.38	1.61	0.44	0.24	0.13
Illegal Activities Scale	0.81	0.66	0.56	0.53	0.43
Criminal Justice System Invlmt	0.37	0.59	0.29	0.13	0.18
Days Institutionalized	0.35	4.15	0.19	0.09	0.07
Mental Health Functioning					
Internal Mental Distress Scale	0.15	1.26	0.27	0.11	0.08
Problem Orientation Scale	0.14	1.82	0.16	0.02	0.046
Emotional Problems Scale	0.28	0.97	0.32	0.14	0.13
Behavioral Complexity Scale	0.22	1.78	0.44	0.16	0.11
Previous Mental Health Trt	0.12	1.11	0.29	0.11	0.05
Mental Health Trt Scale	0.04	1.39	0.41	0.07	0.03
Environmental Risk					
Social Risk Scale	0.22	2.29	0.24	0.29	0.19
Controlled Environment Scale	0.64	3.59	0.23	0.18	0.24
Environmental Risk Scale	0.46	0.12	0.28	0.07	0.33
Living in house/apartment	0.13	2.2	0.06	0.06	0.06
Living in Jail/Correctional Facility	0.23	0.24	0.08	0.04	0.04
Other Living Situation	0.19	3.1	0.07	0.09	0.09
Training Activity Scale	0.29	0.07	0.18	0.1	0.17

Table 3. Unweighted and parametric and nonparametric weighted maximum absolute standardized mean differences for each pretreatment covariate in our case study data when interest lies in estimating pairwise ATTs, shown for each type of ATT in the case study

	Community ATT			MET/CBT-5 ATT			SCY ATT		
Covariate	Unwt	ML	GBM	Unwt	ML	GBM	Unwt	ML	GBM
Demographics (Race/Ethnicity)									
Percent non-Hispanic White	0.49	0.10	0.24	0.34	0.14	0.12	0.23	0.20	0.04
Percent African American (nonH)	0.27	0.13	0.15	0.21	0.14	0.14	0.15	0.20	0.02
Percent Hispanic	0.69	0.26	0.28	0.33	0.13	0.16	0.67	0.12	0.26
Percent Other	0.52	0.17	0.14	0.87	0.17	0.09	0.51	0.14	0.28
Substance Use									
In Recovery	0.10	0.10	0.13	0.17	0.16	0.01	0.19	0.03	0.05
Substance Frequency Scale	0.26	0.07	0.04	0.38	0.40	0.12	0.31	0.19	0.09
Subs Prob Scale (past month)	0.12	0.04	0.10	0.21	0.15	0.07	0.19	0.16	0.06
Subs Prob Scale (past year)	0.36	0.14	0.16	0.37	0.33	0.12	0.19	0.36	0.13
Subs Dep Scale (past month)	0.11	0.04	0.07	0.21	0.10	0.05	0.18	0.22	0.07
Subs Dep Scale (past year)	0.33	0.12	0.16	0.34	0.32	0.09	0.17	0.40	0.15
Days Drunk/High (past 90 days)	0.23	0.12	0.11	0.35	0.25	0.03	0.27	0.10	0.08
Criminal Activities									
Crime Violence Scale	0.52	0.12	0.10	0.59	0.43	0.13	0.31	0.48	0.10
Illegal Activities Scale	0.71	0.21	0.11	1.22	0.74	0.44	0.73	0.54	0.43
Criminal Justice System Invlmt	0.47	0.21	0.27	0.52	0.24	0.12	0.25	0.32	0.16
Days Institutionalized	0.42	0.33	0.32	0.66	0.19	0.04	0.36	0.15	0.11
Mental Health Functioning									
Internal Mental Distress Scale	0.23	0.07	0.10	0.23	0.13	0.12	0.21	0.44	0.08
Problem Orientation Scale	0.17	0.08	0.12	0.25	0.12	0.02	0.21	0.23	0.08
Emotional Problems Scale	0.39	0.10	0.07	0.41	0.29	0.13	0.19	0.37	0.10
Behavioral Complexity Scale	0.33	0.06	0.03	0.32	0.31	0.09	0.15	0.60	0.13
Previous Mental Health Trt	0.17	0.04	0.05	0.18	0.24	0.06	0.12	0.34	0.05
Mental Health Trt Scale	0.04	0.06	0.05	0.07	0.25	0.06	0.07	0.70	0.00
Environmental Risk									
Social Risk Scale	0.28	0.08	0.17	0.35	0.38	0.34	0.24	0.07	0.16
Controlled Environment Scale	0.53	0.42	0.44	1.17	0.25	0.08	0.46	0.19	0.19
Environmental Risk Scale	0.52	0.18	0.28	0.47	0.19	0.25	0.66	0.38	0.35
Living in house/apartment	0.13	0.18	0.03	0.23	0.04	0.16	0.18	0.11	0.17
Living in Jail/Correctional Facility	0.19	0.16	0.17	0.27	0.08	0.09	0.26	0.13	0.01
Other Living Situation	0.42	0.15	0.33	0.31	0.04	0.13	0.25	0.04	0.19
Training Activity Scale	0.38	0.26	0.20	0.45	0.17	0.03	0.26	0.18	0.05

2. Example Code for Estimating Pairwise ATE and ATT

This section provides example code for estimating ATE and ATT using a simulated data set with three treatments. We provide code for SAS, R, and Stata. We do not provide code for estimating the propensity scores. The `twang` package in R provides functions for propensity score estimation using GBM and for checking the balance with multiple treatments using the metrics provided in the paper.

