**\*\*\* Stata Code for “Prediction of whole blood methylmercury using blood total mercury measurements among fish and shellfish consumers”**

**\*\* 2020**

\* Create dataset with demographics, nutrient intakes first day, body measurements, total mercury – blood, and methyl mercury-blood datasets from NHANES 2011-2, 2013-4, 2015-6

\*data management\*

rename lbxthg thg

rename lbxbgm mehg

gen lthg=ln(thg)

gen lmehg=ln(mehg)

\*MeHg/THg ratio\*

gen mehgthg=mehg/thg

label variable mehgthg "MeHg/THg"

rename ridageyr age

gen agecat15=.

replace agecat15=0 if age>14

replace agecat15=1 if age>29

replace agecat15=2 if age>44

replace agecat15=3 if age>59

label variable agecat15 "0:15-29 yrs; 1:30-44 yrs; 2:45-59 yrs; 3:60+ yrs"

gen male=.

replace male=1 if riagendr==1

replace male=0 if riagendr==2

gen race=.

replace race=0 if ridreth3==3

replace race=1 if ridreth3==4

replace race=2 if ridreth3==1

replace race=2 if ridreth3==2

replace race=3 if ridreth3==6

replace race=4 if ridreth3==7

label variable race "0-NHwhite; 1-NHblack; 2-Hispanic; 3-NH Asian; 4-multi/other"

\*annual household income\*

gen income=0

replace income=1 if indhhin2==5

replace income=1 if indhhin2==6

replace income=1 if indhhin2==7

replace income=2 if indhhin2==8

replace income=2 if indhhin2==9

replace income=2 if indhhin2==10

replace income=0 if indhhin2==13

replace income=3 if indhhin2==14

replace income=3 if indhhin2==15

replace income=. if indhhin2==77

replace income=. if indhhin2==99

replace income=. if indhhin2==12

replace income=. if indhhin2==.

label variable income "0:<20K; 1:20 to <45K; 2:45K to 75K; 3:75K+"

rename bmxbmi bmi

gen bmicat=0

replace bmicat=1 if bmi>=25

replace bmicat=1 if bmdbmic==3

replace bmicat=2 if bmi>=30

replace bmicat=2 if bmdbmic==4

replace bmicat=. if (bmi==. & age>19)

replace bmicat=. if (bmdbmic==. & age<20)

label variable bmicat "0:under/normal wt; 1:overwt; 2:obese"

\*fish consumption\*

gen anyfish=0

replace anyfish=1 if drd340==1

replace anyfish=1 if drd360==1

label variable anyfish "reported eating fish or shellfish past 30 days"

gen shell=0

replace shell=1 if drd340==1

label variable shell "reported eating shellfish past 30 days"

gen fish=0

replace fish=1 if drd360==1

label variable fish "reported eating fish past 30 days"

\*\* # fish eaten in past 30 days \*\*

egen fishnum=rowtotal(drd370aq drd370bq drd370cq drd370dq drd370eq drd370fq ///

drd370gq drd370hq drd370iq drd370jq drd370kq drd370lq drd370mq drd370nq ///

drd370oq drd370pq drd370qq drd370rq drd370sq drd370tq drd370uq)

\*\* # shellfish eaten in past 30 days \*\*

egen shellnum=rowtotal(drd350aq drd350bq drd350cq drd350dq drd350eq ///

drd350fq drd350gq drd350hq drd350iq drd350jq)

\*\* total # fish+seafood eaten \*\*

egen fishshellnum=rowtotal(fishnum shellnum)

label variable fishnum "# fish eaten in past 30 days"

label variable shellnum "# shellfish eaten in past 30 days"

label variable fishshellnum "# fish or shellfish eaten in past 30 days"

xtile seanum4=fishshellnum if fishshellnum!=0,nq(4)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\* Generation of study datasets

gen pop=1

replace pop=0 if thg==.

replace pop=0 if mehg==.

replace pop=0 if income==.

replace pop=0 if bmicat==.

replace pop=0 if dr1drstz==2

replace pop=0 if age<15

replace pop=0 if anyfish==0

gen x=rnormal(0) if pop==1

gen test=0 if pop==1

replace test=1 if (x>.5 & pop==1)

label variable test "0:training; 1:test"

