Advanced Methods in Health Services Research: Analysis - 309.716
Tuesday and Thursday 9:00-10:20
Instructors: Darrell J. Gaskin, Ph.D. and Roland J. Thorpe, Ph.D.
250 Hampton House

Computer Exercise #3: Estimating the Effect of Health Status, Poverty Status, and Census Region on having an Office Based Physician Visits using LPM, Logistic, and Probit Regression Models

Due: September 20, 2016

Answer Guide

Use the analysis file you built in computer exercise #1. For this analysis, limit the sample to adults, i.e., persons over the age of 17.

Using the office based physician variable *obtotv08* create a dummy variable indicating if a person had an office based physician visit during the year.

1. What are the strengths and weaknesses of the linear probability model?
   The main strength of the linear probability model is that the interpretation of coefficients on your independent variable are straightforward and easier to interpret as you would a usual OLS regression – that they are marginal effects of the covariates on the outcome. It has a number of weaknesses. First, predictions of the outcome variable may fall outside of one and zero, i.e., less than zero or more than one. These would be nonsensical since probabilities must fall between zero and one. Another weakness is the non-normality of errors – they are binomial. Heteroskedasticity related to the β is also present - β is inefficient and the standard errors are biased. Thus, some predictions and hypothesis tests may be biased and inaccurate. In addition, as with OLS, the marginal effects are constant unless the model is nonlinear in the parameters.

2. What is the fundamental difference between the logistic regression and probit regression models?
   The main difference lies in the assumption regarding the error terms in the model. In logistic regression, the errors are assumed to follow the logistic distribution while in probit regression, the errors are assumed to follow the normal distribution. That is, the link functions, i.e., the function that relates the mean of the response to the predictors of the model, of the two are different. Graphically, the logistic function has flatter tails and the probit function approaches the axes more quickly than the curve. However, in practice the logit and the probit link functions give very similar outputs when given the same inputs. Thus, researchers sometimes use logistic and probit regression interchangeably.

3. Estimate the following model using linear probability model (LPM), logistic, and probit regression techniques.

Any obvisit = f(health status, age, race/ethnicity, sex, poverty status, education, insurance status and location) *Use the following reference categories: excellent health, age category 25-44, male, white non-Hispanic, HS diploma/GED, high income, privately insured, msa, East*

a. Create table 1 that displays the coefficients for each model for the race/ethnicity and sex.

TABLE 1. LPM, Logit, Probit Model Coefficients

	(1) LPM	(2) Logit	(3) Probit
visit			
Black	-0.118^{***}	-0.690^{***}	-0.409^{***}
	(0.0104)	(0.0582)	(0.0345)
Hispanic	-0.115^{***}	-0.620^{***}	-0.367^{***}
	(0.0131)	(0.0716)	(0.0423)
Asian	-0.141^{***}	-0.853^{***}	-0.493^{***}
	(0.0160)	(0.0847)	(0.0500)
Other	-0.0335	-0.242	-0.144
	(0.0342)	(0.210)	(0.118)
Female	0.139^{***}	0.863^{***}	0.496^{***}
	(0.00568)	(0.0374)	(0.0214)
N	31323	31323	31323
R^2	0.205		

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

b. Interpret the coefficients for race/ethnicity and sex from each model.

For the linear probability model, being Black, Hispanic, or Asian decreases the predicted probability of having any physician office visit by 0.118, 0.115, and 0.141 compared to being White, respectively, holding all other variables constant. These are statistically significant. The effect of belonging to other races is not statistically significant. On the other hand, being female increases the predicated probability of having any physician office visit by 0.139 compared to being male, holding all other things constant and this is also statistically significant.

For the logit model, being Black, Hispanic, or Asian decreases the log odds of having any physician office visit by 0.69, 0.62, and 0.853 compared to being White, respectively, holding all other variables constant. These are statistically significant. The effect of belonging to other races is also not statistically significant. On the other hand, being female increases the log odds of having any physician office visit by 0.863 compared to being male and this is also statistically significant.

For the probit model, being Black, Hispanic, or Asian, the probability of having any physician office visit is associated with having a change in predicted z-score of -0.409, -0.367, and -0.493 compared to being White, holding all other things constant. These are all statistically significant. The effect of belonging to other races is also not statistically significant. On the other hand, for females, the probability of having any physician office visit is associated with having a predicted z-score of 0.869 compared to being male. This is also statistically significant.

c. Compute and display in table 2 the odds ratios from the logistic regression model for race/ethnicity and sex. Interpret the odds ratios.

TABLE 2. Logistic Regression Odds Ratios.

	(1) Odds Ratio
visit	
Black	0.502***
	(0.0292)
Hispanic	0.538***
	(0.0385)
Asian	0.426***
	(0.0361)
Other Race	0.785
	(0.165)
Female	2.371***
	(0.0887)
N	31323

Exponentiated coefficients; Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

For Blacks, the odds of having any physician office visits is 0.502 times that of the odds for Whites. For Hispanics, the odds of having any physician office visits is 0.538 times that of the odds for Whites. For Asians, the odds of having any physician office visits is 0.426 times that of the odds for Whites. These are all statistically significant. The odds ratio for other races is not found to be statistically significant.

The odds of women having any physician office visits is 2.37 times the odds of men having any physician office visits. This is found to be statistically significant.

d. Compute and display in table 3 the marginal effects for race/ethnicity and sex from the probit and logistic regression models using the margin command in Stata? Interpret the marginal effects.

TABLE 3. Logistic and Probit Marginal Effects.

	(1) Logit, Marginal Effects	(2) Probit, Marginal Effects
Black	-0.121***	-0.122***
	(0.0106)	(0.0107)
Hispanic	-0.108***	-0.108***
	(0.0130)	(0.0131)
Asian	-0.152***	-0.149***
	(0.0161)	(0.0161)
Other Race	-0.0399	-0.0406
	(0.0358)	(0.0342)
Female	0.146***	0.143***
	(0.00619)	(0.00607)
N	26790	26790

Marginal effects; Standard errors in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We first interpret the logit model marginal effects. Relative to white individuals, black, Hispanic, and Asian individuals have 0.121, 0.108, and 0.152 lower probabilities of having any office-based provider visit. These have been found to be statistically significant. In addition, compared to males, females have a 0.146 higher probability of having any office-based provider visit.

We next interpret the probit model marginal effects. Relative to white individuals, black, Hispanic, and Asian individuals have 0.122, 0.108, and 0.149 lower probabilities of having any office-based provider visit. These have been found to be statistically significant. In addition, compared to males, females have a 0.143 higher probability of having any office-based provider visit.

4. Predict the dependent variable using all three techniques. Compute the mean, standard deviations, minimum and maximum values the actual and predicted dependent variables. Display your results in table 4. Discuss how and why they are similar and/or different.

We see that the predicted means for the LPM, logit, and probit models are quite similar to the actual mean from our data. However, we see that the standard deviation for the actual data is much higher than those computed for the three models. This may be due to the fact that the actual data only contains 0s and 1s while the three models have predicted values in the (0,1) interval since represent latent probabilities. We also note that the maximum predicted value in the LPM is 1.268, which is greater than one and does not make sense in the context of our analysis. This represents one of the weaknesses of the LPM that was discussed in number 1.

TABLE 4. Actual and Predicted Values.