2.1. SAS Code

```
*************************************************************************************;
** The example uses the data set dsname.csv which is available upon request from   **;
** the authors. The data are simulated to demonstrate syntax.                      **;
**                                                                                 **;
**  The data set contains 600 observations for "youth" in three treatment groups:  **;
**  community, MET/CBT-5, and SCY.  The data include an outcome variable substance **;
**  use frequency at the 12 month follow-up, covariates, indicators for the treat- **;
**  ment groups, and ATE and ATT weights.                                          **;
**  The variables are:                                                             **;
**     community -- indicator for being in the community care treatment group      **;
**     metcbt5 -- indicator for being in the MET/CBT-5 treatment group             **;
**     scy -- indicator for being in the SCY treatment group                       **;
**     atewgt -- ATE weight variable                                               **;
**     attwgt_com_met -- ATT weight for community vs. MET/CBT-5, community target  **;
**     attwgt_com_scy -- ATT weight for community vs. SCY, community target        **;
**     attwgt_met_com -- ATT weight for MET/CBT-5 vs. community, MET/CBT-5 target  **;
**     attwgt_met_scy -- ATT weight for MET/CBT-5 vs. SCY, MET/CBT-5 target        **;
**     attwgt_scy_com -- ATT weight for SCY vs. community, SCY target              **;
**     attwgt_scy_met -- ATT weight for SCY vs. MET/CBT-5, SCY target              **;
**     suf12 -- outcome variable, substance use frequency at 12 month follow-up   **;
**     illact -- covariate, illicit activities scale                               **;
**     crimjust -- covariate, criminal justice involvement                         **;
**     subprob -- covariate, substance use problem scale                           **;
**     subdep -- covariate, substance use dependence scale                         **;
**     white -- covariate, indicator for white non-Hispanic youth                  **;
*************************************************************************************;

   **  Read in the example data set. **;

proc import file="dsname.csv" out=dsname dbms=csv replace;
run;

*************************************************************************************;
**                              ATE Example SAS Code                               **;
** We assume there are three treatmeant groups, community, MET/CBT-5, and SCY. The **;
** data contain indicator variables for each treatment (community, metctb5, scy).  **;
** They also include an ATE weight variable, atewgt, that equals the reciprocal    **;
** of the propensity score for the treatment each youth actually received, e.g.,   **;
** the propensity score for community for youth who received community care, etc.  **;
** The data also include an outcome variable, substance use frequency at the 12-   **;
** month follow-up, suf12, and five covariates that will be used for doubly robust **;
** estimation: illact (illict activities), crimjust (criminal justice system       **;
** involvement), subprob (substance use problems), subdep (substance use           **;
** dependence), and white (an indicator for white/non-Hispanic youth).             **;
** The data set is named dsname.                                                   **;
** The code estimates treatment effects and group means using a regression model   **;
** without covariate adjustment and with the additional covariate adjustment.      **;
*************************************************************************************;
```

```
   ** Estimate ATE without covariate adjustment -- include indicators for   **;
   ** MET/CBT-5 and SCY. The coefficients on the indicators estimate ATE of **;
   ** MET/CBT-5 relative to community and SCY relative to community. Their  **;
   ** difference estimates the ATE of MET/CBT-5 relative to SCY. The        **;
   ** population means for the potential outcomes are estimated by the      **;
   ** intercept (community) and the sum of the intercept and one the of     **;
   ** indicators (MET/CBT-5 and SCY).                                       **;

   ** PROC SURVEYREG uses standard error estimates that account for the     **;
   ** weighting.                                                            **;

proc surveyreg data=dsname;
   model suf12 = metcbt5 scy;
   weight atewgt;
   estimate 'ATE MET/CBT-5 vs SCY' metcbt5 1 scy -1;    ** Estimates the ATE for
                                                           MET/CBT-5 vs SCY ;
   estimate 'Pop mean MET/CBT-5' intercept 1 metcbt5 1; ** Estimates the population
                                                           mean of MET/CBT-5
                                                           potential outcomes ;
   estimate 'Pop mean SCY'       intercept 1 scy 1;     ** Estimates the population
                                                           mean of SCY potential
                                                           outcomes;
title "Estimate ATE, no covariates";
run;

  ************************************************************************;
  ** Estimate ATE with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;

  ** Centering the covariates at the unweighted mean of the overall sample  **;
  ** will facilitate the estimation of the counterfactual means for the     **;
  ** population had all youth received each of the three treatments.        **;

  ** We use proc standard to mean center the covariates in a temporary      **;
  ** data set to avoid any confusion about how variables are centered for   **;
  ** later analyses in this example.                                       **;

proc standard data=dsname out=dtemp mean=0;  ** dtemp is the temporary
                                                dataset and mean=0 is the
                                                syntax for mean centering.  ;
    var illact crimjust subprob subdep white;
run;

proc surveyreg data=dtemp;
   model suf12 = metcbt5 scy illact crimjust subprob subdep white;
   weight atewgt;
   estimate 'ATE MET/CBT-5 vs SCY' metcbt5 1 scy -1;    ** Estimates the ATE for
                                                           MET/CBT-5 vs SCY ;
   estimate 'Pop mean MET/CBT-5' intercept 1 metcbt5 1; ** Estimate the population
                                                           mean of MET/CBT-5
                                                           potential outcomes ;
   estimate 'Pop mean SCY'       intercept 1 scy 1;     ** Estimate the population
                                                           mean of SCY potential
                                                           outcomes ;
title "Estimate ATE, with covariates -- doubly robust";
run;
```

```
