

MCBS_HMO_weights_readme.txt

Purpose: Create HMO adjustment weights for each of the 5 multiply imputed datasets by population and multiple. Within imputed survey data, including participants with both 'pure' and 'non-pure' Medicare enrollment, is used for creation of the adjustment weights.

Community population--Program/code required:
HMO_Weights_MCBS2009.sas

Institutionalized population--Program/code required:
HMO_Weights_MCBS2009_INST.sas

```

*****
*****
Program: HMO_Weights_MCBS2009_INST.sas
Purpose:
1. Calculate weights to adjust for MCBS where not considered "Pure" Medicare
enrollment (MCBS "pure"
selected as follows: 1. no participation in Medicare Advantage program, and
2. enrolled in Medicare parts A & B
for the full 12-months unless participant died during the year (if
mcadv_months=0 and if mcab_months_3grp=2 or died=1).
2. Incorporate adjustment to year 2000 and 2010 dollars:
    a. 2000: If year=2009 then adj=88.723/109.729
    b. 2010: If year=2009 then adj=110.992/109.729
3. Incorporate NHEA cost adjustment.
*****
*****;
%let l=1;
*%let l=2;
*%let l=3;
*%let l=4;
*%let l=5;

%let yr=09;
%let year=2009;

libname mc&yr. "Insert file path";
libname mcwgt "Insert file path";
libname mcc "Insert file path";
libname nhea "Insert file path";

proc contents data=mcwgt.inst_mcbs_i&yr.;
run;

proc contents data=mcc.Calibr_inst_mcbs&yr.;
run;

*1 multiple dataset that needs to be split into 5;
*Within imputed*;
Data mcwgt.inst_mcbs_i&yr._cl1 mcwgt.inst_mcbs_i&yr._cl2
mcwgt.inst_mcbs_i&yr._cl3 mcwgt.inst_mcbs_i&yr._cl4
mcwgt.inst_mcbs_i&yr._cl5;
Set mcwgt.inst_mcbs_i&yr.;
If _mult_=1 then output mcwgt.inst_mcbs_i&yr._cl1;
If _mult_=2 then output mcwgt.inst_mcbs_i&yr._cl2;
If _mult_=3 then output mcwgt.inst_mcbs_i&yr._cl3;
If _mult_=4 then output mcwgt.inst_mcbs_i&yr._cl4;

```

```
If _mult_=5 then output mcwgt.inst_mcbs_i&yr._cl5;  
Run;
```

```
*Full calibrated sets*;
```

```
Data mcwgt.calibr_inst_mcbs&yr._cl1 mcwgt.calibr_inst_mcbs&yr._cl2  
mcwgt.calibr_inst_mcbs&yr._cl3 mcwgt.calibr_inst_mcbs&yr._cl4  
mcwgt.calibr_inst_mcbs&yr._cl5;  
Set mcc.calibr_inst_mcbs&yr.;  
If _mult_=1 then output mcwgt.calibr_inst_mcbs&yr._cl1;  
If _mult_=2 then output mcwgt.calibr_inst_mcbs&yr._cl2;  
If _mult_=3 then output mcwgt.calibr_inst_mcbs&yr._cl3;  
If _mult_=4 then output mcwgt.calibr_inst_mcbs&yr._cl4;  
If _mult_=5 then output mcwgt.calibr_inst_mcbs&yr._cl5;  
Run;
```

```
proc sort data=mcwgt.inst_mcbs_i&yr._cl&l. out=inst_mcbs_i&yr._cl&l.; by  
baseid; run;
```

```
*Recode categorical variables to binary;
```

```
data mc_all_rv1_cl&l.;  
    set inst_mcbs_i&yr._cl&l.;
```

```
*Recode categorical variables for inclusion in modeling*;
```

```
if race=4 then race4=1;  
else race4=0;  
if race=3 then race3=1;  
else race3=0;  
if race=2 then race2=1;  
else race2=0;  
if ed5=5 then ed_5=1;  
else ed_5=0;  
if ed5=4 then ed_4=1;  
else ed_4=0;  
if ed5=3 then ed_3=1;  
else ed_3=0;  
if ed5=2 then ed_2=1;  
else ed_2=0;  
if comphealth=3 then comphealth3=1;  
else comphealth3=0;  
if comphealth=2 then comphealth2=1;  
else comphealth2=0;  
if healthstat=5 then healthstat5=1;  
else healthstat5=0;  
if healthstat=4 then healthstat4=1;  
else healthstat4=0;  
if healthstat=3 then healthstat3=1;
```

```

else healthstat3=0;
if healthstat=2 then healthstat2=1;
else healthstat2=0;
if dif_stoop=4 then dif_stoop4=1;
else dif_stoop4=0;
if dif_stoop=3 then dif_stoop3=1;
else dif_stoop3=0;
if dif_stoop=2 then dif_stoop2=1;
else dif_stoop2=0;
if dif_lift=4 then dif_lift4=1;
else dif_lift4=0;
if dif_lift=3 then dif_lift3=1;
else dif_lift3=0;
if dif_lift=2 then dif_lift2=1;
else dif_lift2=0;
if povcat=5 then povcat5=1;
else povcat5=0;
if povcat=4 then povcat4=1;
else povcat4=0;
if povcat=3 then povcat3=1;
else povcat3=0;
if povcat=2 then povcat2=1;
else povcat2=0;
if maritals=4 then maritals4=1;
else maritals4=0;
if maritals=5 then maritals5=1;
else maritals5=0;
if maritals=2 then maritals2=1;
else maritals2=0;
run;

*Confirm "pure";
proc freq data=mc_all_rv1_cl&l.;
tables pure;
title "Pure Medicare";
run;

*Run logistic regression of pure on covariates, output predicted
probabilities;
*2009: multiple models to fit calibrated sets.
*For calibrated set 1, 2, 3 ONLY: For better model fit: dif_stoop, dif_lift,
healthstat,
comphealth, as continuous, and ed5, povcat as categorical;
ods trace on;
proc logistic data=mc_all_rv1_cl&l.;
class race4 race3 race2 flushot pneushot didserv dif_walk