\* save dataset for further use

\*\*NOTE: generation of random variable varies each time it is run. Thus, when these commands are repeated the training/test dataset distribution may vary slightly.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*modify detection limit for THg 2011-2012

replace thg=.2 if (thg!=. & thg<.28 & pop==1)

\*detection limit for THg now effectively 0.28 across all years (.2 is LOD/sqrt(2))

\*generate subset population of women 15-44 yrs

gen popw=pop

replace popw=0 if male==1

replace popw=0 if agecat==2

replace popw=0 if agecat==3

tab popw pop

\*\*\* explore Hg values <LOD

\*MeHg < LOD

tab lbdbgmlc pop

tab lbdbgmlc popw

\*gen new lod marker for 2011-2012 (since we altered this manually above)

gen lodthg=lbdthglc

replace lodthg=1 if (thg<0.28 & pop==1)

replace lodthg=. if thg==.

tab lbdthglc lodthg if (pop==1 & test==0)

tab lodthg if (pop==1 & test==0)

tab lodthg if (popw==1 & test==0)

tab lodthg if (pop==1 & test==1)

tab lodthg if (popw==1 & test==1)

tab lbdbgmlc if (pop==1 & test==0)

tab lbdbgmlc if (popw==1 & test==0)

tab lbdbgmlc if (pop==1 & test==1)

tab lbdbgmlc if (popw==1 & test==1)

\*THg < LOD

tab lodthg pop

tab lodthg popw

\* is MeHg larger than THg?

gen ratio1=0

replace ratio1=1 if mehgthg>1

\*MeHg/THg >1

tab ratio1 pop

tab ratio1 popw

\*\*\*\*\*\*\*\* TABLE 1

tabulate agecat test if pop==1, chi2 column exact

tabulate male test if pop==1, chi2 column exact

tabulate race test if pop==1, chi2 column exact

tabulate income test if pop==1, chi2 column exact

tabulate bmicat test if pop==1, chi2 column exact

tabulate anyfish test if pop==1, chi2 column exact

tabulate seanum4 test if pop==1, chi2 column exact

tabulate agecat test if popw==1, chi2 column exact

tabulate race test if popw==1, chi2 column exact

tabulate income test if popw==1, chi2 column exact

tabulate bmicat test if popw==1, chi2 column exact

tabulate anyfish test if popw==1, chi2 column exact

tabulate seanum4 test if popw==1, chi2 column exact

\*\*\*\*\*\*\*\*\*\*\* TABLE 2

ameans thg mehg mehgthg if (pop==1 & test==0)

ameans thg mehg mehgthg if (pop==1 & test==1)

ameans thg mehg mehgthg if (pop==1)

regress lthg test if pop==1

regress lmehg test if pop==1

regress mehgthg test if pop==1

ameans thg mehg mehgthg if (popw==1 & test==0)

ameans thg mehg mehgthg if (popw==1 & test==1)

ameans thg mehg mehgthg if (popw==1)

regress lthg test if popw==1

regress lmehg test if popw==1

regress mehgthg test if popw==1

\*\* SOME GRAPHICS

\*option 1: regular graph

twoway (scatter mehg thg if (pop==1 & test==0), mcolor(gs12) msize(small) msymbol(circle\_hollow)) ///

(lowess mehg thg if pop==1, lcolor(maroon) lwidth(medium) lpattern(solid)) ///

(lfit mehg thg if pop==1, lcolor(black) lwidth(medium) lpattern(dash)) ///

, ytitle("Methylmercury, µg/L") xtitle("Total mercury, µg/L") ///

legend(off) scheme(s1color) fxsize(100) fysize(100) name(hg, replace)

\*option 2: log-transformed Hg

twoway (scatter mehg thg if (pop==1 & test==0), mcolor(gs12) msize(small) msymbol(circle\_hollow)) ///

(lowess mehg thg if (pop==1 & test==0), lcolor(maroon) lwidth(medium) lpattern(solid)) ///