  ** The coefficients on metcbt5 and scy estimate the pairwise ATE compared **;
  ** with community for MET/CBT-5 and SCY respectively.                     **;

  ** Because we centered the covariates at the overall unweighted mean,     **;
  ** the intercept estimates the population mean had all youth received     **;
  ** community care, the intercept plus the coefficient on MET/CBT-5        **;
  ** estimates the mean had all youth received MET/CBT-5, and the intercept **;
  ** plus the coefficient on SCY estimates the mean had all youth received  **;
  ** SCY.  By centering the covariates, the coefficients have the same      **;
  ** interpretation for models with or without covariate adjustment.        **;


*************************************************************************************;
**                              ATT Example SAS Code                                 **;
** The example assumes the same data as the ATE example plus six ATT weights.      **;
** attwgt_com_met -- the ATT weight for estimating the ATT of community vs         **;
**                   MET/CBT-5 for youth like those receiving community care       **;
** attwgt_com_scy -- the ATT weight for estimating the ATT of community vs SCY     **;
**                   for youth like those receiving community care                 **;
** attwgt_met_com -- the ATT weight for estimating the ATT of MET/CBT-5 vs         **;
**                   community for youth like those receiving MET/CBT-5            **;
** attwgt_met_scy -- the ATT weight for estimating the ATT of MET/CBT-5 vs         **;
**                   scy for youth like those receiving MET/CBT-5                  **;
** attwgt_scy_com -- the ATT weight for estimating the ATT of SCY vs community     **;
**                   for youth like those receiving SCY                            **;
** attwgt_scy_met -- the ATT weight for estimating the ATT of SCY vs MET/CBT-5     **;
**                   for youth like those receiving SCY                            **;
**                                                                                 **;
** The procedures are similar to ATE but they are restricted to subsamples of the  **;
** full data set defined by pairs of treatment                                     **;
*************************************************************************************;

   ** Estimate ATT. We start with code for community vs. MET/CBT-5 and SCY  **;
   ** and then repeat for using the other treatments                        **;

   ** %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**;
   ** %%%%%%%%%%%%%%%%  ATT with Community as the target %%%%%%%%%%%%%%%%%%%% **;
   ** %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**;

   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;
   ** Select the community and MET/CBT-5 subsample                            **;
   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;

data com_met;
   set dsname;
   if community = 1 or metcbt5 = 1;
run;

proc surveyreg data=com_met;
   model suf12 = community;
   weight attwgt_com_met;
title "ATT Community vs. MET/CBT-5 -- Community Target";
run;

  ** The coefficient on community estimates the ATT, the mean for youth     **;
  ** like those in community had they received community care minus the     **;
```

```
  ** counterfactual mean had they received MET/CBT-5.                    **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the community group had they received MET/CBT-5.  The unweighted mean  **;
  ** of the community group estimates the mean for youth like those         **;
  ** receiving community care when they received community care.            **;

proc means data=dsname mean;
    where community = 1;        ** Select only the community sample **;
    var suf12;
run;

  ************************************************************************;
  ** Estimate ATT with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;

  ** If we subtract the unweighted mean for the community group from each   **;
  ** covariate, then the intercept of the model with covariates estimates   **;
  ** the counterfactual mean for the community group had it received        **;
  ** MET/CBT-5. The coefficient has the same interpretation as it did in    **
  ** the model without covariates.                                          **;

  ** Obtain community group means.                                          **;

proc summary data=dsname nway;
   where community = 1;           ** Select the community sample **;
   var illact crimjust subprob subdep white;;
   output out=commean mean=illact_m crimjust_m subprob_m subdep_m white_m;
run;

  ** Subtract them from the covariates.                                    **;

data com_met;
   set com_met;
   if _n_ = 1 then set commean;
   array vars illact crimjust subprob subdep white;
   array means illact_m crimjust_m subprob_m subdep_m white_m;
   do over vars;
      vars = vars - means;
      end;
run;

  ** Fit the model with the centered covariates.                           **;

proc surveyreg data=com_met;
   model suf12 = community illact crimjust subprob subdep white;;
   weight attwgt_com_met;
run;

  ** The coefficient on community estimates the ATT, the mean for youth     **;
  ** like those in community had they received community care minus the     **;
  ** counterfactual mean had they received MET/CBT-5                        **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the community group had they received MET/CBT-5. The unweighted mean   **;
  ** of the community group (calculated previously) estimates the mean for  **;
  ** youth like those receiving community care when they received community **;
```

```
  ** care.                                                                **;


   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;
   ** Select the community and SCY subsample                             **;
   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;

data com_scy;
   set dsname;
   if community = 1 or scy = 1;
run;

proc surveyreg data=com_scy;
   model suf12 = community;
   weight attwgt_com_scy;
title "ATT Community vs. SCY -- Community Target";
run;

  ** The coefficient on community estimates the ATT, the mean for youth     **;
  ** like those in community had they received community care minus the     **;
  ** counterfactual mean had they received SCY.                            **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the community care group had they received SCY.  The unweighted mean   **;
  ** of the community care group (calculated previously) estimates the mean **;
  ** for youth like those receiving community care when they received       **;
  ** community care.                                                       **;

  ************************************************************************;
  ** Estimate ATT with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;


  ** Subtract community care means from the covariates.                    **;

data com_scy;
   set com_scy;
   if _n_ = 1 then set commean;
   array vars illact crimjust subprob subdep white;
   array means illact_m crimjust_m subprob_m subdep_m white_m;
   do over vars;
      vars = vars - means;
      end;
run;

  ** Fit the model with the centered covariates.                           **;

proc surveyreg data=com_scy;
   model suf12 = community illact crimjust subprob subdep white;;
   weight attwgt_com_scy;
run;

  ** The coefficient on community estimates the ATT, the mean for youth     **;
  ** like those in community had they received community care minus the     **;
  ** counterfactual mean had they received SCY.                            **;

  ** The intercept from the model estimates the counterfactual mean for     **;
```

```
  ** the community care group had they received SCY.  The unweighted mean   **;
  ** of the community care group (calculated previously) estimates the mean **;
  ** for youth like those receiving community care when they received       **;
  ** community care.                                                        **;


   ** %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**;
   ** %%%%%%%%%%%%%%%%  ATT with MET/CBT-5 as the target %%%%%%%%%%%%%%%%%%% **;
   ** %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**;

   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;
   ** Select the MET/CBT-5 and community subsample                           **;
   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;

data met_com;
   set dsname;
   if community = 1 or metcbt5 = 1;
run;

proc surveyreg data=met_com;
   model suf12 = metcbt5;
   weight attwgt_met_com;
title "ATT  MET/CBT-5 vs. Community -- MET/CBT-5 Target";
run;

  ** The coefficient on metcbt5 estimates the ATT, the mean for youth like  **;
  ** those in MET/CBT-5 had they received MET/CBT-5 minus the counter-      **;
  ** factual mean had they received community care.                         **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** in the MET/CBT-5 group had they received community care.  The          **;
  ** unweighted mean of the MET/CBT-5 group estimates the mean for youth    **;
  ** like those receiving MET/CBT-5 when they received MET/CBT-5.           **;

proc means data=dsname mean;