```

```

ed_5 ed_4 ed_3 ed_2  ever smoke hearingaid hyst male mammogram
maritals2 marital4 marital5 pap_smear
povcat5 povcat4 povcat3 povcat2 prb_dres prb_eat psa1yr smokenow died
;
model pure (desc)=age height weightkg inpatnights inpatstays
race4 race3 race2 flushot pneushot didserv dif_stoop
dif_lift dif_walk ed_5 ed_4 ed_3 ed_2
ever smoke comphealth healthstat
hear_inst hearingaid hyst male mammogram marital2 marital4 marital5
pap_smear
povcat5 povcat4 povcat3 povcat2 prb_dres prb_eat psa1yr smokenow died
/ rsq lackfit;
output out=predstat_cl&l. prob=p xbeta=xbeta;
title 'Model Pure Medicare on Covariates';
run;
ods trace off;

*For calibrated set 4, 5 ONLY: For better model fit: dif_stoop, dif_lift,
comphealth, povcat as continuous, and ed5, healthstat as categorical;
ods trace on;
proc logistic data=mc_all_rv1_cl&l.;
class race4 race3 race2 flushot pneushot didserv
dif_walk ed_5 ed_4 ed_3 ed_2  ever smoke
healthstat2 healthstat3 healthstat4 healthstat5
hearingaid hyst male mammogram marital2 marital4 marital5 pap_smear
prb_dres prb_eat psa1yr smokenow died
;
model pure (desc)=age height weightkg inpatnights inpatstays
race4 race3 race2 flushot pneushot didserv dif_stoop
dif_lift dif_walk ed_5 ed_4 ed_3 ed_2
ever smoke comphealth healthstat2 healthstat3 healthstat4 healthstat5
hear_inst hearingaid hyst male mammogram marital2 marital4 marital5
pap_smear
povcat prb_dres prb_eat psa1yr smokenow died
/ rsq lackfit;
output out=predstat_cl&l. prob=p xbeta=xbeta;
title 'Model Pure Medicare on Covariates';
run;
ods trace off;

*Create 'pure' weight*;
data test_cl&l.;
    set predstat_cl&l.;
pred=1/p;
pure_wgt=wgt*pred;
run;

```

```

proc print data=test_cl&l. (obs=10);
var wgt p pred pure_wgt;
run;

proc univariate data=test_cl&l. normal plots;
var p;
class pure;
histogram p/normal;
title 'Predicted Probability';
run;

proc univariate data=test_cl&l. normal plots;
var pred;
where pure=1;
histogram pred/normal;
title '1/Predicted Probability, Pure=1';
run;

proc print data=test_cl&l.;
var BASEID pred cost;
where pred gt 2 and pure=1;
run;

proc univariate data=test_cl&l. normal plots;
var cost;
where pure=1;
title "Cost prior to adjustments, where pure=1";
run;

*Investigate covariates that may be driving force of weights*;
*Check for balance;
%macro balance (var,varname);
proc reg data=test_cl&l.;
model &var.=p;
output out=&var._cl&l. r=resid;
run;
quit;

proc sort data=&var._cl&l.;
by pure;
run;

proc means data=&var._cl&l.;
class pure;
var resid;

