

Advanced Methods in Health Services Research: Analysis - 309.716

Tuesday and Thursday 9:00-10:20

Instructors: Darrell J. Gaskin, Ph.D.

Roland J. Thorpe, Ph.D.

250 Hampton House

Computer Exercise #4: Estimating the Effect of Sex and Race/Ethnicity on having an office based physician visit using LPM, Logistic, and Probit Regression Models

Due: September 27, 2016

Answer Guide/Suggested Answers

1. Consider the following model

Hospital Readmissions = $B_0 + B_1 \text{Pct_Medicare_Patients} + B_2 \text{Teach_Hospital} + B_3 \text{Pct_Medicare_Patient} \times \text{Teach_Hospital}$

a. What is the marginal effect of percentage of Medicare patients on the number of readmissions?

The marginal effect can be taken by computing the partial derivative of Readmissions with respect to Pct_Medicare_Patient:

$$\frac{\partial \text{Readmissions}}{\partial \text{Pct_Medicare_Patient}} = \beta_1 + \beta_3 \times \text{Teach_Hospital}$$

We see here that the marginal effect of the percentage of Medicare patients is β_1 if the hospital is not a teaching hospital and $\beta_1 + \beta_3$ if the hospital is a teaching hospital.

b. What is the marginal effect of being a teaching hospitals on the number of readmissions?

The marginal effect can be taken by computing the partial derivative of Readmissions with respect to Teach_Hospital:

$$\frac{\partial \text{Readmissions}}{\partial \text{Teach_Hospital}} = \beta_2 + \beta_3 \times \text{Pct_Medicare_Patient}$$

We see here that the marginal effect of the percentage of Medicare patients is $\beta_2 + \beta_3 \times (\text{percentage of Medicare patients})$.

c. What does B3 tell us?

β_3 gives us the interaction effect between Pct_Medicare_Patient or the percentage of Medicare patients and Teach_Hospital or being a teaching hospital. That is, we can interpret β_3 as the additional effect of changes in the percentage of Medicare patients on readmissions if the hospital is a teaching hospital. Alternatively, we can also interpret β_3 as the additional effect of being a teaching hospital on readmissions at different levels of the percentage of Medicare patients.

Use the analysis file you built in computer exercises 1-3. For this analysis, limit the sample to adults, i.e., persons over the age of 17. Note you will have to create a variable that identifies the adult sample. Use the subpop command with the svy procedures to estimate the models correctly in Stata.

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

Create sex-insurance status interaction variables

Create sex-insurance categorical variables (male private insured, female private insured, male Medicare insured, female Medicare insured, male Medicaid-other public insured, female Medicaid-other public insured, male uninsured, female uninsured)

2. Estimate the following models using linear probability model (LPM) and logistic regression techniques.

Model 1: Any obvisit = f(age, race/ethnicity, sex, poverty status, education, insurance status, **sex-insurance status interaction**, health status, MSA status and region)

a. Create table 1 that displays the model 1 coefficients for LPM regressions the sex and insurance variables and interpret the coefficients for LPM.

TABLE 1. LPM Coefficients, Model 1

	(1) LPM
Public Insurance	0.00361 (0.0250)
Uninsured	-0.241*** (0.0176)
Medicare	0.0699*** (0.0147)

Female	0.147*** (0.00728)
Public Insurance-Female	0.0438 (0.0241)
Uninsured-Female	-0.0121 (0.0211)
Medicare-Female	-0.0889*** (0.0152)
<i>N</i>	31258
<i>R</i> ²	0.206

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficients on Female represents the main effect of gender. The coefficient implies that being female increases the predicted probability of having any office visit by 0.147 compared to being male, holding all other things constant. We also note that it is the effect of being female and having private insurance.

The coefficients on public insurance, uninsured, and Medicare, represent the effect of having these types of insurance on the predicted probability of having any office relative to having private insurance. We note that for men, these are the only effects of insurance status on their predicted probabilities. From above, the regression results imply that for men, being uninsured decreases the probability of having any physician office visit by -0.241 and being on Medicare increases the probability by 0.0699 compared to having private insurance, holding all other things constant. Both are statistically significant. The effect of Medicaid/public insurance was not statistically significant. To examine the effect of insurance status conditional on being female, we have to account for the gender/interaction term.

The interaction terms represent the modification of the effect of insurance status due to gender. In particular, in our case it represents the modification of the effect of insurance status on the predicted probability of having an office visit conditional on being female. The coefficient on Medicare*Female, for example, implies that females on Medicare decreases have 0.0889 lower probability of an office visit compared to men on Medicare. Thus for females, being on Medicare decreases the probability of an office visit by 0.0699 (coefficient on Medicare) – 0.0889 (coefficient on Medicare*Female) = 0.019 compared to having private insurance. Additionally, we note that women on Medicare have $0.147 + 0.0699$ (the coefficient on Medicare) - 0.0889 (the coefficient on Medicare*Female) = 0.128 higher probability of having any office visit compared to privately insured men. The other interaction variables can be interpreted similarly but are not statistically significant.

- b. Create table 2 that displays the model 1 coefficients and odds ratios for the logistic regression model.