(lfit mehg thg if (pop==1 & test==0), lcolor(black) lwidth(medium) lpattern(dash) range(0.2 50.81)) ///

, ytitle("Methylmercury, µg/L") yscale(log) ///

ylabel(0.5 5 50) xtitle("Total mercury, µg/L") xscale(log) ///

xlabel(0.5 5 50) legend(off) scheme(s1color) fxsize(100) fysize(100) name(hglog, replace)

graph combine hg hglog, col(2) scheme(s1color) imargin(1 1 1 1) xsize(8) ysize(4)

\*\* box plot

xtile thgquin=thg, nq(5)

\*box plot

graph box mehgthg if test==0, over(thgquin) box(1, fcolor(maroon) ///

fintensity(inten70) lcolor(maroon) lwidth(thin) lpattern(solid)) ///

marker(1, mcolor(maroon) msize(small) msymbol(smcircle)) ytitle("MeHg/THg") ///

caption("THg, µg/L", position(6)) legend(off) scheme(s1color)

sort thgquin

by thgquin: summ thg if (test==0 & pop==1)

by thgquin: summ mehgthg if (test==0 & pop==1)

\*\*\*\*\*\*\*\*\*\* SUPPLEMENTAL TABLE 1

\*\*15+ seafood consumers

\*age category

sort agecat15

by agecat15: ameans thg mehg if pop==1

mean mehgthg if pop==1, over(agecat15)

regress lthg i.agecat15 if pop==1

regress lmehg i.agecat15 if pop==1

regress mehgthg i.agecat15 if pop==1

\*sex

sort male

by male: ameans thg mehg if pop==1

mean mehgthg if pop==1, over(male)

regress lthg male if pop==1

regress lmehg male if pop==1

regress mehgthg male if pop==1

\*race

sort race

by race: ameans thg mehg if pop==1

mean mehgthg if pop==1, over(race)

regress lthg i.race if pop==1

regress lmehg i.race if pop==1

regress mehgthg i.race if pop==1

\*income

sort income

by income: ameans thg mehg if pop==1

mean mehgthg if pop==1, over(income)

regress lthg i.income if pop==1

regress lmehg i.income if pop==1

regress mehgthg i.income if pop==1

\*bmi

sort bmicat

by bmicat: ameans thg mehg if pop==1

mean mehgthg if pop==1, over(bmicat)

regress lthg i.bmicat if pop==1

regress lmehg i.bmicat if pop==1

regress mehgthg i.bmicat if pop==1

\*\*15-44 female seafood consumers

\*age category

sort agecat15

by agecat15: ameans thg mehg if popw==1

mean mehgthg if popw==1, over(agecat15)

regress lthg i.agecat15 if popw==1

regress lmehg i.agecat15 if popw==1

regress mehgthg i.agecat15 if popw==1

\*race

sort race

by race: ameans thg mehg if popw==1

mean mehgthg if popw==1, over(race)

regress lthg i.race if popw==1

regress lmehg i.race if popw==1

regress mehgthg i.race if popw==1

\*income

sort income

by income: ameans thg mehg if popw==1

mean mehgthg if popw==1, over(income)

regress lthg i.income if popw==1

regress lmehg i.income if popw==1

regress mehgthg i.income if popw==1

\*bmi

sort bmicat

by bmicat: ameans thg mehg if popw==1

mean mehgthg if popw==1, over(bmicat)

regress lthg i.bmicat if popw==1

regress lmehg i.bmicat if popw==1

regress mehgthg i.bmicat if popw==1

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\*\*\*\*\*\* make splines

\* make tHg splines

mkspline thg5a .5 thg5b = thg

mkspline thg1a 1 thg1b = thg

mkspline thg2a 2 thg2b = thg

mkspline thg25a 25 thg25b = thg

mkspline thgcub=thg, cubic

\*\*\*\*\*\*\*\*\*\*\*\* For table 3

\*training dataset linear models (unadjusted/adjusted) all seafood consumers

regress mehg thg if (pop==1 & test==0)

estimates stats

predict p1, xb

predict r1, resid

predict l1, leverage

summ mehg p1 r1 l1 if (pop==1 & test==0), detail

twoway (scatter r1 p1), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Training Dataset") subtitle("Unadjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Training Dataset") subtitle("Unadjusted linear model") scheme(s1color)