```

```
output out=m_&var._cl&l. mean= std= /autoname;
run;
```

```
Data p_&var._cl&l. (keep=covar pure_m pure_sd);
  Set m_&var._cl&l.;
  If pure=1;
length covar $14;
covar=&varname.;
  rename resid_Mean=pure_m resid_StdDev=pure_sd;
  Run;
```

```
Data np_&var._cl&l. (keep=npure_m npure_sd);
  Set m_&var._cl&l.;
  If pure=0;
  rename resid_Mean=npure_m resid_StdDev=npure_sd;
  Run;
```

```
data m_&var._all_cl&l.;
  merge p_&var._cl&l. np_&var._cl&l.;
  D=((pure_m-npure_m)/(sqrt((pure_sd**2+npure_sd**2)/2)));
run;
```

```
data m_&var._all_cl&l.;
  retain covar pure_m pure_sd npure_m npure_sd D;
  set m_&var._all_cl&l.;
run;
```

```
%mend;
```

```
*For calibrated set 1,2,3 ONLY: For better model fit: dif_stoop, dif_lift,
healthstat,
```

```
comphealth, as continuous, and ed5, povcat as categorical;
```

```
%balance (age, 'age')
```

```
%balance (height, 'height')
```

```
%balance (weightkg, 'weightkg')
```

```
%balance (inpatnights, 'inpatnights')
```

```
%balance (inpatstays, 'inpatstays')
```

```
%balance (race4, 'race4')
```

```
%balance (race3, 'race3')
```

```
%balance (race2, 'race2')
```

```
%balance (flushot, 'flushot')
```

```
%balance (pneushot, 'pneushot')
```

```
%balance (didserv, 'didserv')
```

```
%balance (dif_stoop, 'dif_stoop')
```

```
%balance (dif_lift, 'dif_lift')
```

```
%balance (dif_walk, 'dif_walk')
```

```
%balance (ed_5, 'ed_5')
%balance (ed_4, 'ed_4')
%balance (ed_3, 'ed_3')
%balance (ed_2, 'ed_2')
%balance (eversmoke, 'eversmoke')
%balance (comphealth, 'comphealth')
%balance (healthstat, 'healthstat')
%balance (hear_inst, 'hear_inst')
%balance (hearingaid, 'hearingaid')
%balance (hyst, 'hyst')
%balance (male, 'male')
%balance (mammogram, 'mammogram')
%balance (maritals2, 'maritals2')
%balance (maritals4, 'maritals4')
%balance (maritals5, 'maritals5')
%balance (pap_smear, 'pap_smear')
%balance (povcat5, 'povcat5')
%balance (povcat4, 'povcat4')
%balance (povcat3, 'povcat3')
%balance (povcat2, 'povcat2')
%balance (prb_dres, 'prb_dres')
%balance (prb_eat, 'prb_eat')
%balance (psa1yr, 'psa1yr')
%balance (smokenow, 'smokenow')
%balance (died, 'died')
;
```

```
*Combine for export*;
data all_cohend_cl&l.;
    set m_age_all_cl&l.
m_height_all_cl&l.
m_weightkg_all_cl&l.
m_inpatnights_all_cl&l.
m_inpatstays_all_cl&l.
m_race4_all_cl&l.
m_race3_all_cl&l.
m_race2_all_cl&l.
m_flushot_all_cl&l.
m_pneushot_all_cl&l.
m_didserv_all_cl&l.
m_dif_stoop_all_cl&l.
m_dif_lift_all_cl&l.
m_dif_walk_all_cl&l.
m_ed_5_all_cl&l.
m_ed_4_all_cl&l.
m_ed_3_all_cl&l.
```



```
m_ed_2_all_cl&l.  
m_everSmoke_all_cl&l.  
m_compHealth_all_cl&l.  
m_healthStat_all_cl&l.  
m_hear_inst_all_cl&l.  
m_hearingaid_all_cl&l.  
m_hyst_all_cl&l.  
m_male_all_cl&l.  
m_mammogram_all_cl&l.  
m_maritals2_all_cl&l.  
m_maritals4_all_cl&l.  
m_maritals5_all_cl&l.  
m_pap_smear_all_cl&l.  
m_povcat5_all_cl&l.  
m_povcat4_all_cl&l.  
m_povcat3_all_cl&l.  
m_povcat2_all_cl&l.  
m_prb_dres_all_cl&l.  
m_prb_eat_all_cl&l.  
m_psa1yr_all_cl&l.  
m_smokenow_all_cl&l.  
m_died_all_cl&l.
```

```
;  
run;
```

```
proc export data= all_cohend_cl&l.  
    outfile="Inserrt file path\balance_final_cl&l..xls"  
    dbms=excel replace;  
    sheet="&year._cl&l.";  
run;
```