TABLE 2. Logit Coefficients and Odds Ratios, Model 1

	(1) Coefficients	(2) OR
obtotv08n		
Public Insurance	0.119 (0.121)	1.126 (0.136)
Uninsured	-0.966*** (0.0860)	0.381*** (0.0327)
Medicare	0.238 (0.138)	1.269 (0.175)
Female	0.951*** (0.0497)	2.589*** (0.129)
Public Insurance-Female	0.0800 (0.127)	1.083 (0.137)
Uninsured-Female	-0.367*** (0.106)	0.693*** (0.0736)
Medicare-Female	-0.270 (0.168)	0.763 (0.128)
N	31258	31258

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- c. Discuss the problems associated with interpreting the coefficients on the sex and insurance status interaction variables.

The main problem is that the coefficients on the sex/insurance status interaction variables cannot be interpreted by themselves. As we saw in (a), we have to add coefficients to deduce the effects of belonging to different categories relative to the base categories. Thus it would be more difficult for consumers of the analysis to interpret the results of the regression. In addition, it would be more difficult to ascertain significant differences between groups when we add the pertinent coefficients as some coefficients are significant while others are not. The problem of interpretation is even more acute for logit models. Adding changes in log odds or odds ratios does not make for a convenient interpretation of the results.

Model 2: Any obvisit = f(age, race/ethnicity, **sex-insurance status categories**, poverty status, education, health status, MSA status and region) *Use males with private insurance as the reference group.*

- d. Create table 3 that displays the model 2 coefficients for LPM regressions for the sex and insurance status variables.

TABLE 3. LPM Coefficients, Model 2

	(1) LPM
Private Insurance - Women	0.147*** (0.00728)
Public Insurance - Men	0.00361 (0.0250)
Public Insurance - Women	0.195*** (0.0156)
Uninsured - Men	-0.241*** (0.0176)
Uninsured - Women	-0.105*** (0.0181)
Medicare - Men	0.0699*** (0.0147)
Medicare - Women	0.128*** (0.0118)
<i>N</i>	31258
<i>R</i> ²	0.206

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- e. Interpret the coefficients for LPM.

Here, the interpretations of the coefficients are more straightforward. We recall that we set the base category as men with private insurance. Thus the coefficients on the different sex/insurance status categories represent the change or the difference in the predicted probabilities of having an office visit of the different categories compared to males with private insurance. As an example, women on private insurance have 0.147 higher probability of having an office visit compared to men on private insurance. Similarly, uninsured women have a 0.105 lower probability of an office visit compared to men on private insurance.

The other coefficients can be interpreted accordingly. All coefficients except for that on men with public insurance are statistically significant.

- f. Create table 4 that displays the model 2 coefficients for the logistic regression model for the sex and insurance status variables.

TABLE 4. Logit Coefficients and Odds Ratios, Model 2

	(1) Coefficients	(2) OR
obtotv08n		
Private Insurance - Women	0.951*** (0.0497)	2.589*** (0.129)
Public Insurance - Men	0.119 (0.121)	1.126 (0.136)
Public Insurance - Women	1.150*** (0.102)	3.158*** (0.323)
Uninsured - Men	-0.966*** (0.0860)	0.381*** (0.0327)
Uninsured - Women	-0.382*** (0.0843)	0.683*** (0.0576)
Medicare - Men	0.238 (0.138)	1.269 (0.175)
Medicare - Women	0.920*** (0.160)	2.508*** (0.400)
N	31258	31258

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- g. Interpret the coefficients on the sex and insurance status variables. Do the coefficients from Model 2 present the same problems as those from Model 1? Why or why not?

We interpret the odds ratios as these are more convenient to interpret. From Table 4, we see that the odds of having any office visit for women on private insurance is 2.589 times that of the odds for men on private insurance, holding all other variables constant. We also note that the odds of having any office visit for women on public insurance is 3.158 times that of the odds for men on private insurance. The other coefficients can be interpreted similarly.

The problems we saw with model 1 are no longer present here since we can directly interpret the coefficients on the different categories as opposed to adding the different coefficients for the different variables we used in model 1.

```

-----
name: <unnamed>
log: /Users/jojo/Documents/JHU/TA Folder/Advanced HSR Methods -
Analysis/Answers/Assignment4_2016.1o
> g
log type: text
opened on: 4 Oct 2016, 15:47:42

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

. 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>

.
. /*Variable Creation and Cleaning*/
.
. * Trimming the data
. summ 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
75% 2425 Largest Std. Dev. 9786.619
90% 7453 238659
95% 13582 264510 Variance 9.58e+07
99% 40763 373799 Skewness 14.40863
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
75%             2413.5      Largest
90%              7353      98210
95%             13387      99251      Variance            5.49e+07
99%             38550      99264      Skewness            5.819768
                              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
75%             28838      Largest
90%             52000      164895
95%             70000      164895      Variance            5.81e+08
99%            103084      165391      Skewness            1.717878
                              169564      Kurtosis            6.147364

```

```

.
.
. * Generate Categories
. gen age18over = 0

. replace age18over = 1 if age08x >= 18
(23183 real changes made)

```



```

. gen agecat = 1 if age08x >=0 & age08x <=17
(23434 missing values generated)

. replace agecat = 2 if age08x >=18 & age08x <=24
(3372 real changes made)

. replace agecat = 3 if age08x >=25 & age08x <=44
(8813 real changes made)

. replace agecat = 4 if age08x >=45 & age08x <=64
(7614 real changes made)

. replace agecat = 5 if age08x >=65 & age08x <=74
(1867 real changes made)

. replace agecat = 6 if age08x >=75
(1517 real changes made)

.
. label define agecats 1 "0-18" 2 "18-24" 3 "25-44" 4 "45-64" 5 "65-74" 6 "75+"
. label values agecat agecats /*set label name sexn to the variable sex*/
. tabulate agecat

