Homework 2

Jacob Puthipiroj

April 2, 2020

1 A Theory of Extramarital Affairs

(a) **The regressors of interest are $v1$ to $v8$; however, not necessarily all of them belong in your model. Use these data to build a binary choice model for A. Report all computed results for the model. Compute the marginal effects for the variables you choose. Compare the results you obtain for a probit model to those for a logit model. Are there any substantial differences in the results for the two models?**

The specification of the probit model is

$$p = \Phi(\beta_0 + \mathbf{B}x),\ \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

The specification of the logit model is

$$p = \frac{1}{1+e^{-l}} = \frac{e^l}{1+e^l},\ l = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \mathbf{B}x$$

In both cases, β_0 is the intercept, and we define $\mathbf{Bx}$ as

$$\mathbf{Bx} = \beta_1 \text{Rating} + \beta_2 \text{Age} + \beta_3 \text{Years} + \beta_4 \text{religiosity} + \delta_1 \text{professional} + \delta_2 \text{managerial}$$

Thus the variables of rating, age, years, and religiosity are implicitly thought to be continuous variables, and there are indicator variables for if the wife is in a managerial/administrative/business role, or a professionalism with an advanced degree. These variables were chosen for the model via a process of backward elimination: first all the predictors were included, then the predictor with the highest p-value over $\alpha = 0.05$ was successively removed until the model only contained significant predictors. The probit model had an AIC and BIC of 6944 and 6992 respectively, while the logit model had an AIC and BIC of 6995 respectively, which were among the lowest found.

The negative signs for β_{rating} and $\beta_{\text{religiosity}}$ are unsurprising: The better one rates [satisfaction of] a marriage, the less reason there is to be involved in an extramarital affair. Similarly, the more religious one is, the less one is less likely to cheat, perhaps due to religious beliefs concerning the (im)morality of infidelity. The negative sign for β_{age} implies that older women cheat less, when controlled for the years of the marriage (both are highly correlated with each other at $r = 0.8941$). β_{years} is positive, perhaps because as a marriage drags on, one is more inclined to look elsewhere for emotional or physical fulfillment, such as an extramarital affair. A final interesting note is that females in managerial and professional careers are more likely to have cheated if they were a student.

There are no substantial differences in the results of the two models. The same variables were found to be significant, and the corresponding p-values for said variables in the two tables are nearly identical. While the coefficients are different in the two models, they are used differently and cannot be directly compared. Instead, a comparison between the average marginal effects for the significant predictors in both models shows very similar results. The coefficients are different, because they are parameters of the latent model, which in the case of the probit model is CDF of the standard normal distribution, and in the case of the logit model is the CDF of the logistic function. A further theoretical difference is that the probit model conforms more closely to the assumption of normality, while coefficients of the logit model can be interpreted in terms of odd ratios. Practically, as observed, there is little difference between the two, only that the logistic distribution is less computationally expensive.

The marginal effects of each variable vary at different points, and as such we compute instead the average marginal effects (AME) by calculating the marginal effect for each individual with their observed levels of covariates, which are then averaged across individuals. The AMEs for the variables in the probit model is available here in the appendix. The marginal effects of the logit model were similar, available here in the appendix. There were no substantial differences in the results between the two models. On average, each additional increase in the reported marriage score decreased the probability of cheating by around 13%, each additional year of the female decreased the probability by 1.1%, each additional year of marriage increased the probability by 2%, each additional score of religiosity decreased the probability by 6.8%, being in a managerial, administrative or business occupation increased the probability of extramarital affairs by 9%, and being in a professional career with an advanced degree increased the probability by 10%, both compared to the baseline of being a student.[1]

(b) **Continuing the analysis from part a), we now consider the self-reported rating, $v1$. This is a natural candidate for an ordered choice model, because the simple five-item coding is a censored version of what would be a continuous scale on some subjective satisfaction variable. Analyze this variable using an ordered probit model. What variables appear to explain the response to this survey question? Can you obtain the marginal effects for your model? Report them as well. What do they suggest about the impact of the different independent variables on the reported ratings?**