summ r1 l1 if (pop==1 & test==0), detail

gen ex1=0 if (pop==1 & test==0)

replace ex1=1 if (r1<-1.002586 & l1>.0037383 & pop==1 & test==0)

replace ex1=1 if (r1>.774267 & l1>.0037383 & pop==1 & test==0)

sort thg

list thg mehg lbxihg if ex1==1

regress mehg thg age male i.race i.income i.bmicat if (pop==1 & test==0)

estimates stats

predict p2, xb

predict r2, resid

predict l2, leverage

summ mehg p2 r2 l2 if (pop==1 & test==0), detail

twoway (scatter r2 p2), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Training Dataset") subtitle("Adjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Training Dataset") subtitle("Adjusted linear model") scheme(s1color)

summ r2 l2 if (pop==1 & test==0), detail

gen ex2=0 if (pop==1 & test==0)

replace ex2=1 if (r2<-1.020614 & l2>.0071184 & pop==1 & test==0)

replace ex2=1 if (r2>.7618847 & l2>.0071184 & pop==1 & test==0)

sort thg

list thg mehg lbxihg if ex2==1

\*training dataset spline models (unadjusted/adjusted) all seafood consumers

regress mehg thg1a thg1b if (pop==1 & test==0)

estimates stats

predict p3, xb

predict r3, resid

predict l3, leverage

summ mehg p3 r3 l3 if (pop==1 & test==0), detail

twoway (scatter r3 p3), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Training Dataset") subtitle("Unadjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Training Dataset") subtitle("Unadjusted spline model") scheme(s1color)

summ r3 l3 if (pop==1 & test==0), detail

gen ex3=0 if (pop==1 & test==0)

replace ex3=1 if (r3<-1.009971 & l3>.0040523 & pop==1 & test==0)

replace ex3=1 if (r3>.771104 & l3>.0040523 & pop==1 & test==0)

sort thg

list thg mehg lbxihg if ex3==1

regress mehg thg1a thg1b age male i.race i.income i.bmicat if (pop==1 & test==0)

estimates stats

predict p4, xb

predict r4, resid

predict l4, leverage

summ mehg p4 r4 l4 if (pop==1 & test==0), detail

twoway (scatter r4 p4), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Training Dataset") subtitle("Adjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Training Dataset") subtitle("Adjusted spline model") scheme(s1color)

summ r4 l4 if (pop==1 & test==0), detail

gen ex4=0 if (pop==1 & test==0)

replace ex4=1 if (r4<-1.019492 & l4>.0074372 & pop==1 & test==0)

replace ex4=1 if (r4> .7491542 & l4>.0074372 & pop==1 & test==0)

sort thg

list thg mehg lbxihg if ex4==1

\*training dataset linear models (unadjusted/adjusted) reprodutive

regress mehg thg if (popw==1 & test==0)

estimates stats

predict p5, xb

predict r5, resid

predict l5, leverage

summ mehg p5 r5 l5 if (popw==1 & test==0), detail

twoway (scatter r5 p5), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Training Dataset") subtitle("Unadjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Training Dataset") subtitle("Unadjusted linear model") scheme(s1color)

summ r5 l5 if (popw==1 & test==0), detail

gen ex5=0 if (popw==1 & test==0)

replace ex5=1 if (r5<-.9828732 & l5> .0153277 & popw==1 & test==0)

replace ex5=1 if (r5>.792838 & l5>.0153277 & popw==1 & test==0)

sort thg

list thg mehg lbxihg if ex5==1

regress mehg thg age i.race i.income i.bmicat if (popw==1 & test==0)

estimates stats

predict p6, xb

predict r6, resid

predict l6, leverage

summ mehg p6 r6 l6 if (popw==1 & test==0), detail

twoway (scatter r6 p6), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Training Dataset") subtitle("Adjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Training Dataset") subtitle("Adjusted linear model") scheme(s1color)

summ r6 l6 if (popw==1 & test==0), detail

gen ex6=0 if (popw==1 & test==0)