*For calibrated set 4, 5 ONLY: For better model fit: dif_stoop, dif_lift, compHealth, povcat as continuous, and ed5, healthStat as categorical;

```
%balance (age, 'age')  
%balance (height, 'height')  
%balance (weightkg, 'weightkg')  
%balance (inpatnights, 'inpatnights')  
%balance (inpatstays, 'inpatstays')  
%balance (race4, 'race4')  
%balance (race3, 'race3')  
%balance (race2, 'race2')  
%balance (flushot, 'flushot')  
%balance (pneushot, 'pneushot')  
%balance (didserv, 'didserv')  
%balance (dif_stoop, 'dif_stoop')  
%balance (dif_lift, 'dif_lift')
```

```
%balance (dif_walk, 'dif_walk')
%balance (ed_5, 'ed_5')
%balance (ed_4, 'ed_4')
%balance (ed_3, 'ed_3')
%balance (ed_2, 'ed_2')
%balance (eversmoke, 'eversmoke')
%balance (comphealth, 'comphealth')
%balance (healthstat2, 'healthstat2')
%balance (healthstat3, 'healthstat3')
%balance (healthstat4, 'healthstat4')
%balance (healthstat5, 'healthstat5')
%balance (hear_inst, 'hear_inst')
%balance (hearingaid, 'hearingaid')
%balance (hyst, 'hyst')
%balance (male, 'male')
%balance (mammogram, 'mammogram')
%balance (maritals2, 'maritals2')
%balance (maritals4, 'maritals4')
%balance (maritals5, 'maritals5')
%balance (pap_smear, 'pap_smear')
%balance (povcat, 'povcat')
%balance (prb_dres, 'prb_dres')
%balance (prb_eat, 'prb_eat')
%balance (psa1yr, 'psa1yr')
%balance (smokenow, 'smokenow')
%balance (died, 'died')
;
```

```
*Combine for export*;
data all_cohend_cl&l.;
    set m_age_all_cl&l.
m_height_all_cl&l.
m_weightkg_all_cl&l.
m_inpatnights_all_cl&l.
m_inpatstays_all_cl&l.
m_race4_all_cl&l.
m_race3_all_cl&l.
m_race2_all_cl&l.
m_flushot_all_cl&l.
m_pneushot_all_cl&l.
m_didserv_all_cl&l.
m_dif_stoop_all_cl&l.
m_dif_lift_all_cl&l.
m_dif_walk_all_cl&l.
m_ed_5_all_cl&l.
m_ed_4_all_cl&l.
```

```

m_ed_3_all_cl&l.
m_ed_2_all_cl&l.
m_everSmoke_all_cl&l.
m_compHealth_all_cl&l.
m_healthstat2_all_cl&l.
m_healthstat3_all_cl&l.
m_healthstat4_all_cl&l.
m_healthstat5_all_cl&l.
m_hear_inst_all_cl&l.
m_hearingaid_all_cl&l.
m_hyst_all_cl&l.
m_male_all_cl&l.
m_mammogram_all_cl&l.
m_maritals2_all_cl&l.
m_maritals4_all_cl&l.
m_maritals5_all_cl&l.
m_pap_smear_all_cl&l.
m_povcat_all_cl&l.
m_prb_dres_all_cl&l.
m_prb_eat_all_cl&l.
m_psa1yr_all_cl&l.
m_smokenow_all_cl&l.
m_died_all_cl&l.
;
run;

proc export data= all_cohend_cl&l.
    outfile="Insert file path\balance_final_cl&l..xls"
    dbms=excel replace;
    sheet="&year._cl&l.";
run;

*****Create final dataset and calculate mean cost with new
weights*****;
*Subset for only pure Medicare*;
data pure1_cl&l. (keep=baseid p pred pure_wgt);
    set test_cl&l.;
    if pure=1;
run;

*NHEA cost adjustment factor dataset*;
data nhea_&year. (keep=baseid NHEA_Adjustment);
set nhea.nhea_adjustment&year.;
run;