Variable	Mean	SD	Min	Max
Actual	0.667	0.471	0.000	1.000
Linear	0.657	0.202	0.104	1.268
Logit	0.652	0.220	0.097	0.997
Probit	0.652	0.216	0.093	0.999

Do-file

```
* Advanced methods in health services research: analysis - 309.716
* September 20, 2016
* Assignment # 3: Suggested code


set varabbrev off
log using "/Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods -
Analysis/Answers/Assignment3_2016.log", replace
use "/Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods - Analysis/meps08.dta"
cd "/Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods - Analysis/Answers"
svyset varpsu [pweight = perwt08f], strata(varstr)

** Data preparation
* Trimming/cleaning the data
summ totexp08
summ totexp08, detail

gen healthexp = totexp08
replace healthexp = . if totexp08 < 0 /*replaces negative values with missing*/
replace healthexp = . if totexp08 > 100000 /*replaces values > $100,000 with missing*/
summ healthexp
summ healthexp, detail

gen income = ttlp08x
replace income = . if ttlp08x < 0 /*sets negative values to missing*/
replace income = . if ttlp08x > 170000 /*sets very large values to missing*/
/* Note: If you did not exclude those with incomes > 170000, it's fine*/
summarize income
summarize income, detail

* Generate categories
summarize age08x
gen age = age08x
replace age = . if age08x < 0 /*sets negative values to missing*/
summarize age

gen agecat = 1 if age08x >=0 & age08x <=24
replace agecat = 2 if age08x >=25 & age08x <=44
replace agecat = 3 if age08x >=45 & age08x <=64
replace agecat = 4 if age08x >=65 & age08x <=74
replace agecat = 5 if age08x >=75
label define agecats 1 "18-24" 2  "25-44" 3 "45-64" 4 "65-74" 5 "75+"
label values agecat agecats /*set label name sexn to the variable sex*/

gen female = 0 if sex == 1
replace female = 1 if sex == 2

label define sexn 0 "Male" 1 "Female"
label values female sexn

label define raceethn 1 "Hispanic" 2 "Black non-Hispanic" 3 "Asian non-Hispanic" 4 "Other race/not Hispanic"
```

```
label values racethnx raceethn

label define racexn 1 "White" 2 "Black" 3 "Amer Indian/Alaska Native" 4 "Asian" 5 "Native Hawaiian/Pacific Islander" 6 "Multiple races reported"
label values racex racexn

gen race = 1 if racex == 1 & racethnx != 1
replace race = 2 if racex == 2 & racethnx != 1
replace race = 3 if racethnx == 1
replace race = 4 if racex == 4 & racethnx != 1
replace race = 5 if (racex == 3 | racex == 5 | racex == 6) & racethnx != 1

label define racexn2 1 "White" 2 "Black" 3 "Hispanic" 4 "Asian" 5 "Other"
label values race racexn2

gen education = 1 if educyr >=0 & educyr <=8
replace education = 2 if educyr >=9 & educyr <=12 & hideg == 1 /*Some High School*/
replace education = 2 if educyr >=9 & educyr <=11 & hideg <0 /* Some High School but didn't answer hideg question*/
replace education = 3 if hideg == 2 | hideg == 3 /*High School*/
replace education = 4 if educyr >=13 & educyr <=17 & hideg == 3 /*Some college-In school but only have high school diploma*/
replace education = 5 if hideg == 4
replace education = 6 if hideg == 5 | hideg ==6
replace education = 6 if educyr == 17 & hideg == 7 /*Advanced Degree* - implies law degree or similar type*/

label define educn 1 "No High School" 2 "Some High" 3 "High School/GED" 4 "Some College/Tech School/AA degree" 5 "College"  6 "Advanced Degree"
label values education educn

gen insurance = 1 if inscov08 ==1
replace insurance = 2 if inscov08 == 2 & mcrev08 == 2 /*Had public insurance but not medicare*/
replace insurance = 3 if inscov08 == 3
replace insurance = 4 if inscov08 == 2 & mcrev08 == 1 /* Medicare */

label define insr 1 "Private Insurance" 2 "Public Insurance" 3 "Uninsured" 4 "Medicare"
label values insurance insr

gen msa = 0 if msa08 == 0
replace msa = 1 if msa08 == 1

label define msan 0 "Not in MSA" 1 "In MSA"
label values msa msan

gen region = 1 if region08 == 1
replace region = 2 if region08 == 3
replace region = 3 if region08 == 2
replace region = 4 if region08 == 4

label define regi 1 "North East" 2 "South" 3 "Midwest" 4 "West"
label values region regi

gen healthstatus = 1 if rthlth42 == 1
```

```
replace healthstatus = 2 if rthlth42 == 2
replace healthstatus = 3 if rthlth42 == 3
replace healthstatus = 4 if rthlth42 == 4
replace healthstatus = 5 if rthlth42 == 5

label define health 1 "Excellent" 2 "Very Good" 3 "Good" 4 "Fair" 5 "Poor"
label values healthstatus health

gen fpl = 1 if povcat08 == 1
replace fpl = 2 if povcat08 == 2
replace fpl = 3 if povcat08 == 3
replace fpl = 4 if povcat08 == 4
replace fpl = 5 if povcat08 == 5

label define fplstat 1 "Poor" 2 "Near Poor" 3 "Low Income" 4 "Middle Income" 5 "High Income"
label values fpl fplstat

* Generate indicator variable for adult
gen adult = 0
replace adult = 1 if age08x > 17

* Generate indicator variable for office visit
gen visit = 0
replace visit = 1 if obtotv08 > 0

** Number 3 **
* Table 1 Construction
* Linear probability model
svy, subpop(adult): reg visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region

eststo lpm

* Logit regression
svy, subpop(adult): logit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region

eststo logit1

* Probit regression
svy, subpop(adult): probit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region

eststo probit1

* Actual Table 1
esttab lpm logit1 probit1 using "hw3table1.rtf", se replace r2 keep(1.female 2.race 3.race 4.race 5.race) ///
mtitles(LPM Logit Probit) coeflabels(2.race Black 3.race Hispanic 4.race Asian ///
5.race Other 1.female Female) title(TABLE 1. LPM, Logit, Probit Model ///
Coefficients)

* Table 2 Construction
* Logistic Regression for Odds Ratios
```

```
svy, subpop(adult): logistic visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education i.insurance ib1.msa ib1.region

eststo logistic1

* Actual Table 2
esttab logistic1 using "hw3table2.rtf", se replace eform keep(1.female 2.race 3.race 4.race 5.race) ///
mtitles(Odds Ratios) coeflabels(2.race Black 3.race Hispanic 4.race Asian ///
5.race "Other Race" 1.female Female) title(TABLE 2. Logistic Regression Odds Ratios.)

* Table 3 Construction
* Logit regression, marginal effects
svy, subpop(adult): logit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region

margins, dydx(race female) post
est store margins1

* Probit regression
svy, subpop(adult): probit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region

margins, dydx(race female) post
est store margins2

* Actual Table 3
esttab margins1 margins2 using hw3table3.rtf, se replace margin keep(1.female 2.race 3.race 4.race 5.race) ///
mtitles("Logit, Marginal Effects" "Probit, Marginal Effects") ///
coeflabels(2.race Black 3.race Hispanic 4.race Asian ///
5.race "Other Race" 1.female Female) title(TABLE 3. Logistic and Probit Marginal Effects.)

** Number 4 **
svy, subpop(adult): reg visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
predict predlpm
label var predlpm "Predicted Values - LPM"

svy, subpop(adult): logit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
predict predlogit
label var predlogit "Predicted Values - Logit"

svy, subpop(adult): probit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
predict predprobit
label var predprobit "Predicted Values - Probit"

tabstat visit predlpm predlogit predprobit, statistics(mean sd min max) col(stat) ///
format(%9.3f)
```