```

agecat	Freq.	Percent	Cum.
-----+-----			
0-18	9,632	29.35	29.35
18-24	3,372	10.28	39.63
25-44	8,813	26.86	66.48
45-64	7,614	23.20	89.69
65-74	1,867	5.69	95.38
75+	1,517	4.62	100.00
-----+-----			
Total	32,815	100.00	

```

.
. 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

.
. label define sexn 1 "male" 2 "female"
. label values sex sexn

. tabulate sex

```

sex	Freq.	Percent	Cum.
-----+-----			
male	15,885	48.04	48.04
female	17,181	51.96	100.00
-----+-----			
Total	33,066	100.00	

```

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

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

.
. gen racesexcat = 1 if race == 1 & female == 0
(26095 missing values generated)

. replace racesexcat = 2 if race == 1 & female == 1
(7320 real changes made)

. replace racesexcat = 3 if race == 2 & female == 0
(2894 real changes made)

. replace racesexcat = 4 if race == 2 & female == 1
(3582 real changes made)

. replace racesexcat = 5 if race == 3 & female == 0
(4612 real changes made)

. replace racesexcat = 6 if race == 3 & female == 1
(4780 real changes made)

. replace racesexcat = 7 if race == 4 & female == 0
(965 real changes made)

. replace racesexcat = 8 if race == 4 & female == 1
(1032 real changes made)

. replace racesexcat = 9 if race == 5 & female == 0
(443 real changes made)

. replace racesexcat = 10 if race == 5 & female == 1
(467 real changes made)

.
.
. gen whitemale = 0 /*okay since no missings for race and sex*/

. replace whitemale = 1 if race == 1 & female == 0
(6971 real changes made)

. gen whitefemale = 0

. replace whitefemale = 1 if race == 1 & female == 1
(7320 real changes made)

. gen blackmale = 0

. replace blackmale = 1 if race == 2 & female == 0
(2894 real changes made)

```

```

. gen blackfemale = 0

. replace blackfemale = 1 if race == 2 & female == 1
(3582 real changes made)

. gen hispanicmale = 0

. replace hispanicmale = 1 if race == 3 & female == 0
(4612 real changes made)

. gen hispanicfemale = 0

. replace hispanicfemale = 1 if race == 3 & female == 1
(4780 real changes made)

. gen asianmale = 0

. replace asianmale = 1 if race == 4 & female == 0
(965 real changes made)

. gen asianfemale = 0

. replace asianfemale = 1 if race == 4 & female == 1
(1032 real changes made)

. gen othermale = 0

. replace othermale = 1 if race == 5 & female == 0
(443 real changes made)

. gen otherfemale = 0

. replace otherfemale = 1 if race == 5 & female == 1
(467 real changes made)

.
. label define racesex 1 "White Male" 2 "White Female" 3 "Black Male" 4 "Black Female"
5 "Hispanic Male" 6 "
> Hispanic Female" 7 "Asian Male" 8 "Asian Female" 9 "Other Male" 10 "Other Female"

. label values racesexcat racesex

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

. replace education = 2 if educyr >=9 & (educyr <=12 & educyr >=0) & hideg == 1
(3732 real changes made)

. replace education = 3 if hideg == 2 | hideg == 3
(11173 real changes made)

. replace education = 4 if educyr >=13 & educyr <=17 & hideg == 3
(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)

.
. label define educ 1 "Grade 8 and below" 2 "Some High School" 3 "HighSchool" 4 "Some
College" 5 "College" 6
> "Advanced Degree"

. label values education educ

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

. replace insurance = 2 if inscov08 == 2 & mcrev08 == 2
(6642 real changes made)

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

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

.
. label define insure 1 "Private Ins" 2 "Medicaid" 3 "Uninsured" 4 "Medicare"

. label values insurance insure

.
. gen inssexcat = 1 if insurance == 1 & female == 0
(24151 missing values generated)

. replace inssexcat = 2 if insurance == 1 & female == 1
(9378 real changes made)

. replace inssexcat = 3 if insurance == 2 & female == 0
(2953 real changes made)

. replace inssexcat = 4 if insurance == 2 & female == 1
(3689 real changes made)

. replace inssexcat = 5 if insurance == 3 & female == 0

```

```

(3008 real changes made)

. replace inssexcat = 6 if insurance == 3 & female == 1
(2654 real changes made)

. replace inssexcat = 7 if insurance == 4 & female == 0
(1009 real changes made)

. replace inssexcat = 8 if insurance == 4 & female == 1
(1460 real changes made)

.
. label define insex 1 "Priv Ins Men" 2 "Priv Ins Women" 3 "Medicaid Men" 4 "Medicaid
Women" 5 "Uninsured Me
> n" 6 "Uninsured Women" 7 "Medicare Men" 8 "Medicare Women"

. label values inssexcat insex

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

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

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

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

.
. label define region2 1 "Northeast" 2 "Midwest" 3 "South" 4 "West"

. label values region region2

.
.
.
. 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

.
. gen notmsa = 1 if msa08 == 0
(28406 missing values generated)