```

```

***Merge new survey weights and NHEA cost adjustment factors to pure Medicare
dataset***;
proc sort data=mcwgt.calibr_inst_mcbs&yr._cl&l.
out=calibr_inst_mcbs&yr._cl&l.; by baseid; run;
proc sort data=pure1_cl&l.; by baseid; run;
proc sort data=nhea_&year.; by baseid; run;

data mcc.mc_inst_pure&yr._cl&l.
(rename=(NHEA_Adjustment=nhea_adjustment_factor wgt=survey_wgt
pred=puremc_wgt pure_wgt=final_wgt)); *Pure sample;
merge calibr_inst_mcbs&yr._cl&l. (in=a) pure1_cl&l. nhea_&year.;
cost_2000d=(cost*88.723)/109.729; *remember to change factor for each
year update;
cost_2000d_nhea=(cost_2000d*NHEA_Adjustment);
cost_2010d=(cost*110.992)/109.729; *remember to change factor for each
year update;
cost_2010d_nhea=(cost_2010d*NHEA_Adjustment);
by baseid;
label NHEA_Adjustment="NHEA cost adjustment factor"
wgt="Survey Weight (MCBS)"
pred="Pure Medicare adjustment weight"
pure_wgt="Product of survey weight and pure adjustment weight"
cost_2000d="Expenditure adjusted to year 2000 dollars"
cost_2000d_nhea="Expenditure adjusted to year 2000 dollars and NHEA
cost adjustment factor"
cost_2010d="Expenditure adjusted to year 2010 dollars"
cost_2010d_nhea="Expenditure adjusted to year 2010 dollars and NHEA
cost adjustment factor"
/*Add any labels missing from dataset*/;
if a;
run;

ods rtf file="Insert file path\MC_Inst_Pure&yr._Contents.rtf";
proc contents data=mcc.mc_inst_pure&yr._cl&l.;
title "Contents: MC_Inst_Pure&yr.";
run;
ods rtf close;

proc surveymeans data=mcc.mc_inst_pure&yr._cl&l.;
stratum psu;
cluster strat;
weight final_wgt;
var cost_2010d_nhea;
title "With new survey weight, CL&l. (adjusted to yr 2010 $$ with NHEA
adjustment factor)";
run;

```

```
proc surveymeans data=mcc.mc_inst_pure&yr._cl&l.;  
stratum psu;  
cluster strat;  
weight survey_wgt;  
var cost_2010d_nhea;  
title "With old survey weight, CL&l. (adjusted to yr 2010 $$ with NHEA  
adjustment factor)";  
run;
```

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

Program: HMO_Weights_MCBS2009.sas

Purpose:

1. Calculate weights to adjust for MCBS where not considered "Pure" Medicare enrollment (MCBS "pure"

selected as follows: 1. no participation in Medicare Advantage program, and

2. enrolled in Medicare parts A & B

for the full 12-months unless participant died during the year (if

mcadv_months=0 and if mcab_months_3grp=2 or died=1).

2. Incorporate adjustment to year 2000 dollars:

 If year=2009 then adj=88.723/109.729

 Incorporate adjustment to year 2010 dollars:

 If year=2009 then adj=110.992/109.729

3. Incorporate NHEA cost adjustment.

```
*****
*****;
```

```
%let l=1;
```

```
    *%let l=2;
```

```
    *%let l=3;
```

```
    *%let l=4;
```

```
    *%let l=5;
```

```
%let yr=09;
```

```
%let year=2009;
```

```
libname mc&yr. "Insert file path";
```

```
libname mcwgt "Insert file path";
```

```
libname wgts "Insert file path";
```

```
libname mcp "Insert file path";
```

```
libname nhea "Insert file path";
```

```
proc contents data=wgts.mcbs_ni_i&yr.;
```

```
run;
```

*First create dataset with indicator for "pure" Medicare enrollment, using within imputed datasets for total sample;