    where metcbt5 = 1;        ** Select only the MET/CBT-5 sample **;
    var suf12;
run;

  ************************************************************************;
  ** Estimate ATT with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;

  ** If we subtract the unweighted mean for the MET/CBT-5 group from each   **;
  ** covariate, then the intercept of the model with covariates estimates   **;
  ** the counterfactual mean for the MET/CBT-5 group had it received com-   **;
  ** munity care. The coefficient has the same interpretation as it did in  **;
  ** the model without covariates.                                          **;

  ** Obtain MET/CBT-5 group means.                                          **;

proc summary data=dsname nway;
   where metcbt5 = 1;           ** Select the MET/CBT-5 sample **;
   var illact crimjust subprob subdep white;;
   output out=metmean mean=illact_m crimjust_m subprob_m subdep_m white_m;
run;
```

```
  ** Subtract MET/CBT-5 means from the covariates.                          **;

data met_com;
   set met_com;
   if _n_ = 1 then set metmean;
   array vars illact crimjust subprob subdep white;
   array means illact_m crimjust_m subprob_m subdep_m white_m;
   do over vars;
      vars = vars - means;
      end;
run;

  ** Fit the model with the centered covariates.                           **;

proc surveyreg data=met_com;
   model suf12 = metcbt5 illact crimjust subprob subdep white;;
   weight attwgt_met_com;
run;

  ** The coefficient on metcbt5 estimates the ATT, the mean for youth like  **;
  ** those in MET/CBT-5 had they received MET/CBT-5 minus the counter-      **;
  ** factual mean had they received community care.                         **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the MET/CBT-5 group had they received community care.  The unweighted  **;
  ** mean of the MET/CBT-5 group mean (calculated previously) estimates the **;
  ** mean for youth like those receiving MET/CBT-5 when the received        **;
  ** MET/CBT-5.                                                             **;


   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;
   ** Select the MET/CBT-5 and SCY subsample                               **;
   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;

data met_scy;
   set dsname;
   if metcbt5 = 1 or scy = 1;
run;

proc surveyreg data=met_scy;
   model suf12 = metcbt5;
   weight attwgt_met_scy;
title "ATT MET/CBT-5 vs. SCY -- MET/CBT-5 Target";
run;

  ** The coefficient on metcbt5 estimates the ATT, the mean for youth like  **;
  ** those in MET/CBT-5 had they received MET/CBT-5 minus the counter-      **;
  ** factual mean had they received SCY.                                    **;

  ** The intercept from the model estimates the counterfactual mean for the **;
  ** MET/CBT-5 group had they received SCY. The unweighted mean of the      **;
  ** MET/CBT-5 group (calculated previously) estimates the mean for youth   **;
  ** like those receiving MET/CBT-5 when they received MET/CBT-5.           **;

  ************************************************************************;
  ** Estimate ATT with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;
```

```
  ** Subtract MET/CBT-5 means from the covariates.                          **;

data met_scy;
   set met_scy;
   if _n_ = 1 then set metmean;
   array vars illact crimjust subprob subdep white;
   array means illact_m crimjust_m subprob_m subdep_m white_m;
   do over vars;
      vars = vars - means;
      end;
run;

  ** Fit the model with the centered covariates.                            **;

proc surveyreg data=met_scy;
   model suf12 = metcbt5 illact crimjust subprob subdep white;;
   weight attwgt_met_scy;
run;

  ** The coefficient on metcbt5 estimates the ATT, the mean for youth like  **;
  ** those in MET/CBT-5 had they received MET/CBT-5 minus the counter-      **;
  ** factual mean had they received SCY.                                    **;

  ** The intercept from the model estimates the counterfactual mean for the **;
  ** MET/CBT-5 group had they received SCY. The  unweighted mean of the     **;
  ** MET/CBT-5 group mean (calculated previously) estimates the mean for    **;
  ** youth like those receiving MET/CBT-5 when they received MET/CBT-5.     **;


   ** %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**;
   ** %%%%%%%%%%%%%%%     ATT with SCY as the target    %%%%%%%%%%%%%%%%%% **;
   ** %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**;

   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;
   ** Select the SCY and community subsample                                **;
   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;

data scy_com;
   set dsname;
   if community = 1 or scy = 1;
run;

proc surveyreg data=scy_com;
   model suf12 = scy;
   weight attwgt_scy_com;
title "ATT  SCY vs. Community -- SCY Target";
run;

  ** The coefficient on scy estimates the ATT, the mean for youth like      **;
  ** those in SCY had they received SCY minus the counterfactual mean had   **;
  ** they received community care.                                          **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the SCY group had they received community care.  The unweighted mean   **;
  ** of the SCY group estimates the mean for youth like those receiving SCY **;
  ** when they received SCY.                                                **;
```

```
proc means data=dsname mean;
    where scy = 1;        ** Select only the SCY sample **;
    var suf12;
run;

  ************************************************************************;
  ** Estimate ATT with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;

  ** If we subtract the unweighted mean for the SCY group from each         **;
  ** covariate, then the intercept of the model with covariates estimates   **;
  ** the counterfactual mean for the SCY group had it received community    **;
  ** care. The coefficient has the same interpretation as it did in the     **;
  ** model without covariates.                                              **;

  ** Obtain SCY group means.                                                **;

proc summary data=dsname nway;
   where scy = 1;           ** Select the SCY sample **;
   var illact crimjust subprob subdep white;;
   output out=scymean mean=illact_m crimjust_m subprob_m subdep_m white_m;
run;

  ** Subtract SCY means from the covariates.                                **;

data scy_com;
   set scy_com;
   if _n_ = 1 then set scymean;
   array vars illact crimjust subprob subdep white;
   array means illact_m crimjust_m subprob_m subdep_m white_m;
   do over vars;
      vars = vars - means;
      end;
run;

  ** Fit the model with the centered covariates.                            **;

proc surveyreg data=scy_com;
   model suf12 = scy illact crimjust subprob subdep white;;
   weight attwgt_scy_com;
run;

  ** The coefficient on scy estimates the ATT, the mean for youth like      **;
  ** those in SCY had they received SCY minus the counterfactual mean had   **;
  ** they received community care.                                          **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the SCY group had they received community care. The unweighted         **;
  ** the SCY group mean (calculated previously) estimates the mean for      **;
  ** youth like those receiving SCY when they received SCY.                 **;

   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;
   ** Select the MET/CBT-5 and SCY subsample                              **;
   ** ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++ **;

data scy_met;
```

```
   set dsname;
   if metcbt5 = 1 or scy = 1;
run;

proc surveyreg data=scy_met;
   model suf12 = scy;
   weight attwgt_scy_met;
title "ATT SCY vs. MET/CBT-5 -- SCY Target";
run;