replace ex6=1 if (r6<-.9607773 & l6>.0264746 & popw==1 & test==0)

replace ex6=1 if (r6>.7446032 & l6>.0264746 & popw==1 & test==0)

sort thg

list thg mehg lbxihg if ex6==1

\*training dataset spline models (unadjusted/adjusted) reprodutive

regress mehg thg1a thg1b if (popw==1 & test==0)

estimates stats

predict p7, xb

predict r7, resid

predict l7, leverage

summ mehg p7 r7 l7 if (popw==1 & test==0), detail

twoway (scatter r7 p7), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Training Dataset") subtitle("Unadjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Training Dataset") subtitle("Unadjusted spline model") scheme(s1color)

summ r7 l7 if (popw==1 & test==0), detail

gen ex7=0 if (popw==1 & test==0)

replace ex7=1 if (r7<-.9483767 & l7>.0176363 & popw==1 & test==0)

replace ex7=1 if (r7>.7993576 & l7>.0176363 & popw==1 & test==0)

sort thg

list thg mehg lbxihg if ex7==1

regress mehg thg1a thg1b age i.race i.income i.bmicat if (popw==1 & test==0)

estimates stats

predict p8, xb

predict r8, resid

predict l8, leverage

summ mehg p8 r8 l8 if (popw==1 & test==0), detail

twoway (scatter r8 p8), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Training Dataset") subtitle("Adjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Training Dataset") subtitle("Adjusted spline model") scheme(s1color)

summ r8 l8 if (popw==1 & test==0), detail

gen ex8=0 if (popw==1 & test==0)

replace ex8=1 if (r8<-.9377695 & l8>.0281342 & popw==1 & test==0)

replace ex8=1 if (r8>.7523971 & l8>.0281342 & popw==1 & test==0)

sort thg

list thg mehg lbxihg if ex8==1

\*\*\*\*\*\*\*\*\*\*\*\* for Table 4

\* adult population, linear models

\*original model

regress mehg thg if (pop==1 & test==0)

\*predictions

predict testp1 if (pop==1 & test==1), xb

predict testr1 if (pop==1 & test==1), resid

predict testl1 if (pop==1 & test==1), leverage

summ mehg testp1 testr1 testl1 if (pop==1 & test==1), detail

twoway (scatter testr1 testp1), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Test Dataset") subtitle("Unadjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Test Dataset") subtitle("Unadjusted linear model") scheme(s1color)

\*original model

regress mehg thg age male i.race i.income i.bmicat if (pop==1 & test==0)

\*predictions

predict testp2 if (pop==1 & test==1), xb

predict testr2 if (pop==1 & test==1), resid

predict testl2 if (pop==1 & test==1), leverage

summ mehg testp2 testr2 testl2 if (pop==1 & test==1), detail

twoway (scatter testr2 testp2), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Test Dataset") subtitle("Adjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Test Dataset") subtitle("Adjusted linear model") scheme(s1color)

\* adult population, spline models

\*original model

regress mehg thg1a thg1b if (pop==1 & test==0)

\*predictions

predict testp3 if (pop==1 & test==1), xb

predict testr3 if (pop==1 & test==1), resid

predict testl3 if (pop==1 & test==1), leverage

summ mehg testp3 testr3 testl3 if (pop==1 & test==1), detail

twoway (scatter testr3 testp3), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Test Dataset") subtitle("Unadjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Test Dataset") subtitle("Unadjusted spline model") scheme(s1color)

\*original model

regress mehg thg1a thg1b age male i.race i.income i.bmicat if (pop==1 & test==0)

\*predictions

predict testp4 if (pop==1 & test==1), xb

predict testr4 if (pop==1 & test==1), resid

predict testl4 if (pop==1 & test==1), leverage

summ mehg testp4 testr4 testl4 if (pop==1 & test==1), detail

twoway (scatter testr4 testp4), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Adults - Test Dataset") subtitle("Adjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Adults - Test Dataset") subtitle("Adjusted spline model") scheme(s1color)

\* reproductive population, linear models

\*original model

regress mehg thg if (popw==1 & test==0)