*1 multiple dataset that needs to be split into 5;

```
Data mcwgt.mcbs_ni_iwt&yr._cl1 mcwgt.mcbs_ni_iwt&yr._cl2
```

```
mcwgt.mcbs_ni_iwt&yr._cl3 mcwgt.mcbs_ni_iwt&yr._cl4
```

```
mcwgt.mcbs_ni_iwt&yr._cl5;
```

```
Set wgts.mcbs_ni_i&yr.;
```

```
If _mult_=1 then output mcwgt.mcbs_ni_iwt&yr._cl1;
```

```
If _mult_=2 then output mcwgt.mcbs_ni_iwt&yr._cl2;
```

```
If _mult_=3 then output mcwgt.mcbs_ni_iwt&yr._cl3;
```

```
If _mult_=4 then output mcwgt.mcbs_ni_iwt&yr._cl4;
```

```
If _mult_=5 then output mcwgt.mcbs_ni_iwt&yr._cl5;  
Run;
```

```
proc sort data=mcwgt.mcbs_ni_iwt&yr._cl&l. out=mcbs_ni_iwt&yr._cl&l.; by  
baseid; run;
```

```
*Recode categorical variables to binary;
```

```
data mc_all_rv1_cl&l.;  
    set mcbs_ni_iwt&yr._cl&l.;
```

```
*Recode categorical variables for inclusion in modeling*;
```

```
if race=4 then race4=1;  
else race4=0;  
if race=3 then race3=1;  
else race3=0;  
if race=2 then race2=1;  
else race2=0;  
if ed5=5 then ed_5=1;  
else ed_5=0;  
if ed5=4 then ed_4=1;  
else ed_4=0;  
if ed5=3 then ed_3=1;  
else ed_3=0;  
if ed5=2 then ed_2=1;  
else ed_2=0;  
if comphealth=3 then comphealth3=1;  
else comphealth3=0;  
if comphealth=2 then comphealth2=1;  
else comphealth2=0;  
if healthstat=5 then healthstat5=1;  
else healthstat5=0;  
if healthstat=4 then healthstat4=1;  
else healthstat4=0;  
if healthstat=3 then healthstat3=1;  
else healthstat3=0;  
if healthstat=2 then healthstat2=1;  
else healthstat2=0;  
if dif_stoop=4 then dif_stoop4=1;  
else dif_stoop4=0;  
if dif_stoop=3 then dif_stoop3=1;  
else dif_stoop3=0;  
if dif_stoop=2 then dif_stoop2=1;  
else dif_stoop2=0;  
if dif_lift=4 then dif_lift4=1;  
else dif_lift4=0;  
if dif_lift=3 then dif_lift3=1;
```

```
else dif_lift3=0;
if dif_lift=2 then dif_lift2=1;
else dif_lift2=0;
if bc_taken=5 then bc_taken5=1;
else bc_taken5=0;
if bc_taken=4 then bc_taken4=1;
else bc_taken4=0;
if bc_taken=3 then bc_taken3=1;
else bc_taken3=0;
if bc_taken=2 then bc_taken2=1;
else bc_taken2=0;
if bp_taken=4 then bp_taken4=1;
else bp_taken4=0;
if bp_taken=3 then bp_taken3=1;
else bp_taken3=0;
if bp_taken=2 then bp_taken2=1;
else bp_taken2=0;
if dwel=4 then dwel4=1;
else dwel4=0;
if dwel=3 then dwel3=1;
else dwel3=0;
if dwel=2 then dwel2=1;
else dwel2=0;
if povcat=5 then povcat5=1;
else povcat5=0;
if povcat=4 then povcat4=1;
else povcat4=0;
if povcat=3 then povcat3=1;
else povcat3=0;
if povcat=2 then povcat2=1;
else povcat2=0;
if maritals=4 then maritals4=1;
else maritals4=0;
if maritals=5 then maritals5=1;
else maritals5=0;
if maritals=2 then maritals2=1;
else maritals2=0;
i_dayssq=i_days**2;
pov_care=povcat*havecare;
care_job=havecare*hasjob;
race3_job=race3*hasjob;
race2_pov=race2*povcat;
mar2_race3=maritals2*race3;
mar4_race2=maritals4*race2;
mar5_race3=maritals5*race3;
race3_stays=race3*inpatstays;
```



```

agesq=age**2;
race2_serve=didserv*race2;
patstaysq=inpatstays**2;
run;

*Confirm "pure";
proc freq data=mc_all_rv1_cl&l.;
tables pure;
title 'Pure';
run;

*Run logistic regression of pure on covariates, output predicted
probabilities;
ods trace on;
proc logistic data=mc_all_rv1_cl&l.;
class race4 race3 race2 flushot pneushot asthma_emphysema
bc_taken5 bc_taken4 bc_taken3 bc_taken2 bp_taken4 bp_taken3 bp_taken2
comphealth3 comphealth2 didserv dif_stoop4 dif_stoop3 dif_stoop2
dif_lift4 dif_lift3 dif_lift2 dif_walk ed_5 ed_4 ed_3 ed_2
eversmoke hasjob havecare healthstat2 healthstat3 healthstat4 healthstat5
hearing hearingaid hyst male mammogram marital2 marital4 marital5
pap_smear
povcat5 povcat4 povcat3 povcat2 prb_dres prb_eat psa1yr smokenow died
race3_job race2_serve mar2_race3 mar5_race3 mar4_race2 care_job ;
model pure (desc)=age height weightkg inpatnights inpatstays nbrpeopl i_days
race4 race3 race2 flushot pneushot asthma_emphysema
bc_taken5 bc_taken4 bc_taken3 bc_taken2 bp_taken4 bp_taken3 bp_taken2
comphealth3 comphealth2 didserv dif_stoop4 dif_stoop3 dif_stoop2
dif_lift4 dif_lift3 dif_lift2 dif_walk ed_5 ed_4 ed_3 ed_2
eversmoke hasjob havecare healthstat2 healthstat3 healthstat4 healthstat5
hearing hearingaid hyst male mammogram marital2 marital4 marital5
pap_smear
/*povcat*/ povcat5 povcat4 povcat3 povcat2 prb_dres prb_eat psa1yr smokenow
died
race3_job race2_serve mar2_race3 mar5_race3 mar4_race2 race2_pov agesq
race3_stays i_dayssq patstaysq care_job pov_care
/ rsq lackfit;
output out=predstat_cl&l. prob=p xbeta=xbeta;
title 'Model Pure Medicare on Covariates';
run;
ods trace off;

*Create 'pure' weight*;
data test_cl&l.;
set predstat_cl&l.;
pred=1/p;