  ** The coefficient on scy estimates the ATT, the mean for youth like      **;
  ** those in SCY had they received SCY minus the counterfactual mean had   **;
  ** they received MET/CBT-5.                                               **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the SCY group had they received MET/CBT-5. The unweighted SCY group    **;
  ** mean (calculated previously) estimates the mean for youth like those   **;
  ** receiving SCY when they received SCY.                                  **;

  ************************************************************************;
  ** Estimate ATT with weighting and covariate adjustment -- doubly robust  **;
  ************************************************************************;

  ** Subtract SCY means from the covariates.                                **;

data scy_met;
   set scy_met;
   if _n_ = 1 then set scymean;
   array vars illact crimjust subprob subdep white;
   array means illact_m crimjust_m subprob_m subdep_m white_m;
   do over vars;
      vars = vars - means;
      end;
run;

  ** Fit the model with the centered covariates.                           **;

proc surveyreg data=scy_met;
   model suf12 = scy illact crimjust subprob subdep white;;
   weight attwgt_scy_met;
run;

  ** The coefficient on scy estimates the ATT, the mean for youth like      **;
  ** those in SCY had they received SCY minus the counterfactual mean had   **;
  ** they received MET/CBT-5.                                               **;

  ** The intercept from the model estimates the counterfactual mean for     **;
  ** the SCY group had they received MET/CBT-5. The unweighted SCY group    **;
  ** mean (calculated previously) estimates the mean for youth like those   **;
  ** receiving SCY when they received SCY.                                  **;
```

2.2. *R Code*

```
  ## Read in the data, dsname.csv. described in SAS code        ##

dsname <- read.csv("dsname.csv")
```

```
  ## Load survey package -- this package does weighted modeling ##
  ## and estimates standard errors that account for weighting.  ##
  ## It treats the weights as known rather than estimated from  ##
  ## the data.                                                  ##
library(survey)


  ## +++++++++++++++++++++++++ ATE ++++++++++++++++++++++++++++ ##
  ## ATE Estimates.                                             ##

  ## Create a survey design object to use with the survey       ##
  ## package linear regression function. Use the ATE weights    ##
  ## atewgt.                                                    ##

dsd <- svydesign(id=~1, weights=~atewgt, data=dsname)

ates <- svyglm(suf12 ~ metcbt5 + scy, design=dsd)

summary(ates)

  ## The coefficients on metcbt5 and scy estimate the pairwise  ##
  ## ATE compared with community for MET/CBT-5 and SCY           ##
  ## respectively. Their difference estimates the ATE of SCY    ##
  ## relative to MET/CBT-5.                                     ##

  ## svycontrast is a function for estimating linear            ##
  ## combinations of model coefficients.  The function returns  ##
  ## estimates and standard errors. It does not conduct a test  ##
  ## of the null hypothesis. We use the code below to conduct   ##
  ## the test.                                                  ##

ate_met_scy <- data.frame(svycontrast(ates, quote(metcbt5 - scy)))
names(ate_met_scy) <- c("Estimate", "Std. Error")
ate_met_scy$"t value" <- ate_met_scy$Estimate/ate_met_scy$"Std. Error"
ate_met_scy$"Pr(>|t|)" <- 2*(1-pt(abs(ate_met_scy$"t value"),
                                  df=summary(ates)$df.residual))

  ## Combine the three ATE estimates.                           ##

ates_all <- rbind(summary(ates)$coef[-1,], ate_met_scy)
rownames(ates_all) <- c("MET/CBT-5 vs. community",
                        "SCY vs. community", "MET/CBT-5 vs. SCY")

ates_all

  ## Estimate the means.                                        ##
  ## The intercept estimates the mean for community.            ##

com_mean <- summary(ates)$coef[1,]

  ## The estimated means for the other two treatments equal the ##
  ## the coefficient on the indicator plus the intercept.  We   ##
  ## again use svycontrast to estimate the linear combinations  ##
  ## of coefficients and their standard errors.                 ##

pop_means <- data.frame(svycontrast(ates,
```

```
                                        list("MET/CBT-5"=c(1,1,0), SCY=c(1,0,1))))
names(pop_means) <- c("Estimate", "Std. Error")
pop_means$"t value" <- pop_means$Estimate/pop_means$"Std. Error"
pop_means$"Pr(>|t|)" <- 2*(1-pt(abs(pop_means$"t value"),
                                df=summary(ates)$df.residual))

   ## Combine all three means.                                          ##
pop_means <- rbind(community=com_mean, pop_means)

pop_means

  ######################################################################
  ## Repeat with model that includes covariates--doubly robust         ##
  ## estimation.                                                       ##
  ######################################################################

  ## Centering the covariates at the unweighted mean of the            ##
  ## overall sample will facilitate the estimation of the              ##
  ## counterfactual means for the population had all youth             ##
  ## received each of the three treatments.                            ##

  ## The scale function will center variables at their means.          ##

covs <- c("illact", "crimjust", "subprob", "subdep", "white")
dtemp <- dsname
dtemp[,covs] <- scale(dtemp[,covs], center=T, scale=F)

  ## Create a svydesign with the centered data.                        ##
dsdtmp <- svydesign(id=~1, weights=~atewgt, data=dtemp)

ates <- svyglm(suf12 ~ metcbt5 + scy + illact + crimjust +
                       subprob + subdep + white, design=dsd)

summary(ates)

  ## By centering the variables the coefficients have the same         ##
  ## interpretation for models with and without covariates. We         ##
  ## can use the code used for the models without covariates           ##
  ## to estimates the ATE and the population means from the            ##
  ## model with covariates.                                            ##

ate_met_scy <- data.frame(svycontrast(ates, quote(metcbt5 - scy)))
names(ate_met_scy) <- c("Estimate", "Std. Error")
ate_met_scy$"t value" <- ate_met_scy$Estimate/ate_met_scy$"Std. Error"
ate_met_scy$"Pr(>|t|)" <- 2*(1-pt(abs(ate_met_scy$"t value"),
                                  df=summary(ates)$df.residual))

  ## Combine the three ATE estimate.                                   ##

ates_all <- rbind(summary(ates)$coef[2:3,], ate_met_scy)
rownames(ates_all) <- c("MET/CBT-5 vs. community",
                        "SCY vs. community", "MET/CBT-5 vs. SCY")

ates_all

  ## Estimate the means.                                               ##
  ## The intercept estimates the mean for community.                   ##
```

```
com_mean <- summary(ates)$coef[1,]

  ## The estimated means for the other two treatments equal the ##
  ## the coefficients on the indicators plus the intercept. We  ##
  ## again use svycontrast to estimate the linear combinations  ##
  ## of coefficients and their standard errors                  ##

pop_means <- data.frame(svycontrast(ates, list("MET/CBT-5"=c(1,1,0,0,0,0,0,0),
                                               SCY=c(1,0,1,0,0,0,0,0))))
names(pop_means) <- c("Estimate", "Std. Error")
pop_means$"t value" <- pop_means$Estimate/pop_means$"Std. Error"
pop_means$"Pr(>|t|)" <- 2*(1-pt(abs(pop_means$"t value"), df=summary(ates)$df.residual))

   ## Combine all three means                                   ##
pop_means <- rbind(community=com_mean, pop_means)

pop_means



  ## ++++++++++++++++++++++++++ ATT ++++++++++++++++++++++++++ ##

  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## ATT community vs. MET/CBT-5 for youth like those receiving ##
  ## community.                                                 ##
  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##

  ## Create an svydesign that includes only community and       ##
  ## MET/CBT-5                                                  ##

  ## If we center the covariates at the mean for the target     ##
  ## population, community, then coefficients have the same     ##
  ## interpretation in models with and without covariates and   ##
  ## we can use the same code to estimate population means. We  ##
  ## use this approach for all the ATT estimates.               ##

  ## We cannot use scale to center at the target mean.  The     ##
  ## following code conducts the centering.                     ##

dsd_com_met <- subset(dsname, subset=(community==1 | metcbt5 == 1))