\*predictions

predict testp5 if (popw==1 & test==1), xb

predict testr5 if (popw==1 & test==1), resid

predict testl5 if (popw==1 & test==1), leverage

summ mehg testp5 testr5 testl5 if (popw==1 & test==1), detail

twoway (scatter testr5 testp5), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Test Dataset") subtitle("Unadjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Test Dataset") subtitle("Unadjusted linear model") scheme(s1color)

\*original model

regress mehg thg age male i.race i.income i.bmicat if (popw==1 & test==0)

\*predictions

predict testp6 if (popw==1 & test==1), xb

predict testr6 if (popw==1 & test==1), resid

predict testl6 if (popw==1 & test==1), leverage

summ mehg testp6 testr6 testl6 if (popw==1 & test==1), detail

twoway (scatter testr6 testp6), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Test Dataset") subtitle("Adjusted linear model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Test Dataset") subtitle("Adjusted linear model") scheme(s1color)

\* reproductive population, spline models

\*original model

regress mehg thg1a thg1b if (popw==1 & test==0)

\*predictions

predict testp7 if (popw==1 & test==1), xb

predict testr7 if (popw==1 & test==1), resid

predict testl7 if (popw==1 & test==1), leverage

summ mehg testp7 testr7 testl7 if (popw==1 & test==1), detail

twoway (scatter testr7 testp7), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Test Dataset") subtitle("Unadjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Test Dataset") subtitle("Unadjusted spline model") scheme(s1color)

\*original model

regress mehg thg1a thg1b age male i.race i.income i.bmicat if (popw==1 & test==0)

\*predictions

predict testp8 if (popw==1 & test==1), xb

predict testr8 if (popw==1 & test==1), resid

predict testl8 if (popw==1 & test==1), leverage

summ mehg testp8 testr8 testl8 if (popw==1 & test==1), detail

twoway (scatter testr8 testp8), yline(0, lwidth(thin) lpattern(longdash) ///

lcolor(cranberry)) title("Reproductive - Test Dataset") subtitle("Adjusted spline model") ///

scheme(s1mono)

lvr2plot, mcolor(black) title("Reproductive - Test Dataset") subtitle("Adjusted spline model") scheme(s1color)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\* TABLE: 5 Predicted MeHg vs measured THg

\*\*\* use blank dataset\*\*

set obs 500

generate x=\_n

gen thg=x/10

\*\* predicted MeHg values

\*unadjusted continuous

gen linad=-0.17+(thg\*1.01)

gen linrep=-0.15+(thg\*0.98)

\*unadjusted spline

gen spad=-0.08+(thg\*.88)

replace spad=0.80+((thg-1)\*1.01) if thg>=1

gen sprep=-0.08+(thg\*.88)

replace sprep=.80+((thg-1)\*0.99) if thg>=1

\*for values not shown in estimates above

\*adults, unadjusted spline

display -0.08+(0.75\*.88)

display 0.80+((1.25-1)\*1.01)

display 0.80+((1.75-1)\*1.01)

\*reproductive, unadjusted spline

display -0.08+(0.75\*.88)

display 0.80+((1.25-1)\*0.99)

display 0.80+((1.75-1)\*0.99)

\*\*\*\*\*\*\*\*\*\*

\*\*plot for graphical abstract -- thg vs predited mehg values

sort thg

twoway (scatter mehg thg if (pop==1 & test==1 & thg<10), mcolor(gs12) msymbol(circle\_hollow)) ///

(connected p3 thg if (pop==1 & test==1 & thg<10), msymbol(none) lcolor(cranberry) lwidth(medium)) ///

, ytitle("Methylmercury,µg/L") xtitle("Total mercury, µg/L") scheme(s1color) ///

legend(on cols(1) position(5) ring(0))

twoway (scatter mehg thg if (pop==1 & test==1), mcolor(gs12) msymbol(circle\_hollow)) ///

(connected p3 thg if (pop==1 & test==1), msymbol(none) lcolor(cranberry) lwidth(medthin)) ///

, ytitle("Methylmercury,µg/L") xtitle("Total mercury, µg/L") scheme(s1color) legend(off)