```

```
pure_wgt=wgt*pred;  
run;
```

```
proc print data=test_cl&l. (obs=10);  
var wgt p pred pure_wgt;  
run;
```

```
proc univariate data=test_cl&l. normal plots;  
var p;  
class pure;  
histogram p/normal;  
title 'Predicted Probability';  
run;
```

```
proc univariate data=test_cl&l. normal plots;  
var pred;  
where pure=1;  
histogram pred/normal;  
title '1/Predicted Probability, Pure=1';  
run;
```

```
proc print data=test_cl&l.;  
var BASEID pred cost;  
where pred gt 3 and pure=1;  
run;
```

Investigate covariates that may be driving force of weights;

```
proc rank data=test_cl&l. out=deciles_cl&l. groups=10;  
var xbeta;  
ranks r_pred;  
Run;
```

```
data deciles_cl&l.;  
    set deciles_cl&l.;  
r_pred=r_pred+1;  
run;
```

```
proc freq data=deciles_cl&l.;  
tables pure*r_pred;  
run;
```

```
proc sort data=deciles_cl&l.; by pure r_pred; run;  
proc means data=deciles_cl&l.;  
var age height weightkg inpatnights inpatstays nbrpeopl i_days  
race4 race3 race2 flushot pneushot asthma_emphysema  
bc_taken5 bc_taken4 bc_taken3 bc_taken2 bp_taken4 bp_taken3 bp_taken2
```

```

comphealth3 comphealth2 didserv dif_stoop4 dif_stoop3 dif_stoop2
dif_lift4 dif_lift3 dif_lift2 dif_walk ed_5 ed_4 ed_3 ed_2
eversmoke hasjob havecare healthstat2 healthstat3 healthstat4 healthstat5
hearing hearingaid hyst male mammogram maritals2 maritals4 maritals5
pap_smear
povcat5 povcat4 povcat3 povcat2 prb_dres prb_eat psa1yr smokenow died
race3_job race2_serve mar2_race3 mar5_race3 mar4_race2 race2_pov agesq
race3_stays i_dayssq patstaysq care_job pov_care ;
class pure r_pred;
output out=by_group;
run;

proc export data=by_group
      outfile="Insert file path\mean_by_group_cl&l..csv"
      dbms=csv replace;
run;

*Run models to obtain F-statistic;
proc glm data=deciles_cl&l. outstat=fstat_cl&l.;
class pure r_pred;
model age height weightkg inpatnights inpatstays nbrpeopl i_days
race4 race3 race2 flushot pneushot asthma_emphysema
bc_taken5 bc_taken4 bc_taken3 bc_taken2 bp_taken4 bp_taken3 bp_taken2
comphealth3 comphealth2 didserv dif_stoop4 dif_stoop3 dif_stoop2
dif_lift4 dif_lift3 dif_lift2 dif_walk ed_5 ed_4 ed_3 ed_2
eversmoke hasjob havecare healthstat2 healthstat3 healthstat4 healthstat5
hearing hearingaid hyst male mammogram maritals2 maritals4 maritals5
pap_smear
/*povcat*/ povcat5 povcat4 povcat3 povcat2 prb_dres prb_eat psa1yr smokenow
died
race3_job race2_serve mar2_race3 mar5_race3 mar4_race2 race2_pov agesq
race3_stays i_dayssq patstaysq care_job pov_care
=pure r_pred pure*r_pred;
lsmeans pure;
run;
quit;

Data F_st_pure_cl&l. (keep=varname F_pure);
  Set fstat_cl&l.;
  If _type_='SS3' ;
  If _source_='pure';
  rename F=F_pure prob=Prob_pure _name_=varname;
  Run;

Data F_st_r_pred_cl&l. (keep=varname F_pclass);
  Set fstat_cl&l.;