```

```

. replace notmsa = 0 if msa08 == 1
(28155 real changes made)

.
. gen obtotv08n = 0 if obtotv08 == 0
(22054 missing values generated)

. replace obtotv08n = 1 if obtotv08 >=1
(22054 real changes made)

.
.
. /*Number 2*/
.
. *help fvvarlist
.
.
. ** Model 1
. * Linear Probability Model
. svy, subpop(ager18over): reg obtotv08n ib1.insurance##sex i.healthstatus ib3.agecat
i.race ///
> ib5.fpl ib3.education notmsa ib1.region
(running regress on estimation sample)

```

Survey: Linear regression

Number of strata	=	165	Number of obs	=	31258
Number of PSUs	=	370	Population size	=	284307457
			Subpop. no. of obs	=	20107
			Subpop. size	=	207295725
			Design df	=	205
			F(32, 174)	=	173.38
			Prob > F	=	0.0000
			R-squared	=	0.2064

		Linearized					
obtotv08n		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
insurance							
Medicaid		.0036096	.0250178	0.14	0.885	-.0457156	.0529348
Uninsured		-.2405578	.0176057	-13.66	0.000	-.2752693	-.2058464
Medicare		.0698667	.0147064	4.75	0.000	.0408715	.0988618
sex							
female		.1472871	.0072804	20.23	0.000	.132933	.1616412
insurance#sex							
Medicaid#female		.0437597	.0241271	1.81	0.071	-.0038094	.0913288
Uninsured#female		-.012076	.0211177	-0.57	0.568	-.0537118	.0295598
Medicare#female		-.0888763	.0151969	-5.85	0.000	-.1188386	-.058914
healthstatus							
Very Good		.0737853	.0105081	7.02	0.000	.0530675	.094503
Good		.1159634	.0113728	10.20	0.000	.0935408	.1383859
Fair		.2015123	.0123947	16.26	0.000	.1770749	.2259496
Poor		.2521038	.0154487	16.32	0.000	.2216451	.2825625
agecat							
18-24		-.0258016	.0134958	-1.91	0.057	-.05241	.0008068
45-64		.0869574	.0091178	9.54	0.000	.0689807	.1049342

65-74		.162363	.0116351	13.95	0.000	.1394231	.1853029
75+		.189536	.0116863	16.22	0.000	.1664952	.2125767
race							
Black		-.1178243	.0103882	-11.34	0.000	-.1383058	-.0973428
Hispanic		-.1150025	.0131188	-8.77	0.000	-.1408676	-.0891373
Asian		-.1407749	.0161656	-8.71	0.000	-.172647	-.1089028
Other		-.032871	.0346938	-0.95	0.345	-.1012734	.0355313
fpl							
Poor		-.0565723	.0134472	-4.21	0.000	-.0830849	-.0300598
Near Poor		-.0510409	.0181924	-2.81	0.006	-.0869091	-.0151727
Low Income		-.0391438	.0113795	-3.44	0.001	-.0615798	-.0167079
Middle Income		-.0371092	.0098996	-3.75	0.000	-.0566273	-.0175912
education							
Grade 8 and below		.0024143	.0155315	0.16	0.877	-.0282076	.0330363
Some High School		-.0162994	.0117257	-1.39	0.166	-.0394178	.006819
Some College		.0488912	.0096691	5.06	0.000	.0298276	.0679547
College		.0901982	.0099331	9.08	0.000	.0706141	.1097823
Advanced Degree		.103468	.0126997	8.15	0.000	.0784293	.1285068
notmsa		-.0135874	.0088065	-1.54	0.124	-.0309502	.0037754
region							
Midwest		.0003883	.0115426	0.03	0.973	-.0223691	.0231458
South		-.0122769	.0100813	-1.22	0.225	-.0321532	.0075993
West		-.0192716	.0110263	-1.75	0.082	-.041011	.0024678
_cons		.5879307	.0151007	38.93	0.000	.5581581	.6177033

```
.
. eststo lpm
```

```
.
. * Logit regression
. svy, subpop(age18over): logit obtotv08n ib1.insurance##sex i.healthstatus ib3.agecat
i.race i.sex ///
> ib5.fpl ib3.education notmsa ib1.region
(running logit on estimation sample)
```

Survey: Logistic regression

Number of strata	=	165	Number of obs	=	31258
Number of PSUs	=	370	Population size	=	284307457
			Subpop. no. of obs	=	20107
			Subpop. size	=	207295725
			Design df	=	205
			F(32, 174)	=	75.83
			Prob > F	=	0.0000

		Linearized				
	obtotv08n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
insurance						
Medicaid		.1186804	.1208557	0.98	0.327	-.1195992 .35696
Uninsured		-.9656823	.0859632	-11.23	0.000	-1.135168 -.7961969
Medicare		.2383294	.1377179	1.73	0.085	-.0331958 .5098545
sex						
female		.9512115	.0496533	19.16	0.000	.8533148 1.049108