Log file

```
------------------------------------------------------------------------------------------------------
------
      name:  <unnamed>
       log:  /Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods -
Analysis/Answers/Assignment3_2016.lo
> g
  log type:  text
 opened on:  22 Sep 2016, 01:30:34

. use "/Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods - Analysis/meps08.dta"

. cd "/Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods - Analysis/Answers"
/Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods - Analysis/Answers

. svyset varpsu [pweight = perwt08f], strata(varstr)

      pweight: perwt08f
          VCE: linearized
  Single unit: missing
     Strata 1: varstr
         SU 1: varpsu
        FPC 1: <zero>

.
. ** Data preparation
. * Trimming/cleaning the data
. summm totexp08

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    totexp08 |     33066    3142.069    9786.619          0     553493

. summ totexp08, detail

                  total health care exp 08
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs               33066
25%           67              0       Sum of Wgt.       33066

50%        528.5                      Mean           3142.069
                        Largest       Std. Dev.      9786.619
75%         2425         238659
90%         7453         264510       Variance       9.58e+07
95%        13582         373799       Skewness       14.40863
99%        40763         553493       Kurtosis        469.447

.
. gen healthexp = totexp08

. replace healthexp = . if totexp08 < 0 /*replaces negative values with missing*/
(0 real changes made)

. replace healthexp = . if totexp08 > 100000 /*replaces values > $100,000 with missing*/
(42 real changes made, 42 to missing)

. summ healthexp

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   healthexp |     33024    2937.216    7412.106          0      99988

. summ healthexp, detail

                          healthexp
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs               33024
25%           66              0       Sum of Wgt.       33024
```

```
50%          526                      Mean           2937.216
                        Largest       Std. Dev.      7412.106
75%       2413.5          98210
90%         7353          99251       Variance       5.49e+07
95%        13387          99264       Skewness       5.819768
99%        38550          99988       Kurtosis       48.38959

.
. gen income = ttlp08x

. replace income = . if ttlp08x < 0 /*sets negative values to missing*/
(25 real changes made, 25 to missing)

. replace income = . if ttlp08x > 170000 /*sets very large values to missing*/
(325 real changes made, 325 to missing)

. /* Note: If you did not exclude those with incomes > 170000, it's fine*/
. summarize income

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      income |     32716    18715.83    24109.08          0     169564

. summarize income, detail

                           income
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs               32716
25%            0              0       Sum of Wgt.       32716

50%        10000                      Mean           18715.83
                        Largest       Std. Dev.      24109.08
75%        28838         164895
90%        52000         164895       Variance       5.81e+08
95%        70000         165391       Skewness       1.717878
99%       103084         169564       Kurtosis       6.147364

.
. * Generate categories
. summarize age08x

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      age08x |     33066    33.69395    22.38051         -1         85

. gen age = age08x

. replace age = . if age08x < 0 /*sets negative values to missing*/
(251 real changes made, 251 to missing)

. summarize age

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         age |     32815    33.95932    22.25851          0         85

.
. gen agecat = 1 if age08x >=0 & age08x <=24
(20062 missing values generated)

. replace agecat = 2 if age08x >=25 & age08x <=44
(8813 real changes made)

. replace agecat = 3 if age08x >=45 & age08x <=64
(7614 real changes made)

. replace agecat = 4 if age08x >=65 & age08x <=74
(1867 real changes made)

. replace agecat = 5 if age08x >=75
(1517 real changes made)

. label define agecats 1 "18-24" 2  "25-44" 3 "45-64" 4 "65-74" 5 "75+"
```

```
. label values agecat agecats /*set label name sexn to the variable sex*/

. 
. gen female = 0 if sex == 1
(17181 missing values generated)

. replace female = 1 if sex == 2
(17181 real changes made)

. 
. label define sexn 0 "Male" 1 "Female"

. label values female sexn

. 
. label define raceethn 1 "Hispanic" 2 "Black non-Hispanic" 3 "Asian non-Hispanic" 4 "Other race/not 
Hispani
> c"

. label values racethnx raceethn

. 
. label define racexn 1 "White" 2 "Black" 3 "Amer Indian/Alaska Native" 4 "Asian" 5 "Native 
Hawaiian/Pacific
>  Islander" 6 "Multiple races reported"

. label values racex racexn

. 
. gen race = 1 if racex == 1 & racethnx != 1
(18775 missing values generated)

. replace race = 2 if racex == 2 & racethnx != 1
(6476 real changes made)

. replace race = 3 if racethnx == 1
(9392 real changes made)

. replace race = 4 if racex == 4 & racethnx != 1
(1997 real changes made)

. replace race = 5 if (racex == 3 | racex == 5 | racex == 6) & racethnx != 1
(910 real changes made)

. 
. label define racexn2 1 "White" 2 "Black" 3 "Hispanic" 4 "Asian" 5 "Other"

. label values race racexn2

. 
. gen education = 1 if educyr >=0 & educyr <=8
(25465 missing values generated)

. replace education = 2 if educyr >=9 & educyr <=12 & hideg == 1 /*Some High School*/
(3732 real changes made)

. replace education = 2 if educyr >=9 & educyr <=11 & hideg <0 /* Some High School but didn't answer 
hideg q
> uestion*/
(3 real changes made)

. replace education = 3 if hideg == 2 | hideg == 3 /*High School*/
(11173 real changes made)

. replace education = 4 if educyr >=13 & educyr <=17 & hideg == 3 /*Some college-In school but only 
have hig
> h school diploma*/
(3629 real changes made)

. replace education = 5 if hideg == 4
(3245 real changes made)

. replace education = 6 if hideg == 5 | hideg ==6
(1590 real changes made)
```

```
. replace education = 6 if educyr == 17 & hideg == 7 /*Advanced Degree* - implies law degree or
similar type
> */
(66 real changes made)

.
. label define educn 1 "No High School" 2 "Some High" 3 "High School/GED" 4 "Some College/Tech
School/AA deg
> ree" 5 "College"  6 "Advanced Degree"

. label values education educn

.
. gen insurance = 1 if inscov08 ==1
(14773 missing values generated)

. replace insurance = 2 if inscov08 == 2 & mcrev08 == 2 /*Had public insurance but not medicare*/
(6642 real changes made)

. replace insurance = 3 if inscov08 == 3
(5662 real changes made)

. replace insurance = 4 if inscov08 == 2 & mcrev08 == 1 /* Medicare */
(2469 real changes made)

.
. label define insr 1 "Private Insurance" 2 "Public Insurance" 3 "Uninsured" 4 "Medicare"

. label values insurance insr

.
. gen msa = 0 if msa08 == 0
(28406 missing values generated)

. replace msa = 1 if msa08 == 1
(28155 real changes made)

.
. label define msan 0 "Not in MSA" 1 "In MSA"

. label values msa msan

.
. gen region = 1 if region08 == 1
(28080 missing values generated)

. replace region = 2 if region08 == 3
(12424 real changes made)

. replace region = 3 if region08 == 2
(6499 real changes made)

. replace region = 4 if region08 == 4
(8906 real changes made)

.
. label define regi 1 "North East" 2 "South" 3 "Midwest" 4 "West"

. label values region regi

.
. gen healthstatus = 1 if rthlth42 == 1
(23215 missing values generated)

. replace healthstatus = 2 if rthlth42 == 2
(10188 real changes made)

. replace healthstatus = 3 if rthlth42 == 3
(8628 real changes made)

. replace healthstatus = 4 if rthlth42 == 4
(2777 real changes made)

. replace healthstatus = 5 if rthlth42 == 5
(815 real changes made)
```

```
.
. label define health 1 "Excellent" 2 "Very Good" 3 "Good" 4 "Fair" 5 "Poor"

. label values healthstatus health

.
. gen fpl = 1 if povcat08 == 1
(26099 missing values generated)

. replace fpl = 2 if povcat08 == 2
(2171 real changes made)

. replace fpl = 3 if povcat08 == 3
(5667 real changes made)

. replace fpl = 4 if povcat08 == 4
(9595 real changes made)

. replace fpl = 5 if povcat08 == 5
(8666 real changes made)

.
. label define fplstat 1 "Poor" 2 "Near Poor" 3 "Low Income" 4 "Middle Income" 5 "High Income"

. label values fpl fplstat

.
. * Generate indicator variable for adult
. gen adult = 0

. replace adult = 1 if age08x > 17
(23183 real changes made)

.
. * Generate indicator variable for office visit
. gen visit = 0

. replace visit = 1 if obtotv08 > 0
(22054 real changes made)

.
. ** Number 3 **
. * Table 1 Construction
. * Linear probability model
. svy, subpop(adult): reg visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =     197.62