cmeans <- apply(dsname[dsname$community==1, covs], 2, mean)
dsd_com_met[,covs] <- t(t(dsd_com_met[,covs])-cmeans)

  ## Create the svydesign for ATT.                              ##

dsd_com_met <- svydesign(id=~1, weights=~attwgt_com_met, data=dsd_com_met)

  ## Estimate the ATT.                                          ##

att_com_met <- svyglm(suf12 ~ community, design=dsd_com_met)

summary(att_com_met)

  ## The coefficient on community estimates the ATT for com-    ##
  ## munity care vs. MET/CBT-5 for youth like those in the com- ##
  ## munity sample.                                             ##
```

```
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving community care had they received      ##
  ## MET/CBT-5.                                                 ##

met_mean <- summary(att_com_met)$coef[1,]

  ## The unweighted mean for the community group estimates the  ##
  ## mean for youth like those in the community group when they ##
  ## received community care.                                   ##

  ## t.test estimates the mean and gives the test statistics.   ##
  ## It does not give the standard errors so we back it out.    ##

com_mean <- t.test(dsname$suf12[dsname$community==1])
com_mean <- com_mean[c("estimate","statistic","p.value")]
com_mean$se <- com_mean$estimate/com_mean$statistic
com_mean <- unlist(com_mean[c("estimate","se","statistic","p.value")])
names(com_mean) <- c("Estimate", "Std. Error", "t value", "Pr(>|t|")

   ## Combine both means.                                       ##
pop_means <- rbind(community=com_mean, "MET/CBT-5"=met_mean)

pop_means

  ################################################################
  ## Repeat with model that includes covariates--doubly robust  ##
  ## estimation.                                                ##
  ################################################################

att_com_met <- svyglm(suf12 ~ community + illact + crimjust +
                              subprob + subdep + white, design=dsd_com_met)

summary(att_com_met)

  ## The coefficient on community estimates the ATT for com-    ##
  ## munity care vs. MET/CBT-5 for youth like those in the com- ##
  ## munity sample.                                             ##

  ## Estimate the means.                                        ##
  ## The intercept is the counterfactual mean for youth like    ##
  ## those receiving community care had they received           ##
  ## MET/CBT-5.                                                 ##

met_mean <- summary(att_com_met)$coef[1,]

  ## Combine both means                                         ##
pop_means <- rbind(community=com_mean, "MET/CBT-5"=met_mean)

pop_means


  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## ATT community vs. SCY for youth like those receiving       ##
  ## community.                                                 ##
  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
```

```
  ## Create an svydesign that includes only community and       ##
  ## SCY.                                                       ##

  ## Center the covariates at the target group mean.            ##

dsd_com_scy <- subset(dsname, subset=(community==1 | scy == 1))
cmeans <- apply(dsname[dsname$community==1, covs], 2, mean)
dsd_com_scy[,covs] <- t(t(dsd_com_scy[,covs])-cmeans)

  ## Create the svydesign for ATT.                              ##

dsd_com_scy <- svydesign(id=~1, weights=~attwgt_com_scy,
                         data=dsd_com_scy)

  ## Estimate the ATT.                                          ##

att_com_scy <- svyglm(suf12 ~ community, design=dsd_com_scy)

summary(att_com_scy)

  ## The coefficient on community estimates the ATT for com-    ##
  ## munity care vs. SCY for youth like those in the community  ##
  ## sample.                                                    ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving community care had they received SCY. ##

scy_mean <- summary(att_com_scy)$coef[1,]

   ## Combine both means.                                       ##
pop_means <- rbind(community=com_mean, "SCY"=scy_mean)

pop_means

  ################################################################
  ## Repeat with model that includes covariates--doubly robust  ##
  ## estimation.                                                ##
  ################################################################

att_com_scy <- svyglm(suf12 ~ community + illact + crimjust + subprob +
                              subdep + white, design=dsd_com_scy)

summary(att_com_scy)

  ## The coefficient on community estimates the ATT for com-    ##
  ## munity care vs. SCY for youth like those in the community  ##
  ## sample.                                                    ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving community care had they received SCY. ##

scy_mean <- summary(att_com_scy)$coef[1,]

   ## Combine both means.                                       ##
pop_means <- rbind(community=com_mean, "MET/CBT-5"=scy_mean)
```

```
pop_means


  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## ATT MET/CBT-5 vs. community for youth like those receiving ##
  ## MET/CBT-5.                                                  ##
  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## Center the covariates at the target sample mean.            ##

dsd_met_com <- subset(dsname, subset=(community==1 | metcbt5 == 1))
cmeans <- apply(dsname[dsname$metcbt5==1, covs], 2, mean)
dsd_met_com[,covs] <- t(t(dsd_met_com[,covs])-cmeans)

  ## Create the svydesign for ATT.                               ##

dsd_met_com <- svydesign(id=~1, weights=~attwgt_met_com, data=dsd_met_com)

  ## Estimate the ATT.                                           ##

att_met_com <- svyglm(suf12 ~ metcbt5, design=dsd_met_com)

summary(att_met_com)

  ## The coefficient on metcbt5 estimates the ATT for MET/CBT-5 ##
  ## vs. community care for youth like those in the MET/CBT-5   ##
  ## sample.                                                    ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving MET/CBT-5 had they received community ##
  ## care.                                                      ##

com_mean <- summary(att_met_com)$coef[1,]

  ## The unweighted mean for the MET/CBT-5 group estimates the  ##
  ## mean for youth like those in the MET/CBT-5 group when they ##
  ## received MET/CBT-5.                                        ##

met_mean <- t.test(dsname$suf12[dsname$metcbt5==1])
met_mean <- met_mean[c("estimate","statistic","p.value")]
met_mean$se <- met_mean$estimate/met_mean$statistic
met_mean <- unlist(met_mean[c("estimate","se","statistic","p.value")])
names(met_mean) <- c("Estimate", "Std. Error", "t value", "Pr(>|t|")

   ## Combine both means.                                       ##
pop_means <- rbind(community=com_mean, "MET/CBT-5"=met_mean)

pop_means

  ################################################################
  ## Repeat with model that includes covariates--doubly robust  ##
  ## estimation.                                                ##
  ################################################################

att_met_com <- svyglm(suf12 ~ metcbt5 + illact + crimjust +
                              subprob + subdep + white, design=dsd_met_com)
```

```
summary(att_met_com)

  ## The coefficient on metcbt5 estimates the ATT for MET/CBT-5 ##
  ## vs. community care for youth like those in the MET/CBT-5   ##
  ## sample.                                                    ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving MET/CBT-5 had they received community ##
  ## care.                                                      ##

com_mean <- summary(att_met_com)$coef[1,]

   ## Combine both means                                        ##
pop_means <- rbind(community=com_mean, "MET/CBT-5"=met_mean)

pop_means