insurance#sex							
Medicaid#female		.0799721	.1267309	0.63	0.529	-.169891	.3298352
Uninsured#female		-.3674104	.1062548	-3.46	0.001	-.5769027	-.1579181
Medicare#female		-.2699952	.1680248	-1.61	0.110	-.6012735	.061283
healthstatus							
Very Good		.4127457	.0589983	7.00	0.000	.2964245	.5290669
Good		.6793008	.0666499	10.19	0.000	.5478937	.810708
Fair		1.370507	.0898606	15.25	0.000	1.193337	1.547676
Poor		2.15732	.1987641	10.85	0.000	1.765436	2.549204
agecat							
18-24		-.0595924	.0685782	-0.87	0.386	-.1948013	.0756165
45-64		.4947306	.0562597	8.79	0.000	.3838089	.6056524
65-74		1.23893	.1173271	10.56	0.000	1.007608	1.470253
75+		1.785625	.1639472	10.89	0.000	1.462386	2.108864
race							
Black		-.6944178	.0586745	-11.84	0.000	-.8101007	-.578735
Hispanic		-.6218349	.0717762	-8.66	0.000	-.763349	-.4803207
Asian		-.8599299	.0855788	-10.05	0.000	-1.028657	-.6912025
Other		-.2511919	.2096103	-1.20	0.232	-.6644603	.1620766
fpl							
Poor		-.36992	.0816197	-4.53	0.000	-.5308418	-.2089982
Near Poor		-.3545885	.1162897	-3.05	0.003	-.5838656	-.1253113
Low Income		-.2624286	.0725383	-3.62	0.000	-.4054454	-.1194119
Middle Income		-.2496556	.0621756	-4.02	0.000	-.3722413	-.12707
education							
Grade 8 and below		-.0393499	.1030583	-0.38	0.703	-.24254	.1638402
Some High School		-.0955722	.0679421	-1.41	0.161	-.229527	.0383827
Some College		.2877786	.0604707	4.76	0.000	.1685543	.4070029
College		.5670646	.0680307	8.34	0.000	.432935	.7011942
Advanced Degree		.696993	.0995001	7.00	0.000	.5008182	.8931678
notmsa							
		-.1003801	.0593549	-1.69	0.092	-.2174044	.0166442
region							
Midwest		-.0095235	.0781824	-0.12	0.903	-.1636683	.1446212
South		-.0830117	.0682538	-1.22	0.225	-.2175811	.0515578
West		-.12984	.0734891	-1.77	0.079	-.2747315	.0150514
_cons							
		.2607229	.0860613	3.03	0.003	.0910441	.4304017

```

.
. eststo logit1

```

```

.
. svy, subpop(age18over): logit obtotv08n ib1.insurance##sex i.healthstatus ib3.agecat
i.race i.sex ///
> ib5.fpl ib3.education notmsa ib1.region, or
(running logit on estimation sample)

```

Survey: Logistic regression

Number of strata	=	165	Number of obs	=	31258
Number of PSUs	=	370	Population size	=	284307457
			Subpop. no. of obs	=	20107
			Subpop. size	=	207295725
			Design df	=	205

F(32, 174) = 75.83
 Prob > F = 0.0000

	obtotv08n	Odds Ratio	Linearized Std. Err.	t	P> t	[95% Conf. Interval]	
insurance							
Medicaid		1.12601	.1360848	0.98	0.327	.887276	1.428979
Uninsured		.3807233	.0327282	-11.23	0.000	.3213682	.4510411
Medicare		1.269127	.1747816	1.73	0.085	.9673492	1.665049
sex							
female		2.588844	.1285448	19.16	0.000	2.347415	2.855104
insurance#sex							
Medicaid#female		1.083257	.1372821	0.63	0.529	.8437568	1.390739
Uninsured#female		.6925254	.0735841	-3.46	0.001	.5616352	.8539198
Medicare#female		.7633831	.1282673	-1.61	0.110	.5481132	1.0632
healthstatus							
Very Good		1.510961	.089144	7.00	0.000	1.345041	1.697348
Good		1.972498	.1314668	10.19	0.000	1.729606	2.2495
Fair		3.937345	.3538124	15.25	0.000	3.298069	4.700535
Poor		8.64793	1.718898	10.85	0.000	5.844119	12.79691
agecat							
18-24		.9421485	.0646108	-0.87	0.386	.8229981	1.078549
45-64		1.640056	.092269	8.79	0.000	1.467865	1.832447
65-74		3.451919	.4050037	10.56	0.000	2.739041	4.350335
75+		5.963304	.9776669	10.89	0.000	4.316246	8.238873
race							
Black		.4993651	.0293	-11.84	0.000	.4448133	.5606071
Hispanic		.5369583	.0385408	-8.66	0.000	.4661028	.618585
Asian		.4231917	.0362162	-10.05	0.000	.3574866	.5009733
Other		.7778731	.1630502	-1.20	0.232	.5145512	1.17595
fpl							
Poor		.6907896	.0563821	-4.53	0.000	.5881097	.8113967
Near Poor		.7014621	.0815728	-3.05	0.003	.5577382	.8822222
Low Income		.7691813	.0557951	-3.62	0.000	.6666798	.8874422
Middle Income		.779069	.0484391	-4.02	0.000	.689188	.880672
education							
Grade 8 and below		.9614142	.0990817	-0.38	0.703	.7846324	1.178026
Some High School		.9088528	.0617493	-1.41	0.161	.7949095	1.039129
Some College		1.333462	.0806354	4.76	0.000	1.183592	1.502308
College		1.763084	.1199439	8.34	0.000	1.541776	2.016159
Advanced Degree		2.007706	.1997671	7.00	0.000	1.650071	2.442856
notmsa		.9044935	.0536861	-1.69	0.092	.8046045	1.016783
region							
Midwest		.9905217	.0774414	-0.12	0.903	.8490236	1.155602
South		.9203404	.0628167	-1.22	0.225	.8044624	1.05291
West		.8782359	.0645408	-1.77	0.079	.7597761	1.015165
_cons		1.297868	.1116962	3.03	0.003	1.095317	1.537875