```

```

If _type_='SS3' ;
If _source_='r_pred';
rename F=F_pclass prob=Prob_pclass _name_=varname;
Run;

Data F_st_int_cl&l. (keep=varname F_inter);
Set fstat_cl&l.;
If _type_='SS3' ;
If _source_='pure*r_pred';
rename F=F_inter prob=Prob_inter _name_=varname;
Run;

proc sort data=F_st_pure_cl&l.; by varname; run;
proc sort data=F_st_r_pred_cl&l.; by varname; run;
proc sort data=F_st_int_cl&l.; by varname; run;

data f_stat_all_cl&l.;
merge F_st_pure_cl&l. F_st_r_pred_cl&l. F_st_int_cl&l.;
by varname;
run;

PROC EXPORT DATA= WORK.F_stat_all_cl&l.
OUTFILE= "Insert file path\F_stats_cl&l._FINAL.xls"
DBMS=EXCEL REPLACE;
SHEET="cl&l.";
RUN;

*****Create final dataset and calculate mean cost with new
weights*****;
*Subset for only pure Medicare*;
data pure1_cl&l. (keep=baseid p pred pure_wgt);
set test_cl&l.;
if pure=1;
run;

*NHEA cost adjustment factor dataset from Kaushik*;
data nhea_&year. (keep=baseid NHEA_Adjustment);
set nhea.nhea_adjustment&year.;
run;

***Merge new survey weights and NHEA cost adjustment factors to pure Medicare
dataset***;
proc sort data=mcwgt.mcbs&yr._cl&l. out=mcbs&yr._cl&l.; by baseid; run;
proc sort data=pure1_cl&l.; by baseid; run;
proc sort data=nhea_&year.; by baseid; run;

```

```

data mcp.mc_pure&yr._cl&l. (rename=(NHEA_Adjustment=nhea_adjustment_factor
wgt=survey_wgt pred=puremc_wgt pure_wgt=final_wgt)); *Pure sample;
    merge mcbs&yr._cl&l. (in=a) pure1_cl&l. nhea_&year.;
    cost_2000d=(cost*88.723)/109.729; *remember to change factor for each
year update;
    cost_2000d_nhea=(cost_2000d*NHEA_Adjustment);
    cost_2010d=(cost*110.992)/109.729; *remember to change factor for each
year update;
    cost_2010d_nhea=(cost_2010d*NHEA_Adjustment);
by baseid;
label NHEA_Adjustment="NHEA cost adjustment factor"
    wgt="Survey Weight (MCBS)"
    pred="Pure Medicare adjustment weight"
    pure_wgt="Product of survey weight and pure adjustment weight"
    cost_2000d="Expenditure adjusted to year 2000 dollars"
    cost_2000d_nhea="Expenditure adjusted to year 2000 dollars and NHEA
cost adjustment factor"
    cost_2010d="Expenditure adjusted to year 2010 dollars"
    cost_2010d_nhea="Expenditure adjusted to year 2010 dollars and NHEA
cost adjustment factor";
if a;
run;

proc contents data=mcp.mc_pure&yr._cl&l.;
title "Contents: MC_Pure&yr.";
run;

data wgts.mc_pure&yr._wts_cl&l. (keep=baseid survey_wgt puremc_wgt
final_wgt);
    set mcp.mc_pure&yr._cl&l.;
run;

proc surveymeans data=mcp.mc_pure&yr._cl&l.;
stratum psu;
cluster strat;
weight final_wgt;
var cost_2010d_nhea;
title "With new survey weight, CL&l. (adjusted to yr 2010 $$ with NHEA
adjustment factor)";
run;

proc surveymeans data=mcp.mc_pure&yr._cl&l.;
stratum psu;
cluster strat;
weight survey_wgt;
var cost_2010d_nhea;

```

```
title "With old survey weight, CL&l. (adjusted to yr 2010 $$ with NHEA  
adjustment factor)";
```

```
run;
```

```
ods rtf file ="Insert file path\MC_Pure&yr._Contents.rtf";
```

```
proc contents data=mcp.mc_pure&yr._cl&l.;
```

```
title "Contents: MC_Pure&yr.";
```

```
run;
```

```
ods rtf close;
```