                                              Prob > F           =     0.0000
                                              R-squared          =     0.2046

------------------------------------------------------------------------------------------------------
                                    |             Linearized
                              visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------------+-----------------------------------------------------------------
                       healthstatus |
                         Very Good  |   .0728289   .0104661     6.96   0.000     .0521939    .0934639
                              Good  |   .1160581    .011265    10.30   0.000     .0938479    .1382683
                              Fair  |    .200874   .0124207    16.17   0.000     .1763852    .2253628
                              Poor  |   .2531628   .0153289    16.52   0.000     .2229404    .2833852
                                    |
                             agecat |
                             18-24  |  -.0263929    .013486    -1.96   0.052    -.0529819    .0001961
                             45-64  |   .0861886    .009213     9.36   0.000     .0680242    .1043531
                             65-74  |   .1611664    .011619    13.87   0.000     .1382584    .1840745
                               75+  |   .1867627   .0117369    15.91   0.000     .1636221    .2099032
                                    |
                               race |
```

```
                          Black  |  -.1179558   .0103721   -11.37   0.000    -.1384055   -.0975061
                       Hispanic  |  -.1153449   .0130775    -8.82   0.000    -.1411285   -.0895613
                          Asian  |  -.1405633   .0160152    -8.78   0.000    -.1721389   -.1089877
                          Other  |  -.0334904    .034236    -0.98   0.329    -.1009902    .0340095
                                 |
                          female |
                         Female  |   .1391388   .0056773    24.51   0.000     .1279454    .1503323
                                 |
                             fpl |
                           Poor  |   -.055755   .0133805    -4.17   0.000     -.082136   -.0293741
                      Near Poor  |  -.0515948   .0181077    -2.85   0.005    -.0872961   -.0158935
                     Low Income  |  -.0387547   .0113305    -3.42   0.001    -.0610939   -.0164154
                  Middle Income  |  -.0357165   .0098132    -3.64   0.000    -.0550642   -.0163687
                                 |
                       education |
                 No High School  |   .0020783   .0155323     0.13   0.894    -.0285451    .0327017
                      Some High  |  -.0168397   .0116099    -1.45   0.148    -.0397299    .0060506
Some College/Tech School/AA degree  |   .0492705   .0096024     5.13   0.000     .0303384    .0682027
                        College  |   .0908042   .0099263     9.15   0.000     .0712334    .1103749
                Advanced Degree  |   .1032375   .0123633     8.35   0.000      .078862     .127613
                                 |
                       insurance |
               Public Insurance  |   .0352607   .0151494     2.33   0.021      .005392    .0651294
                      Uninsured  |  -.2457508   .0142871   -17.20   0.000    -.2739193   -.2175822
                       Medicare  |   .0202693   .0107681     1.88   0.061    -.0009611    .0414998
                                 |
                             msa |
                     Not in MSA  |  -.0133504   .0088506    -1.51   0.133    -.0308002    .0040993
                                 |
                          region |
                          South  |  -.0116128   .0100957    -1.15   0.251    -.0315174    .0082918
                        Midwest  |   .0012734   .0115397     0.11   0.912    -.0214783    .0240251
                           West  |  -.0190342   .0110501    -1.72   0.086    -.0408206    .0027523
                                 |
                           _cons |   .5914321   .0151045    39.16   0.000      .561652    .6212122
----------------------------------------------------------------------------------------------------

.
. eststo lpm

.
. * Logit regression
. svy, subpop(adult): logit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running logit on estimation sample)

Survey: Logistic regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =      87.36
                                              Prob > F           =     0.0000

----------------------------------------------------------------------------------------------------
                                 |             Linearized
                           visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------------------+------------------------------------------------------------------
                    healthstatus |
                      Very Good  |   .4044001   .0585686     6.90   0.000     .2889261    .5198742
                           Good  |   .6753597   .0662026    10.20   0.000     .5448345    .8058849
                           Fair  |   1.357334   .0905371    14.99   0.000      1.17883    1.535837
                           Poor  |   2.158895   .1990341    10.85   0.000     1.766478    2.551311
                                 |
                          agecat |
                          18-24  |  -.0625399   .0684019    -0.91   0.362    -.1974014    .0723216
                          45-64  |     .48659   .0567706     8.57   0.000     .3746609    .5985191
                          65-74  |   1.227756   .1165171    10.54   0.000     .9980307    1.457482
                            75+  |   1.777876   .1636173    10.87   0.000     1.455288    2.100465
                                 |
                            race |
                          Black  |  -.6897863   .0581504   -11.86   0.000    -.8044358   -.5751368
                       Hispanic  |  -.6199646   .0715948    -8.66   0.000    -.7611211   -.4788081
                          Asian  |  -.8531756   .0846568   -10.08   0.000    -1.020085   -.6862659
```

```
                       Other |  -.2422756   .2097223    -1.16   0.249    -.6557649    .1712137
                             |
                      female |
                      Female |   .8631438   .0374362    23.06   0.000     .7893345    .9369532
                             |
                         fpl |
                        Poor |  -.3689489   .0817402    -4.51   0.000    -.5301081   -.2077897
                   Near Poor |  -.3572039   .1167581    -3.06   0.003    -.5874046   -.1270032
                  Low Income |  -.2590297   .0720621    -3.59   0.000    -.4011076   -.1169517
               Middle Income |  -.2400863   .0615636    -3.90   0.000    -.3614654   -.1187073
                             |
                   education |
              No High School |  -.0462999   .1027464    -0.45   0.653    -.2488751    .1562753
                   Some High |  -.0955991   .0674459    -1.42   0.158    -.2285757    .0373774
Some College/Tech School/AA degree |   .2893854   .0602219     4.81   0.000     .1706516    .4081192
                     College |   .5677377   .0680977     8.34   0.000      .433476    .7019995
             Advanced Degree |   .6817271    .096589     7.06   0.000      .491292    .8721623
                             |
                   insurance |
            Public Insurance |   .1882728   .0906489     2.08   0.039     .0095492    .3669964
                   Uninsured |  -1.122397   .0706839   -15.88   0.000    -1.261757   -.9830359
                    Medicare |   .1359043   .1235898     1.10   0.273    -.1077657    .3795743
                             |
                         msa |
                  Not in MSA |  -.1005684   .0593348    -1.69   0.092    -.2175531    .0164164
                             |
                      region |
                       South |  -.0840724   .0681831    -1.23   0.219    -.2185023    .0503576
                     Midwest |    -.00681   .0780753    -0.09   0.931    -.1607434    .1471235
                        West |  -.1300268   .0734128    -1.77   0.078    -.2747677    .0147142
                             |
                       _cons |   .2967943   .0865633     3.43   0.001     .1261259    .4674628
----------------------------------------------------------------------------------------------------

.
. eststo logit1

.
. * Probit regression
. svy, subpop(adult): probit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running probit on estimation sample)

Survey: Probit regression

Number of strata   =       165                  Number of obs      =      31323
Number of PSUs     =       370                  Population size    =  285026364
                                                Subpop. no. of obs =      20172
                                                Subpop. size       =  208014633