In this section, I employed an ordered probit regression using all variables, available here in the appendix. As seen, the only significant predictors (at the $\alpha = 0.05$ level) were children, religiosity, and the husband's occupation being in farming, agriculture, semi-skilled, or unskilled labor. β_{children} was negative, suggesting that more children made marriages unhappier, while $\beta_{\text{religiosity}} > 0$ suggests that more religiosity made marriages happier. $\delta_{2,hoc} < 0$, suggesting that wives were less satisfied with their marriage when their husbands were involved in menial labor, compared to when they (the husbands) were students.

The average marginal effects are available here. An interesting finding suggested that having children were apparently bad for marriage satisfaction. More specifically, each additional child decreased the probability of the wife rating their marriage as 5 (the highest satisfaction level) by 2.4%, and increased the probabilities

[1] #specification: I use backward elimination to obtain a model, justifying my model with its low AIC and BIC, as well as economic reasoning. I found no significant differences between probit and logit in terms of their results. I explain that there were many similarities and few differences. The only notable difference is the difference in coefficient estimates, which cannot be compared directly. The marginal effects shows little differences between the two models.

of rating their marriage as a 4, 3, 2 or 1 by 0.5%, 1%, 0.5%, and 0.2% respectively. Religiosity had the opposite effect: according to the model, each additional unit of 'religiosity' reported increased the probability of marriages being rated a 5 by 5%, and decreased the probability of rating their marriage as a 4, 3, 2 or 1 by 1%, 2%, 1%, and 0.5% respectively.

This suggests that different independent variables can have different impacts on the reported ratings. There is a difference in sign: additional religiosity tends to increase marriage satisfaction, while additional children tends to decrease it. Furthermore, there is a difference in magnitude: a one-unit increase in religiosity, for example, seems to increase the probability of a high marriage satisfaction almost twice as much as an additional child lowers it.

2 Incentive Effects in the Demand for Health Care

A note about this dataset: there were several mistakes in some of the data. The indicator variable handdum for example, was miscoded in the year 1987. Ones and zeros were swapped, and if not corrected, would imply that 88% of the participants in 1987 were handicapped, when in reality it was $100 - 88 = 12\%$. The full list of coding errors are given here.[1]

(a) **Begin by fitting a Poisson model to this variable. The exogenous variables are listed in Table F7.1. Determine an appropriate specification for the right-hand side of your model. Report the regression results and the marginal effects.**

The specification we have is

$$\mathbf{Bx} = \sum_{i=1}^{4} \beta_i v_i + \sum_{i=1}^{3} \delta_i u_i$$

Where v_i indicate each of the four continuous variables of age, hsat, educ and docvis, and u_i indicate each of the dummy variables handdum, addon, and bluec. The Poisson regression results are available here in the appendix. All variables, save for the constant, were found to be significant. Specifically, the coefficients for age, hsat, educ and bluec were negative, while the coefficients for docvis, handdum, and addon were positive.

The positive signs for $\beta_{\text{docvis}}, \delta_{\text{handdum}}$ and δ_{addon} are unsurprising. Individuals who have frequent doctor visits in the past 3 months are likely to also frequently visit hospitals in the last calendar year. Handicapped individuals may need to visit hospitals more, as a result of a chronic treatment for said handicap, or a new injury causing patients to become handicapped and require treatment. Finally, those who purchase add-on insurance may expect themselves to be more at risk of requiring healthcare.