.
 . eststo logit2

```
.
. * Table 1
. esttab lpm using hw4table1.rtf , se r2 replace keep(2.sex 2.insurance 3.insurance
4.insurance 2.insurance#
> 2.sex ///
> 3.insurance#2.sex 4.insurance#2.sex) mtitles(LPM) coeclabels(2.sex Female
2.insurance "Public Insurance" /
> ///
> 3.insurance Uninsured 4.insurance Medicare 2.insurance#2.sex "Public Insurance-
Female" ///
> 3.insurance#2.sex "Uninsured-Female" 4.insurance#2.sex "Medicare-Female") ///
> title(TABLE 1. LPM Coefficients, Model 1)
(output written to hw4table1.rtf)
```

```
.
. * Table 2
. esttab logit1 logit2 using hw4table2.rtf , se r2 replace keep(2.sex 2.insurance
3.insurance 4.insurance 2.
> insurance#2.sex ///
> 3.insurance#2.sex 4.insurance#2.sex) mtitles(Coefficients OR) coeclabels(2.sex
Female 2.insurance "Public
> Insurance" ///
> 3.insurance Uninsured 4.insurance Medicare 2.insurance#2.sex "Public Insurance-
Female" ///
> 3.insurance#2.sex "Uninsured-Female" 4.insurance#2.sex "Medicare-Female") ///
> title(TABLE 2. Logit Coefficients and Odds Ratios, Model 1) eform(0 1)
(output written to hw4table2.rtf)
```

```
.
. ** Model 2
. * Linear Probability Model
. svy, subpop(age18over): reg obtotv08n ib1.inssexcat ib1.healthstatus ib3.agecat
i.race ///
> ib5.fpl ib3.education notmsa ib1.region
(running regress on estimation sample)
```

Survey: Linear regression

Number of strata	=	165	Number of obs	=	31258
Number of PSUs	=	370	Population size	=	284307457
			Subpop. no. of obs	=	20107
			Subpop. size	=	207295725
			Design df	=	205
			F(32, 174)	=	173.38
			Prob > F	=	0.0000
			R-squared	=	0.2064

		Linearized				
	obtoto8n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

inssexcat						
Priv Ins Women		.1472871	.0072804	20.23	0.000	.132933 .1616412
Medicaid Men		.0036096	.0250178	0.14	0.885	-.0457156 .0529348
Medicaid Women		.1946564	.0155975	12.48	0.000	.1639043 .2254085
Uninsured Men		-.2405578	.0176057	-13.66	0.000	-.2752693 -.2058464
Uninsured Women		-.1053468	.0180801	-5.83	0.000	-.1409936 -.0696999
Medicare Men		.0698667	.0147064	4.75	0.000	.0408715 .0988618
Medicare Women		.1282775	.011826	10.85	0.000	.1049613 .1515937
healthstatus						
Very Good		.0737853	.0105081	7.02	0.000	.0530675 .094503

Good		.1159634	.0113728	10.20	0.000	.0935408	.1383859
Fair		.2015123	.0123947	16.26	0.000	.1770749	.2259496
Poor		.2521038	.0154487	16.32	0.000	.2216451	.2825625
agecat							
18-24		-.0258016	.0134958	-1.91	0.057	-.05241	.0008068
45-64		.0869574	.0091178	9.54	0.000	.0689807	.1049342
65-74		.162363	.0116351	13.95	0.000	.1394231	.1853029
75+		.189536	.0116863	16.22	0.000	.1664952	.2125767
race							
Black		-.1178243	.0103882	-11.34	0.000	-.1383058	-.0973428
Hispanic		-.1150025	.0131188	-8.77	0.000	-.1408676	-.0891373
Asian		-.1407749	.0161656	-8.71	0.000	-.172647	-.1089028
Other		-.032871	.0346938	-0.95	0.345	-.1012734	.0355313
fpl							
Poor		-.0565723	.0134472	-4.21	0.000	-.0830849	-.0300598
Near Poor		-.0510409	.0181924	-2.81	0.006	-.0869091	-.0151727
Low Income		-.0391438	.0113795	-3.44	0.001	-.0615798	-.0167079
Middle Income		-.0371092	.0098996	-3.75	0.000	-.0566273	-.0175912
education							
Grade 8 and below		.0024143	.0155315	0.16	0.877	-.0282076	.0330363
Some High School		-.0162994	.0117257	-1.39	0.166	-.0394178	.006819
Some College		.0488912	.0096691	5.06	0.000	.0298276	.0679547
College		.0901982	.0099331	9.08	0.000	.0706141	.1097823
Advanced Degree		.103468	.0126997	8.15	0.000	.0784293	.1285068
notmsa		-.0135874	.0088065	-1.54	0.124	-.0309502	.0037754
region							
Midwest		.0003883	.0115426	0.03	0.973	-.0223691	.0231458
South		-.0122769	.0100813	-1.22	0.225	-.0321532	.0075993
West		-.0192716	.0110263	-1.75	0.082	-.041011	.0024678
_cons		.5879307	.0151007	38.93	0.000	.5581581	.6177033

```

.
. eststo lpm2

.
. * Table 3
. esttab lpm2 using hw4table3.rtf, se r2 replace keep(2.inssexcat 3.inssexcat ///
> 4.inssexcat 5.inssexcat 6.inssexcat 7.inssexcat 8.inssexcat) ///
> mtitles(LPM) coeflabels(2.inssexcat "Private Insurance - Women" 3.inssexcat "Public
Insurance - Men" ///
> 4.inssexcat "Public Insurance - Women" 5.inssexcat "Uninsured - Men" 6.inssexcat
"Uninsured - Women" ///
> 7.inssexcat "Medicare - Men" 8.inssexcat "Medicare - Women") ///
> title(TABLE 3. LPM Coefficients, Model 2)
(output written to hw4table3.rtf)

.
. * Logit Model
. svy, subpop(age18over): logit obtotv08n ib1.inssexcat i.healthstatus ib3.agecat
i.race ///
> ib5.fpl ib3.education notmsa ib1.region
(running logit on estimation sample)