                                                Design df          =        205
                                                F(  29,    177)    =     100.41
                                                Prob > F           =     0.0000

----------------------------------------------------------------------------------------------------
                             |             Linearized
                       visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------------+----------------------------------------------------------------------
                healthstatus |
                   Very Good |   .2403186   .0346734     6.93   0.000     .1719564    .3086807
                        Good |   .3924217   .0390688    10.04   0.000     .3153935      .46945
                        Fair |   .7827303   .0526081    14.88   0.000      .679008    .8864526
                        Poor |   1.181061   .1045621    11.30   0.000     .9749059    1.387216
                             |
                      agecat |
                       18-24 |  -.0445412   .0408581    -1.09   0.277    -.1250972    .0360148
                       45-64 |   .2807088   .0328794     8.54   0.000     .2158836    .3455341
                       65-74 |     .68546   .0626452    10.94   0.000     .5619485    .8089715
                         75+ |   .9440751   .0821075    11.50   0.000     .7821918    1.105959
                             |
                        race |
                       Black |  -.4088744   .0344914   -11.85   0.000    -.4768778    -.340871
                    Hispanic |  -.3674273   .0422758    -8.69   0.000    -.4507784   -.2840762
                       Asian |  -.4925948   .0500003    -9.85   0.000    -.5911755   -.3940141
                       Other |   -.144393   .1177109    -1.23   0.221    -.3764722    .0876862
                             |
                      female |
```

```
                           Female |   .4964984   .0213784    23.22   0.000     .4543486    .5386482
                                  |
                              fpl |
                            Poor  |  -.2186107    .047191    -4.63   0.000    -.3116526   -.1255689
                       Near Poor  |  -.2083017   .0675185    -3.09   0.002    -.3414214    -.075182
                      Low Income  |  -.1526016   .0412368    -3.70   0.000    -.2339042   -.0712991
                   Middle Income  |  -.1393628   .0357891    -3.89   0.000    -.2099246   -.0688009
                                  |
                        education |
                  No High School  |  -.0281533   .0596458    -0.47   0.637    -.1457511    .0894446
                       Some High  |  -.0632671   .0396675    -1.59   0.112    -.1414757    .0149416
Some College/Tech School/AA degree  |   .1685094    .035125     4.80   0.000     .0992568    .2377621
                         College  |   .3302053   .0390494     8.46   0.000     .2532153    .4071953
                 Advanced Degree  |   .3824849   .0541269     7.07   0.000     .2757681    .4892017
                                  |
                        insurance |
                Public Insurance  |   .1191419   .0534797     2.23   0.027     .0137013    .2245826
                       Uninsured  |   -.678073   .0415783   -16.31   0.000    -.7600488   -.5960971
                        Medicare  |   .0674487   .0665375     1.01   0.312    -.0637369    .1986344
                                  |
                              msa |
                      Not in MSA  |  -.0603327   .0353654    -1.71   0.090    -.1300592    .0093938
                                  |
                           region |
                           South  |  -.0473932    .040414    -1.17   0.242    -.1270736    .0322872
                         Midwest  |  -.0035991   .0461239    -0.08   0.938    -.0945372    .0873389
                            West  |  -.0790752   .0428784    -1.84   0.067    -.1636144    .0054639
                                  |
                            _cons |   .2014991   .0506248     3.98   0.000      .101687    .3013112
----------------------------------------------------------------------------------------------------

.
. eststo probit1

.
. * Actual Table 1
. esttab lpm logit1 probit1 using "hw3table1.rtf", se replace r2 keep(1.female 2.race 3.race 4.race
5.race)
> ///
> mtitles(LPM Logit Probit) coeflabels(2.race Black 3.race Hispanic 4.race Asian ///
> 5.race Other 1.female Female) title(TABLE 1. LPM, Logit, Probit Model ///
> Coefficients)
(output written to hw3table1.rtf)

.
. * Table 2 Construction
. * Logistic Regression for Odds Ratios
. svy, subpop(adult): logistic visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education i.insurance ib1.msa ib1.region
(running logistic on estimation sample)

Survey: Logistic regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =      87.36
                                              Prob > F           =     0.0000

----------------------------------------------------------------------------------------------------
                                  |             Linearized
                            visit | Odds Ratio   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------------------------+-----------------------------------------------------------------
                     healthstatus |
                       Very Good  |   1.498403   .0877594     6.90   0.000     1.334993    1.681816
                            Good  |    1.96474   .1300708    10.20   0.000     1.724323    2.238677
                            Fair  |   3.885818   .3518107    14.99   0.000      3.25057    4.645211
                            Poor  |   8.661558   1.723945    10.85   0.000     5.850214     12.8239
                                  |
                           agecat |
                           18-24  |   .9393756   .0642551    -0.91   0.362      .820861    1.075001
                           45-64  |   1.626759   .0923521     8.57   0.000     1.454498    1.819422
                           65-74  |   3.413561   .3977382    10.54   0.000     2.712934    4.295129
                             75+  |   5.917277   .9681689    10.87   0.000     4.285717    8.169967
```

```
                                  |
                             race |
                            Black  |   .5016833   .0291731   -11.86   0.000     .4473403    .5626279
                         Hispanic  |   .5379635   .0385154    -8.66   0.000     .4671424    .6195214
                            Asian  |   .4260598   .0360689   -10.08   0.000     .3605642    .5034525
                            Other  |   .7848398   .1645984    -1.16   0.249     .5190449    1.186744
                                  |
                           female |
                           Female  |   2.370602   .0887463    23.06   0.000     2.201931    2.552193
                                  |
                              fpl |
                             Poor  |   .6914607   .0565201    -4.51   0.000     .5885413    .8123778
                        Near Poor  |   .6996298   .0816874    -3.06   0.003     .5557679    .8807308
                       Low Income  |   .7718001   .0556176    -3.59   0.000      .669578    .8896281
                    Middle Income  |     .78656   .0484235    -3.90   0.000     .6966547    .8880677
                                  |
                        education |
                   No High School  |   .9547556   .0980977    -0.45   0.653     .7796773    1.169148
                        Some High  |   .9088283   .0612967    -1.42   0.158     .7956661    1.038085
Some College/Tech School/AA degree  |   1.335606   .0804328     4.81   0.000     1.186077    1.503986
                          College  |   1.764271   .1201429     8.34   0.000      1.54261    2.017783
                  Advanced Degree  |    1.97729   .1909844     7.06   0.000     1.634427    2.392078
                                  |
                        insurance |
                 Public Insurance  |   1.207163   .1094279     2.08   0.039     1.009595    1.443393
                        Uninsured  |   .3254988   .0230075   -15.88   0.000      .283156    .3741734
                         Medicare  |   1.145572    .141581     1.10   0.273     .8978379    1.461662
                                  |
                              msa |
                       Not in MSA  |   .9043233   .0536579    -1.69   0.092     .8044849    1.016552
                                  |
                           region |
                            South  |   .9193647   .0626851    -1.23   0.219     .8037216    1.051647
                          Midwest  |   .9932132   .0775454    -0.09   0.931     .8515105    1.158497
                             West  |   .8780719   .0644617    -1.77   0.078     .7597486    1.014823
                                  |
                            _cons |   1.345539   .1164742     3.43   0.001     1.134425     1.59594