  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## ATT MET/CBT-5 vs. SCY for youth like those receiving       ##
  ## MET/CBT-5.                                                 ##
  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##


  ## Center the covariates at the target sample mean.           ##

dsd_met_scy <- subset(dsname, subset=(scy==1 | metcbt5 == 1))
cmeans <- apply(dsname[dsname$metcbt5==1, covs], 2, mean)
dsd_met_scy[,covs] <- t(t(dsd_met_scy[,covs])-cmeans)

  ## Create the svydesign for ATT.                              ##

dsd_met_scy <- svydesign(id=~1, weights=~attwgt_met_scy, data=dsd_met_scy)

  ## Estimate the ATT.                                          ##

att_met_scy <- svyglm(suf12 ~ metcbt5, design=dsd_met_scy)

summary(att_met_scy)

  ## The coefficient on metcbt5 estimates the ATT for MET/CBT-5 ##
  ## vs. SCY for youth like those in the MET/CBT-5 sample.      ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving MET/CBT-5 had they received SCY.      ##

scy_mean <- summary(att_met_scy)$coef[1,]

    ## Combine both means.                                       ##
pop_means <- rbind("MET/CBT-5"=met_mean, SCY=scy_mean)

pop_means

  ################################################################
```

```
  ## Repeat with model that includes covariates--doubly robust  ##
  ## estimation.                                                ##
  ###############################################################

att_met_scy <- svyglm(suf12 ~ metcbt5 + illact + crimjust +
                              subprob + subdep + white, design=dsd_met_scy)

summary(att_met_scy)

  ## The coefficient on metcbt5 estimates the ATT for MET/CBT-5 ##
  ## vs. SCY for youth like those in the MET/CBT-5 sample.      ##

  ## Estimate the means.                                        ##
  ## The intercept is the counterfactual mean for youth like    ##
  ## those receiving MET/CBT-5 had they recevied SCY.           ##

scy_mean <- summary(att_met_scy)$coef[1,]

    ## Combine both means.                                       ##
pop_means <- rbind("MET/CBT-5"=met_mean, SCY=scy_mean)

pop_means


  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## ATT SCY vs. community for youth like those receiving SCY.  ##
  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## Center the covariates at the target sample mean.           ##

dsd_scy_com <- subset(dsname, subset=(community==1 | scy == 1))
cmeans <- apply(dsname[dsname$scy==1, covs], 2, mean)
dsd_scy_com[,covs] <- t(t(dsd_scy_com[,covs])-cmeans)

  ## Create the svydesign for ATT.                              ##

dsd_scy_com <- svydesign(id=~1, weights=~attwgt_scy_com, data=dsd_scy_com)

  ## Estimate the ATT.                                          ##

att_scy_com <- svyglm(suf12 ~ scy, design=dsd_scy_com)

summary(att_scy_com)

  ## The coefficient on scy estimates the ATT for SCY vs. com-  ##
  ## munity care for youth like those in the SCY sample.        ##

  ## Estimate the means                                         ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving SCY had they received community care. ##

com_mean <- summary(att_scy_com)$coef[1,]

  ## The unweighted mean for the scy group estimates the mean   ##
  ## for youth like those in the SCY group when they received   ##
  ## SCY.                                                       ##

scy_mean <- t.test(dsname$suf12[dsname$scy==1])
```

```
scy_mean <- scy_mean[c("estimate","statistic","p.value")]
scy_mean$se <- scy_mean$estimate/scy_mean$statistic
scy_mean <- unlist(scy_mean[c("estimate","se","statistic","p.value")])
names(scy_mean) <- c("Estimate", "Std. Error", "t value", "Pr(>|t|")

   ## Combine both means.                                          ##
pop_means <- rbind("community"=com_mean, SCY=scy_mean)

pop_means

  ################################################################
  ## Repeat with model that includes covariates--doubly robust  ##
  ## estimation.                                                ##
  ################################################################

att_scy_com <- svyglm(suf12 ~ scy + illact + crimjust +
                              subprob + subdep + white, design=dsd_scy_com)

summary(att_scy_com)

  ## The coefficient on scy estimates the ATT for SCY vs. com-  ##
  ## munity care for youth like those in the SCY sample.        ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving SCY had they received community care. ##

com_mean <- summary(att_scy_com)$coef[1,]

   ## Combine both means.                                          ##
pop_means <- rbind("community"=com_mean, scy_mean)

pop_means


  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##
  ## ATT SCY vs. MET/CBT-5 for youth like those receiving SCY   ##
  ## %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ##

  ## Create an svydesign that includes only SCY and             ##
  ## MET/CBT-5                                                  ##

  ## Center the covariates at the target sample mean.           ##

dsd_scy_met <- subset(dsname, subset=(metcbt5==1 | scy == 1))
cmeans <- apply(dsname[dsname$scy==1, covs], 2, mean)
dsd_scy_met[,covs] <- t(t(dsd_scy_met[,covs])-cmeans)

  ## Create the svydesign for ATT.                              ##

dsd_scy_met <- svydesign(id=~1, weights=~attwgt_scy_met, data=dsd_scy_met)

  ## Estimate the ATT.                                          ##
att_scy_met <- svyglm(suf12 ~ scy, design=dsd_scy_met)

summary(att_scy_met)
```

```
  ## The coefficient on scy estimates the ATT for SCY vs.       ##
  ## MET/CBT-5 for youth like those in the SCY sample.          ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving SCY had they received MET/CBT-5.      ##

met_mean <- summary(att_scy_met)$coef[1,]

   ## Combine both means.                                       ##
pop_means <- rbind("MET/CBT-5"=met_mean, scy_mean)

pop_means

  ################################################################
  ## Repeat with model that includes covariates--doubly robust  ##
  ## estimation.                                                ##
  ################################################################

att_scy_met <- svyglm(suf12 ~ scy + illact + crimjust +
                              subprob + subdep + white, design=dsd_scy_met)

summary(att_scy_met)

  ## The coefficient on scy estimates the ATT for SCY vs.       ##
  ## MET/CBT-5 for youth like those in the SCY sample.          ##

  ## Estimate the means.                                        ##
  ## The intercept estimates the counterfactual mean for youth  ##
  ## like those receiving SCY had they received MET/CBT-5.      ##

met_mean <- summary(att_scy_met)$coef[1,]

   ## Combine both means.                                       ##
pop_means <- rbind("MET/CBT-5"=met_mean, scy_mean)

pop_means
```

2.3. Stata Code

```
log using example.stata.log, replace

insheet using dsname.csv

* +++++++++++++++++++++ * ATE * ++++++++++++++++++++++++ *

* ATE Estimates. Data described in SAS code

  ** [pw=atewgt] is syntax for using atewgt as a weight **
  ** By default Stata adjusts for weighting when pw are **
  ** specified.                                          **

regress suf12 metcbt5 scy [pw=atewgt]

  ** The coefficients on metcbt5 and scy estimate the   **
  ** pairwise ATE compared with community for MET/CBT-5 **
  ** and SCY respectively. Their difference estimates   **
```

```
  ** the ATE of SCY relative to MET/CBT-5.              **

  ** lincom is the Stata command for estimating linear  **
  ** combinations of estimated coefficients             **

lincom metcbt5 - scy

 ** The intercept _cons estimates the population mean   **
 ** for community. The population means for the other   **
 ** treatments are estimated by the sum of the         **
 ** intercept and the coefficients on the indicators.   **
lincom _cons + metcbt5
lincom _cons + scy

  ** Repeat ATE with covariate adjustment and weighting **

regress suf12 metcbt5 scy illact crimjust subprob subdep white  [pw=atewgt]

lincom metcbt5 - scy