```

Survey: Logistic regression

Number of obs	=	31258
Population size	=	284307457
Subpop. no. of obs	=	20107
Subpop. size	=	207295725
Design df	=	205
F(32, 174)	=	75.83
Prob > F	=	0.0000

		Linearized				
obtotv08n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
inssexcat						
Priv Ins Women	.9512115	.0496533	19.16	0.000	.8533148	1.049108
Medicaid Men	.1186804	.1208557	0.98	0.327	-.1195992	.35696
Medicaid Women	1.149864	.1021498	11.26	0.000	.9484652	1.351263
Uninsured Men	-.9656823	.0859632	-11.23	0.000	-1.135168	-.7961969
Uninsured Women	-.3818811	.0843203	-4.53	0.000	-.5481273	-.215635
Medicare Men	.2383294	.1377179	1.73	0.085	-.0331958	.5098545
Medicare Women	.9195457	.1595805	5.76	0.000	.6049163	1.234175
healthstatus						
Very Good	.4127457	.0589983	7.00	0.000	.2964245	.5290669
Good	.6793008	.0666499	10.19	0.000	.5478937	.810708
Fair	1.370507	.0898606	15.25	0.000	1.193337	1.547676
Poor	2.15732	.1987641	10.85	0.000	1.765436	2.549204
agecat						
18-24	-.0595924	.0685782	-0.87	0.386	-.1948013	.0756165
45-64	.4947306	.0562597	8.79	0.000	.3838089	.6056524
65-74	1.23893	.1173271	10.56	0.000	1.007608	1.470253
75+	1.785625	.1639472	10.89	0.000	1.462386	2.108864
race						
Black	-.6944178	.0586745	-11.84	0.000	-.8101007	-.578735
Hispanic	-.6218349	.0717762	-8.66	0.000	-.763349	-.4803207
Asian	-.8599299	.0855788	-10.05	0.000	-1.028657	-.6912025
Other	-.2511919	.2096103	-1.20	0.232	-.6644603	.1620766
fpl						
Poor	-.36992	.0816197	-4.53	0.000	-.5308418	-.2089982
Near Poor	-.3545885	.1162897	-3.05	0.003	-.5838656	-.1253113
Low Income	-.2624286	.0725383	-3.62	0.000	-.4054454	-.1194119
Middle Income	-.2496556	.0621756	-4.02	0.000	-.3722413	-.12707
education						
Grade 8 and below	-.0393499	.1030583	-0.38	0.703	-.24254	.1638402
Some High School	-.0955722	.0679421	-1.41	0.161	-.229527	.0383827
Some College	.2877786	.0604707	4.76	0.000	.1685543	.4070029
College	.5670646	.0680307	8.34	0.000	.432935	.7011942
Advanced Degree	.696993	.0995001	7.00	0.000	.5008182	.8931678
notmsa	-.1003801	.0593549	-1.69	0.092	-.2174044	.0166442
region						
Midwest	-.0095235	.0781824	-0.12	0.903	-.1636683	.1446212
South	-.0830117	.0682538	-1.22	0.225	-.2175811	.0515578
West	-.12984	.0734891	-1.77	0.079	-.2747315	.0150514
_cons	.2607229	.0860613	3.03	0.003	.0910441	.4304017

```
. eststo logit3
```

```
.
. * Table 4
. esttab logit3 logit3 using hw4table4.rtf, se r2 replace keep(2.inssexcat 3.inssexcat
///
> 4.inssexcat 5.inssexcat 6.inssexcat 7.inssexcat 8.inssexcat) ///
> mtitles(Coefficients OR) coeclabels(2.inssexcat "Private Insurance - Women"
3.inssexcat "Public Insurance
> - Men" ///
> 4.inssexcat "Public Insurance - Women" 5.inssexcat "Uninsured - Men" 6.inssexcat
"Uninsured - Women" ///
> 7.inssexcat "Medicare - Men" 8.inssexcat "Medicare - Women") ///
> title(TABLE 4. Logit Coefficients and Odds Ratios, Model 2) eform(0 1)
(output written to hw4table4.rtf)
```

```
.
. /*2g*/
. * Quietly
. svy, subpop(ager18over): logit obtotv08n ib1.race ib3.agecat ib5.fpl notmsa ///
> ib1.inssexcat ib3.education ib1.region ib1.healthstatus
(running logit on estimation sample)
```

Survey: Logistic regression

Number of strata	=	165	Number of obs	=	31258
Number of PSUs	=	370	Population size	=	284307457
			Subpop. no. of obs	=	20107
			Subpop. size	=	207295725
			Design df	=	205
			F(32, 174)	=	75.83
			Prob > F	=	0.0000