----------------------------------------------------------------------------------------------------

.
. eststo logistic1

.
. * Actual Table 2
. esttab logistic1 using "hw3table2.rtf", se replace eform keep(1.female 2.race 3.race 4.race 5.race)
///
> mtitles(Odds Ratios) coeflabels(2.race Black 3.race Hispanic 4.race Asian ///
> 5.race "Other Race" 1.female Female) title(TABLE 2. Logistic Regression Odds Ratios.)
(output written to hw3table2.rtf)

.
. * Table 3 Construction
. * Logit regression, marginal effects
. svy, subpop(adult): logit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running logit on estimation sample)

Survey: Logistic regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =      87.36
                                              Prob > F           =     0.0000

----------------------------------------------------------------------------------------------------
                                  |             Linearized
                            visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------------------+----------------------------------------------------------------
                     healthstatus |
                        Very Good  |   .4044001   .0585686     6.90   0.000     .2889261    .5198742
                             Good  |   .6753597   .0662026    10.20   0.000     .5448345    .8058849
                             Fair  |   1.357334   .0905371    14.99   0.000      1.17883    1.535837
                             Poor  |   2.158895   .1990341    10.85   0.000     1.766478    2.551311
```

```
                                     |
                              agecat |
                               18-24  |  -.0625399   .0684019    -0.91   0.362    -.1974014    .0723216
                               45-64  |     .48659   .0567706     8.57   0.000     .3746609    .5985191
                               65-74  |   1.227756   .1165171    10.54   0.000     .9980307    1.457482
                                 75+  |   1.777876   .1636173    10.87   0.000     1.455288    2.100465
                                     |
                                race |
                               Black  |  -.6897863   .0581504   -11.86   0.000    -.8044358   -.5751368
                            Hispanic  |  -.6199646   .0715948    -8.66   0.000    -.7611211   -.4788081
                               Asian  |  -.8531756   .0846568   -10.08   0.000    -1.020085   -.6862659
                               Other  |  -.2422756   .2097223    -1.16   0.249    -.6557649    .1712137
                                     |
                              female |
                              Female  |   .8631438   .0374362    23.06   0.000     .7893345    .9369532
                                     |
                                 fpl |
                                Poor  |  -.3689489   .0817402    -4.51   0.000    -.5301081   -.2077897
                           Near Poor  |  -.3572039   .1167581    -3.06   0.003    -.5874046   -.1270032
                          Low Income  |  -.2590297   .0720621    -3.59   0.000    -.4011076   -.1169517
                       Middle Income  |  -.2400863   .0615636    -3.90   0.000    -.3614654   -.1187073
                                     |
                           education |
                      No High School  |  -.0462999   .1027464    -0.45   0.653    -.2488751    .1562753
                           Some High  |  -.0955991   .0674459    -1.42   0.158    -.2285757    .0373774
 Some College/Tech School/AA degree  |   .2893854   .0602219     4.81   0.000     .1706516    .4081192
                             College  |   .5677377   .0680977     8.34   0.000      .433476    .7019995
                     Advanced Degree  |   .6817271    .096589     7.06   0.000      .491292    .8721623
                                     |
                           insurance |
                    Public Insurance  |   .1882728   .0906489     2.08   0.039     .0095492    .3669964
                           Uninsured  |  -1.122397   .0706839   -15.88   0.000    -1.261757   -.9830359
                            Medicare  |   .1359043   .1235898     1.10   0.273    -.1077657    .3795743
                                     |
                                 msa |
                          Not in MSA  |  -.1005684   .0593348    -1.69   0.092    -.2175531    .0164164
                                     |
                              region |
                               South  |  -.0840724   .0681831    -1.23   0.219    -.2185023    .0503576
                             Midwest  |    -.00681   .0780753    -0.09   0.931    -.1607434    .1471235
                                West  |  -.1300268   .0734128    -1.77   0.078    -.2747677    .0147142
                                     |
                               _cons |   .2967943   .0865633     3.43   0.001     .1261259    .4674628
----------------------------------------------------------------------------------------------------

.
. margins, dydx(race female) post

Average marginal effects                        Number of obs   =      26790
Model VCE    : Linearized

Expression   : Pr(visit), predict()
dy/dx w.r.t. : 2.race 3.race 4.race 5.race 1.female

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        race |
      Black  |  -.1206578   .0106126   -11.37   0.000     -.141458   -.0998576
   Hispanic  |  -.1075431   .0130433    -8.25   0.000    -.1331075   -.0819787
      Asian  |    -.15191   .0161397    -9.41   0.000    -.1835431   -.1202768
      Other  |  -.0398655   .0357731    -1.11   0.265    -.1099795    .0302484
             |
      female |
     Female  |   .1461475   .0061908    23.61   0.000     .1340137    .1582813
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

. est store margins1

.
. * Probit regression
. svy, subpop(adult): probit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running probit on estimation sample)
```

```
Survey: Probit regression

Number of strata   =       165                  Number of obs      =      31323
Number of PSUs     =       370                  Population size    =  285026364
                                                Subpop. no. of obs =      20172
                                                Subpop. size       =  208014633
                                                Design df          =        205
                                                F(  29,    177)    =     100.41
                                                Prob > F           =     0.0000

------------------------------------------------------------------------------------------------------
                                    |             Linearized
                              visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------------+-----------------------------------------------------------------
                       healthstatus |
                          Very Good |   .2403186   .0346734     6.93   0.000     .1719564    .3086807
                               Good |   .3924217   .0390688    10.04   0.000     .3153935      .46945
                               Fair |   .7827303   .0526081    14.88   0.000      .679008    .8864526
                               Poor |   1.181061   .1045621    11.30   0.000     .9749059    1.387216
                                    |
                             agecat |
                              18-24 |  -.0445412   .0408581    -1.09   0.277    -.1250972    .0360148
                              45-64 |   .2807088   .0328794     8.54   0.000     .2158836    .3455341
                              65-74 |     .68546   .0626452    10.94   0.000     .5619485    .8089715
                                75+ |   .9440751   .0821075    11.50   0.000     .7821918    1.105959
                                    |
                               race |
                              Black |  -.4088744   .0344914   -11.85   0.000    -.4768778    -.340871
                           Hispanic |  -.3674273   .0422758    -8.69   0.000    -.4507784   -.2840762
                              Asian |  -.4925948   .0500003    -9.85   0.000    -.5911755   -.3940141
                              Other |   -.144393   .1177109    -1.23   0.221    -.3764722    .0876862
                                    |
                             female |
                             Female |   .4964984   .0213784    23.22   0.000     .4543486    .5386482
                                    |
                                fpl |
                               Poor |  -.2186107    .047191    -4.63   0.000    -.3116526   -.1255689
                          Near Poor |  -.2083017   .0675185    -3.09   0.002    -.3414214    -.075182
                         Low Income |  -.1526016   .0412368    -3.70   0.000    -.2339042   -.0712991
                      Middle Income |  -.1393628   .0357891    -3.89   0.000    -.2099246   -.0688009