The negative signs for β_{hsat} and β_{educ} also make sense. Individuals who are more satisfied with their health are likely healthier, and require fewer hospital visits. More educated individuals may also make wiser health-related decisions, such as in career and lifestyle choices, and visit hospitals less often. The negative sign for δ_{bluec}, however, is surprising, as workplace injuries (which would require hospital visits) should be more commonplace among blue collar workers. Several explanations are possible: Perhaps only physically healthier people would consider blue collar jobs, or, blue collar workers choose to visit hospitals only for the most serious injuries, either out of necessity or because of being desensitized to workplace injuries. $\beta_{\text{age}} < 0$ is similarly surprising. Despite (or because of) having weaker bodies, older people may make lifestyle and professional choices that would limit physical injuries.

The marginal effects are given in the appendix here. Again, the marginal effects of all explanatory variables were significant. On average, being handicapped and being insured by addon insurance increased the expected number of hospital visits by 4% and 5% respectively, and each additional doctor visit in the past 3 months increased expected visits by 0.4%. Each additional year in age and in schooling decreased expected visits by 0.09% and 0.6% respectively. Each additional unit increase of perceived health satisfaction on the 1-10 scale decreased expected visits by 2.5%. Finally, being in a blue-collar job decreased expected visits by 1.4%.

(b) **Estimate the model using ordinary least squares and compare your least squares results to the marginal effects computed in part a). What do you find?**

The output is given here in the appendix. Unlike in the Poisson regression, the variables addon and bluec were found to be insignificant. In the OLS model with no interaction terms, the coefficient estimates are equivalent to the marginal effects, and can be compared with the AME of the Poisson regression. In modeling count data, such as in this example, Poisson regression is preferred for a number of reasons: data is intrinsically integer-valued, which the Poisson model takes into account. OLS instead assumes that the true values are normally distributed around the expected value, and can take any real value, even negative or fractional ones. Thus the specification of OLS is less appropriate.

Like in the Poisson model, being handicapped had large marginal effects on increasing the expected number of visits, at $\hat{\delta}_{\text{handdum}} = 5.8\%$, comparable to the 4% from the Poisson model. However, the coefficients for being on addon insurance, and in a blue-collar occupation were not significant. Each additional doctor visit in the past 3 months increased expected visits by 1.7%, comparable to the 4% from the Poisson model. The marginal effects for each additional year in age and schooling was -0.15% and -0.5% respectively, compared to -0.09% and 0.6% from the Poisson model. Each additional unit increase of perceived health satisfaction on the 1-10 scale decreased expected visits by 2.4%, similar to 2.5% from the Poisson model.

Overall, the marginal effect estimates the were identical in sign and very similar in magnitude, though δ_{addon} and δ_{bluec} were not significant like they were in the Poisson model.[2]

(c) **Is there evidence of overdispersion in the data? Test for overdispersion.**

Overdispersion is where variance is greater than would be expected in a Poisson regression. One test for this is to simply run the Poisson regression, and then test using the `poisgof` command. The output is

```
Deviance goodness-of-fit =  20004.18
Prob > chi2(27322)       =    1.0000

Pearson goodness-of-fit  =    131234
Prob > chi2(27322)       =    0.0000
```

Which suggests the Poisson model to be inappropriate. We can furthermore check for overdispersion by using the `nbreg` command, given here in the appendix, which fits the data with a negative binomial distribution, and gives a likelihood-ratio test, with the hypothesis being that the negative binomial distribution is equivalent to a poisson distribution, under $\alpha = 0$. Since the p-value is significant, α is significantly different from 0, the negative binomial distribution would be significantly difference from the poisson model, and so the latter again inappropriate.