		Linearized				
	obtotv08n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
race						
Black		-.6944178	.0586745	-11.84	0.000	-.8101007 -.578735
Hispanic		-.6218349	.0717762	-8.66	0.000	-.763349 -.4803207
Asian		-.8599299	.0855788	-10.05	0.000	-1.028657 -.6912025
Other		-.2511919	.2096103	-1.20	0.232	-.6644603 .1620766
agecat						
18-24		-.0595924	.0685782	-0.87	0.386	-.1948013 .0756165
45-64		.4947306	.0562597	8.79	0.000	.3838089 .6056524
65-74		1.23893	.1173271	10.56	0.000	1.007608 1.470253
75+		1.785625	.1639472	10.89	0.000	1.462386 2.108864
fpl						
Poor		-.36992	.0816197	-4.53	0.000	-.5308418 -.2089982
Near Poor		-.3545885	.1162897	-3.05	0.003	-.5838656 -.1253113
Low Income		-.2624286	.0725383	-3.62	0.000	-.4054454 -.1194119
Middle Income		-.2496556	.0621756	-4.02	0.000	-.3722413 -.12707
notmsa						
		-.1003801	.0593549	-1.69	0.092	-.2174044 .0166442
inssexcat						
Priv Ins Women		.9512115	.0496533	19.16	0.000	.8533148 1.049108
Medicaid Men		.1186804	.1208557	0.98	0.327	-.1195992 .35696
Medicaid Women		1.149864	.1021498	11.26	0.000	.9484652 1.351263
Uninsured Men		-.9656823	.0859632	-11.23	0.000	-1.135168 -.7961969
Uninsured Women		-.3818811	.0843203	-4.53	0.000	-.5481273 -.215635

Medicare Men		.2383294	.1377179	1.73	0.085	-.0331958	.5098545
Medicare Women		.9195457	.1595805	5.76	0.000	.6049163	1.234175
education							
Grade 8 and below		-.0393499	.1030583	-0.38	0.703	-.24254	.1638402
Some High School		-.0955722	.0679421	-1.41	0.161	-.229527	.0383827
Some College		.2877786	.0604707	4.76	0.000	.1685543	.4070029
College		.5670646	.0680307	8.34	0.000	.432935	.7011942
Advanced Degree		.696993	.0995001	7.00	0.000	.5008182	.8931678
region							
Midwest		-.0095235	.0781824	-0.12	0.903	-.1636683	.1446212
South		-.0830117	.0682538	-1.22	0.225	-.2175811	.0515578
West		-.12984	.0734891	-1.77	0.079	-.2747315	.0150514
healthstatus							
Very Good		.4127457	.0589983	7.00	0.000	.2964245	.5290669
Good		.6793008	.0666499	10.19	0.000	.5478937	.810708
Fair		1.370507	.0898606	15.25	0.000	1.193337	1.547676
Poor		2.15732	.1987641	10.85	0.000	1.765436	2.549204
_cons		.2607229	.0860613	3.03	0.003	.0910441	.4304017

```
. margins, dydx(inssexcat) noestimcheck
```

```
Average marginal effects          Number of obs   =       26721
Model VCE      : Linearized
```

```
Expression      : Pr(obtotv08n), predict()
dy/dx w.r.t.    : 2.inssexcat 3.inssexcat 4.inssexcat 5.inssexcat 6.inssexcat 7.inssexcat
8.inssexcat
```

		Delta-method				
		dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]
inssexcat						
Priv Ins Women		.1594575	.0079347	20.10	0.000	.1439058 .1750092
Medicaid Men		.0227213	.0228466	0.99	0.320	-.0220572 .0674998
Medicaid Women		.1852411	.0135743	13.65	0.000	.1586359 .2118464
Uninsured Men		-.1981353	.0178939	-11.07	0.000	-.2332067 -.1630638
Uninsured Women		-.076672	.0173164	-4.43	0.000	-.1106116 -.0427325
Medicare Men		.0449397	.0252596	1.78	0.075	-.0045683 .0944476
Medicare Women		.1550918	.0228278	6.79	0.000	.1103502 .1998334

Note: dy/dx for factor levels is the discrete change from the base level.

```
.
. est store margins

.
. * Table 5
. esttab margins using hw4table5.rtf, se r2 replace keep(2.inssexcat 3.inssexcat ///
> 4.inssexcat 5.inssexcat 6.inssexcat 7.inssexcat 8.inssexcat) ///
> mtitles(Marginal Effects) coeflabels(2.inssexcat "Private Insurance - Women"
3.inssexcat "Public Insurance
> - Men" ///
> 4.inssexcat "Public Insurance - Women" 5.inssexcat "Uninsured - Men" 6.inssexcat
"Uninsured - Women" ///
> 7.inssexcat "Medicare - Men" 8.inssexcat "Medicare - Women") ///
> title(TABLE 5. Logit Marginal Effects, Model 2)
(output written to hw4table5.rtf)
```

```
.
.
end of do-file

. log close
  name: <unnamed>
  log:    /Users/jojo/Documents/JHU/TA    Folder/Advanced    HSR    Methods    -
Analysis/Answers/Assignment4_2016.lo
> g
  log type: text
  closed on:  4 Oct 2016, 15:48:43
-----
-----
```