                                    |
                          education |
                     No High School |  -.0281533   .0596458    -0.47   0.637    -.1457511    .0894446
                          Some High |  -.0632671   .0396675    -1.59   0.112    -.1414757    .0149416
Some College/Tech School/AA degree  |   .1685094    .035125     4.80   0.000     .0992568    .2377621
                            College |   .3302053   .0390494     8.46   0.000     .2532153    .4071953
                    Advanced Degree |   .3824849   .0541269     7.07   0.000     .2757681    .4892017
                                    |
                          insurance |
                   Public Insurance |   .1191419   .0534797     2.23   0.027     .0137013    .2245826
                          Uninsured |   -.678073   .0415783   -16.31   0.000    -.7600488   -.5960971
                           Medicare |   .0674487   .0665375     1.01   0.312    -.0637369    .1986344
                                    |
                                msa |
                         Not in MSA |  -.0603327   .0353654    -1.71   0.090    -.1300592    .0093938
                                    |
                             region |
                              South |  -.0473932    .040414    -1.17   0.242    -.1270736    .0322872
                            Midwest |  -.0035991   .0461239    -0.08   0.938    -.0945372    .0873389
                               West |  -.0790752   .0428784    -1.84   0.067    -.1636144    .0054639
                                    |
                              _cons |   .2014991   .0506248     3.98   0.000      .101687    .3013112
------------------------------------------------------------------------------------------------------

.
. margins, dydx(race female) post

Average marginal effects                           Number of obs   =      26790
Model VCE    : Linearized

Expression   : Pr(visit), predict()
dy/dx w.r.t. : 2.race 3.race 4.race 5.race 1.female

------------------------------------------------------------------------------
             |            Delta-method
```

```
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        race |
       Black |  -.1216121   .0106677   -11.40   0.000    -.1425203   -.1007039
    Hispanic |  -.1084496   .0130572    -8.31   0.000    -.1340412    -.082858
       Asian |  -.1486131   .0161213    -9.22   0.000    -.1802103   -.1170159
       Other |  -.0406136   .0342184    -1.19   0.235    -.1076805    .0264533
             |
      female |
      Female |   .1431522   .0060689    23.59   0.000     .1312574    .1550471
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

. est store margins2

.
. * Actual Table 3
. esttab margins1 margins2 using hw3table3.rtf, se replace margin keep(1.female 2.race 3.race 4.race
5.race)
>  ///
> mtitles("Logit, Marginal Effects" "Probit, Marginal Effects") ///
> coeflabels(2.race Black 3.race Hispanic 4.race Asian ///
> 5.race "Other Race" 1.female Female) title(TABLE 3. Logistic and Probit Marginal Effects.)
(output written to hw3table3.rtf)

.
. ** Number 4 **
. svy, subpop(adult): reg visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =     197.62
                                              Prob > F           =     0.0000
                                              R-squared          =     0.2046

----------------------------------------------------------------------------------------------------
                                 |             Linearized
                           visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------------------+------------------------------------------------------------------
                    healthstatus |
                       Very Good |   .0728289   .0104661     6.96   0.000     .0521939    .0934639
                            Good |   .1160581    .011265    10.30   0.000     .0938479    .1382683
                            Fair |    .200874   .0124207    16.17   0.000     .1763852    .2253628
                            Poor |   .2531628   .0153289    16.52   0.000     .2229404    .2833852
                                 |
                          agecat |
                           18-24 |  -.0263929    .013486    -1.96   0.052    -.0529819    .0001961
                           45-64 |   .0861886    .009213     9.36   0.000     .0680242    .1043531
                           65-74 |   .1611664    .011619    13.87   0.000     .1382584    .1840745
                             75+ |   .1867627   .0117369    15.91   0.000     .1636221    .2099032
                                 |
                            race |
                           Black |  -.1179558   .0103721   -11.37   0.000    -.1384055   -.0975061
                        Hispanic |  -.1153449   .0130775    -8.82   0.000    -.1411285   -.0895613
                           Asian |  -.1405633   .0160152    -8.78   0.000    -.1721389   -.1089877
                           Other |  -.0334904    .034236    -0.98   0.329    -.1009902    .0340095
                                 |
                          female |
                          Female |   .1391388   .0056773    24.51   0.000     .1279454    .1503323
                                 |
                             fpl |
                            Poor |   -.055755   .0133805    -4.17   0.000     -.082136   -.0293741
                       Near Poor |  -.0515948   .0181077    -2.85   0.005    -.0872961   -.0158935
                      Low Income |  -.0387547   .0113305    -3.42   0.001    -.0610939   -.0164154
                   Middle Income |  -.0357165   .0098132    -3.64   0.000    -.0550642   -.0163687
                                 |
                       education |
                  No High School |   .0020783   .0155323     0.13   0.894    -.0285451    .0327017
                       Some High |  -.0168397   .0116099    -1.45   0.148    -.0397299    .0060506
```

```
Some College/Tech School/AA degree  |   .0492705   .0096024     5.13   0.000     .0303384    .0682027
                           College  |   .0908042   .0099263     9.15   0.000     .0712334    .1103749
                   Advanced Degree  |   .1032375   .0123633     8.35   0.000      .078862     .127613
                                    |
                          insurance |
                  Public Insurance  |   .0352607   .0151494     2.33   0.021      .005392    .0651294
                         Uninsured  |  -.2457508   .0142871   -17.20   0.000    -.2739193   -.2175822
                          Medicare  |   .0202693   .0107681     1.88   0.061    -.0009611    .0414998
                                    |
                                msa |
                        Not in MSA  |  -.0133504   .0088506    -1.51   0.133    -.0308002    .0040993
                                    |
                             region |
                             South  |  -.0116128   .0100957    -1.15   0.251    -.0315174    .0082918
                           Midwest  |   .0012734   .0115397     0.11   0.912    -.0214783    .0240251
                              West  |  -.0190342   .0110501    -1.72   0.086    -.0408206    .0027523
                                    |
                              _cons |   .5914321   .0151045    39.16   0.000      .561652    .6212122
------------------------------------------------------------------------------------------------------

. predict predlpm
(option xb assumed; fitted values)
(6276 missing values generated)

. label var predlpm "Predicted Values - LPM"

.
. svy, subpop(adult): logit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running logit on estimation sample)

Survey: Logistic regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =      87.36
                                              Prob > F           =     0.0000

------------------------------------------------------------------------------------------------------
                                    |             Linearized
                              visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------------+-----------------------------------------------------------------
                       healthstatus |
                         Very Good  |   .4044001   .0585686     6.90   0.000     .2889261    .5198742
                              Good  |   .6753597   .0662026    10.20   0.000     .5448345    .8058849
                              Fair  |   1.357334   .0905371    14.99   0.000      1.17883    1.535837
                              Poor  |   2.158895   .1990341    10.85   0.000     1.766478    2.551311
                                    |
                             agecat |
                             18-24  |  -.0625399   .0684019    -0.91   0.362    -.1974014    .0723216
                             45-64  |     .48659   .0567706     8.57   0.000     .3746609    .5985191