  ** If we create a treatment variable with values for  **
  ** each group and let Stata create the indicators by  **
  ** using i. construction then we can use the margins  **
  ** post-estimation command to estimate the population **
  ** means. The noweights option causes Stata to ignore **
  ** the weights used in model fitting when averaging   **
  ** the predicted values.                              **

gen tx = 1 if community == 1
replace tx = 2 if metcbt5 == 1
replace tx = 3 if scy == 1

regress suf12 i.tx illact crimjust subprob subdep white  [pw=atewgt]

margins tx, noweights

* ++++++++++++++++++++ * ATT * ++++++++++++++++++++++++ *

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *
* ATT community vs MET/CBT-5 on population like those receiving community *
* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *

regress suf12 community [pw=attwgt_com_met] if tx != 3

  ** The coefficient on community estimates the ATT for **
  ** community care vs. MET/CBT-5 for youth like those  **
  ** in the community sample.                           **

  ** Estimate the means. _cons estimates the counter-   **
  ** factual mean for youth like those in the community **
  ** group had they received MET/CBT-5. The intercept   **
  ** plus the coefficient on the community indicator    **
  ** estimates the mean for the community group when    **
  ** they received community care. This also equals the **
  ** unweighted sample mean for the community group.    **

lincom _cons + community
```

```
  ** Repeat with covariate adjustment and weighting     **

  ** Even though community is a dichotomous variable we **
  ** use i.community in the model to allow for using    **
  ** the margins command to estimate the mean for all   **
  ** youth in the community group had they had          **
  ** community care and had they received MET/CBT-5.    **

  ** if tx != 3 excludes the SCY group.                 **

regress suf12 i.community illact crimjust subprob subdep \\\
        white [pw=attwgt_com_met] if tx != 3

  ** The margins command estimates the population       **
  ** means. The subpop(community) options specifies     **
  ** uses of the community sample for estimating the    **
  ** means.                                             **

margins community, subpop(community) noweights

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *
* ATT community vs SCY on population like those receiving community *
* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *

  ** if tx != 2 excludes the MET/CBT-5 group.           **

regress suf12 community [pw=attwgt_com_scy] if tx != 2

  ** The coefficient on community estimates the ATT for **
  ** community care vs. SCY for youth like those in the **
  ** community sample.                                  **

  ** Estimate the means. _cons estimates the counter-   **
  ** factual mean for youth like those in the community **
  ** group had they received SCY. The intercept plus    **
  ** the coefficient on the community indicator         **
  ** estimates the mean for the community group when    **
  ** they received community care. This also equals the **
  ** unweighted sample mean for the community group.    **

lincom _cons + community

  ** Repeat with covariate adjustment and weighting.    **

regress suf12 i.community illact crimjust subprob subdep \\\
        white [pw=attwgt_com_scy] if tx != 2

  ** Estimate the population means using margins        **

margins community, subpop(community) noweights

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *
* ATT MET/CBT-5 vs community on population like those receiving MET/CBT-5 *
* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *

  ** if tx != 3 excludes the SCY group.                 **
```

```
regress suf12 metcbt5 [pw=attwgt_met_com] if tx != 3

  ** The coefficient on metcbt5 estimates the ATT for   **
  ** MET/CBT-5 vs. community care for youth like those  **
  ** in the MET/CBT-5 sample.                           **

  ** Estimate the means. _cons estimates the counter-   **
  ** factual mean for youth like those in the MET/CBT-5 **
  ** group had they received community care. The        **
  ** intercept plus the coefficient on the MET/CBT-5    **
  ** indicator estimates the mean for the MET/CBT-5     **
  ** group when they received MET/CBT-5. This also      **
  ** equals the unweighted sample mean for the          **
  ** MET/CBT-5 group.                                   **


lincom _cons + metcbt5

** Repeat with covariate adjustment and weighting.    **

regress suf12 i.metcbt5 illact crimjust subprob subdep \\\
        white [pw=attwgt_met_com] if tx != 3

  ** Estimate the population means using margins        **

margins metcbt5, subpop(metcbt5) noweights

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *
* ATT MET/CBT-5 vs SCY on population like those receiving MET/CBT-5 *
* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *

  ** if tx != 1 excludes the community group.           **

regress suf12 metcbt5 [pw=attwgt_met_scy] if tx != 1

  ** The coefficient on metcbt5 estimates the ATT for   **
  ** MET/CBT-5 vs. SCY for youth like those in the      **
  ** MET/CBT-5 sample.                                  **

  ** Estimate the means. _cons estimates the counter-   **
  ** factual mean for youth like those in the MET/CBT-5 **
  ** group had they received SCY. The intercept plus    **
  ** the coefficient on the MET/CBT-5 indicator         **
  ** estimates the mean for the MET/CBT-5 group when    **
  ** they received MET/CBT-5. This also equals the      **
  ** unweighted sample mean for the MET/CBT-5 group.    **

lincom _cons + metcbt5

** Repeat with covariate adjustment and weighting.    **

regress suf12 i.metcbt5 illact crimjust subprob subdep \\\
        white [pw=attwgt_met_scy] if tx != 1

  ** Estimate the population means using margins        **
```

```
margins metcbt5, subpop(metcbt5) noweights

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *
* ATT SCY vs community on population like those receiving SCY *
* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *

  ** if tx != 2 excludes the MET/CBT-5 group.           **

regress suf12 scy [pw=attwgt_scy_com] if tx != 2

  ** The coefficient on scy estimates the ATT for SCY   **
  ** vs. community care for youth like those in the SCY **
  ** sample.                                            **

  ** Estimate the means. _cons estimates the counter-   **
  ** factual mean for youth like those in the SCY group **
  ** had they received community care. The intercept    **
  ** plus the coefficient on the SCY indicator          **
  ** estimates the mean for the SCY group when they     **
  ** received SCY. This also equals the unweighted      **
  ** sample mean for the SCY group.                     **

lincom _cons + scy

** Repeat with covariate adjustment and weighting.    **

regress suf12 i.scy illact crimjust subprob subdep \\\
        white [pw=attwgt_scy_com] if tx != 2

  ** Estimate the population means using margins        **

margins scy, subpop(scy) noweights

* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *
* ATT SCY vs MET/CBT-5 on population like those receiving SCY *
* %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% *

  ** if tx != 1 excludes the commuity group.           **

regress suf12 scy [pw=attwgt_scy_met] if tx != 1

  ** The coefficient on scy estimates the ATT for SCY   **
  ** vs. MET/CBT-5 for youth like those in the SCY      **
  ** sample.                                            **

  ** Estimate the means. _cons estimates the counter-   **
  ** factual mean for youth like those in the SCY group **
  ** had they received MET/CBT-5. The intercept plus    **
  ** the coefficient on the SCY indicator estimates     **
  ** estimates the mean for the SCY group when they     **
  ** received SCY. This also equals the unweighted      **
  ** sample mean for the SCY group.                     **

lincom _cons + scy

** Repeat with covariate adjustment and weighting.    **
```

```
regress suf12 i.scy illact crimjust subprob subdep \\\
        white [pw=attwgt_scy_met] if tx != 1

  ** Estimate the population means using margins        **

margins scy, subpop(scy) noweights
```