                             65-74  |   1.227756   .1165171    10.54   0.000     .9980307    1.457482
                               75+  |   1.777876   .1636173    10.87   0.000     1.455288    2.100465
                                    |
                               race |
                             Black  |  -.6897863   .0581504   -11.86   0.000    -.8044358   -.5751368
                          Hispanic  |  -.6199646   .0715948    -8.66   0.000    -.7611211   -.4788081
                             Asian  |  -.8531756   .0846568   -10.08   0.000    -1.020085   -.6862659
                             Other  |  -.2422756   .2097223    -1.16   0.249    -.6557649    .1712137
                                    |
                             female |
                            Female  |   .8631438   .0374362    23.06   0.000     .7893345    .9369532
                                    |
                                fpl |
                              Poor  |  -.3689489   .0817402    -4.51   0.000    -.5301081   -.2077897
                         Near Poor  |  -.3572039   .1167581    -3.06   0.003    -.5874046   -.1270032
                        Low Income  |  -.2590297   .0720621    -3.59   0.000    -.4011076   -.1169517
                     Middle Income  |  -.2400863   .0615636    -3.90   0.000    -.3614654   -.1187073
                                    |
                          education |
                    No High School  |  -.0462999   .1027464    -0.45   0.653    -.2488751    .1562753
                         Some High  |  -.0955991   .0674459    -1.42   0.158    -.2285757    .0373774
Some College/Tech School/AA degree  |   .2893854   .0602219     4.81   0.000     .1706516    .4081192
```

```
                     College |   .5677377   .0680977     8.34   0.000      .433476    .7019995
             Advanced Degree |   .6817271    .096589     7.06   0.000      .491292    .8721623
                             |
                   insurance |
            Public Insurance |   .1882728   .0906489     2.08   0.039     .0095492    .3669964
                   Uninsured |  -1.122397   .0706839   -15.88   0.000    -1.261757   -.9830359
                    Medicare |   .1359043   .1235898     1.10   0.273    -.1077657    .3795743
                             |
                         msa |
                  Not in MSA |  -.1005684   .0593348    -1.69   0.092    -.2175531    .0164164
                             |
                      region |
                       South |  -.0840724   .0681831    -1.23   0.219    -.2185023    .0503576
                     Midwest |    -.00681   .0780753    -0.09   0.931    -.1607434    .1471235
                        West |  -.1300268   .0734128    -1.77   0.078    -.2747677    .0147142
                             |
                       _cons |   .2967943   .0865633     3.43   0.001     .1261259    .4674628
------------------------------------------------------------------------------------------------

. predict predlogit
(option pr assumed; Pr(visit))
(6276 missing values generated)

. label var predlogit "Predicted Values - Logit"

.
. svy, subpop(adult): probit visit ib1.healthstatus ib2.agecat ib1.race i.female ///
> ib5.fpl ib3.education ib1.insurance ib1.msa ib1.region
(running probit on estimation sample)

Survey: Probit regression

Number of strata   =       165                Number of obs      =      31323
Number of PSUs     =       370                Population size    =  285026364
                                              Subpop. no. of obs =      20172
                                              Subpop. size       =  208014633
                                              Design df          =        205
                                              F(  29,    177)    =     100.41
                                              Prob > F           =     0.0000

------------------------------------------------------------------------------------------------
                                   |             Linearized
                             visit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------------------+------------------------------------------------------------
                      healthstatus |
                         Very Good |   .2403186   .0346734     6.93   0.000     .1719564    .3086807
                              Good |   .3924217   .0390688    10.04   0.000     .3153935      .46945
                              Fair |   .7827303   .0526081    14.88   0.000      .679008    .8864526
                              Poor |   1.181061   .1045621    11.30   0.000     .9749059    1.387216
                                   |
                            agecat |
                             18-24 |  -.0445412   .0408581    -1.09   0.277    -.1250972    .0360148
                             45-64 |   .2807088   .0328794     8.54   0.000     .2158836    .3455341
                             65-74 |     .68546   .0626452    10.94   0.000     .5619485    .8089715
                               75+ |   .9440751   .0821075    11.50   0.000     .7821918    1.105959
                                   |
                              race |
                             Black |  -.4088744   .0344914   -11.85   0.000    -.4768778    -.340871
                          Hispanic |  -.3674273   .0422758    -8.69   0.000    -.4507784   -.2840762
                             Asian |  -.4925948   .0500003    -9.85   0.000    -.5911755   -.3940141
                             Other |   -.144393   .1177109    -1.23   0.221    -.3764722    .0876862
                                   |
                            female |
                            Female |   .4964984   .0213784    23.22   0.000     .4543486    .5386482
                                   |
                               fpl |
                              Poor |  -.2186107    .047191    -4.63   0.000    -.3116526   -.1255689
                         Near Poor |  -.2083017   .0675185    -3.09   0.002    -.3414214    -.075182
                        Low Income |  -.1526016   .0412368    -3.70   0.000    -.2339042   -.0712991
                     Middle Income |  -.1393628   .0357891    -3.89   0.000    -.2099246   -.0688009
                                   |
                         education |
                    No High School |  -.0281533   .0596458    -0.47   0.637    -.1457511    .0894446
                         Some High |  -.0632671   .0396675    -1.59   0.112    -.1414757    .0149416
Some College/Tech School/AA degree |   .1685094    .035125     4.80   0.000     .0992568    .2377621
                           College |   .3302053   .0390494     8.46   0.000     .2532153    .4071953
```

```
                 Advanced Degree  |   .3824849   .0541269     7.07   0.000     .2757681    .4892017
                                  |
                        insurance |
                Public Insurance  |   .1191419   .0534797     2.23   0.027     .0137013    .2245826
                       Uninsured  |   -.678073   .0415783   -16.31   0.000    -.7600488   -.5960971
                        Medicare  |   .0674487   .0665375     1.01   0.312    -.0637369    .1986344
                                  |
                              msa |
                      Not in MSA  |  -.0603327   .0353654    -1.71   0.090    -.1300592    .0093938
                                  |
                           region |
                           South  |  -.0473932    .040414    -1.17   0.242    -.1270736    .0322872
                         Midwest  |  -.0035991   .0461239    -0.08   0.938    -.0945372    .0873389
                            West  |  -.0790752   .0428784    -1.84   0.067    -.1636144    .0054639
                                  |
                            _cons |   .2014991   .0506248     3.98   0.000      .101687    .3013112
----------------------------------------------------------------------------------------------------

. predict predprobit
(option pr assumed; Pr(visit))
(6276 missing values generated)

. label var predprobit "Predicted Values - Probit"

.
. tabstat visit predlpm predlogit predprobit, statistics(mean sd min max) col(stat) ///
> format(%9.3f)

    variable |      mean        sd       min       max
-------------+----------------------------------------
       visit |     0.667     0.471     0.000     1.000
     predlpm |     0.657     0.202     0.104     1.268
   predlogit |     0.652     0.220     0.097     0.997
  predprobit |     0.652     0.216     0.093     0.999
------------------------------------------------------

.
end of do-file

. log close
      name:  <unnamed>
       log:  /Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods -
Analysis/Answers/Assignment3_2016.lo
> g
  log type:  text
 closed on:  22 Sep 2016, 01:32:31
-----------------------------------------------------------------------------------------------------
------
```