A further note should be made here that overdispersion is a specific concern relating to the excess variation of an otherwise properly specified model. If there are omitted variables, or significant correlation among the predictor variables, using the negative binomial regression would still result in a misspecified model, and would still be inappropriate. In this particular case, more research and domain knowledge is necessary to determine if the negative binomial model would be sufficient.[3]

[2] #specification: I explain why the poisson model is more appropriate for count data that OLS.
[3] #modeltesting: I explain how the using the Poisson model implicitly assumes that the mean is equal to the variance, which

3 Appendix

Probit Model Output

```
Probit regression                                 Number of obs     =      6,366
                                                  LR chi2(6)        =    1074.09
                                                  Prob > chi2       =     0.0000
Log likelihood = -3465.4864                       Pseudo R2         =     0.1342
```

A	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
rating	-.4287915	.0182404	-23.51	0.000	-.4645421	-.3930409
age	-.0371386	.0058164	-6.39	0.000	-.0485385	-.0257388
years	.0669787	.0054895	12.20	0.000	.0562195	.0777379
religiosity	-.2229334	.0203962	-10.93	0.000	-.2629093	-.1829575
1.managerial	.2820679	.0527357	5.35	0.000	.1787077	.385428
1.professional	.3201245	.131654	2.43	0.015	.0620874	.5781616
_cons	2.218905	.1535921	14.45	0.000	1.91787	2.51994

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	6,366	-4002.53	-3465.486	7	6944.973	6992.284

Note: BIC uses N = number of observations. See [R] BIC note.

Probit Model AMEs

```
Average marginal effects                          Number of obs     =      6,366
Model VCE    : OIM

Expression   : Pr(A), predict()
dy/dx w.r.t. : rating age years religiosity 1.managerial 1.professional
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
rating	-.1318603	.0048907	-26.96	0.000	-.1414459	-.1222747
age	-.0114207	.0017739	-6.44	0.000	-.0148976	-.0079439
years	.020597	.0016348	12.60	0.000	.0173929	.0238012
religiosity	-.0685556	.0061151	-11.21	0.000	-.080541	-.0565703

may not be the case. I test the validity of the model using both the poigof command which computes a pearson goodness-of-fit chi-squared statistic, which was significant, and a likelihood ratio test, which was also significant. This suggests the poisson model to be inappropriate.

```
  1.managerial |   .0905521     .01749     5.18   0.000     .0562724    .1248318
1.professional |   .1038312   .0444333     2.34   0.019     .0167435    .1909189
------------------------------------------------------------------------------
```

Logit Model Output

```
Logistic regression                             Number of obs     =      6,366
                                                LR chi2(6)        =    1070.49
                                                Prob > chi2       =     0.0000
Log likelihood = -3467.2854                     Pseudo R2         =     0.1337

------------------------------------------------------------------------------
           A |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      rating |  -.7165643    .031318   -22.88   0.000    -.7779464   -.6551822
         age |  -.0632411   .0099064    -6.38   0.000    -.0826572   -.0438249
       years |   .1126725   .0093673    12.03   0.000     .0943129    .1310321
 religiosity |  -.3749785    .034622   -10.83   0.000    -.4428364   -.3071205
1.managerial |   .4730845   .0876404     5.40   0.000     .3013124    .6448566
1.professional |   .5271217   .2217386     2.38   0.017     .0925221    .9617213
       _cons |   3.751573   .2620388    14.32   0.000     3.237987     4.26516
------------------------------------------------------------------------------

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |      6,366   -4002.53  -3467.285       7    6948.571   6995.882
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

Logit Model AMEs

```
Average marginal effects                        Number of obs     =      6,366
Model VCE    : OIM

Expression   : Pr(A), predict()
dy/dx w.r.t. : rating age years religiosity 1.managerial 1.professional

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      rating |  -.1306759   .0048316   -27.05   0.000    -.1401456   -.1212062
         age |  -.0115329    .001788    -6.45   0.000    -.0150373   -.0080286
       years |   .0205475    .001644    12.50   0.000     .0173254    .0237696
 religiosity |  -.0683828   .0061268   -11.16   0.000    -.0803911   -.0563744
1.managerial |    .090605   .0174113     5.20   0.000     .0564795    .1247306
1.professional |   .1020851    .044947     2.27   0.023     .0139906    .1901796
------------------------------------------------------------------------------
```

Ordered Probit Model

```
Ordered probit regression                       Number of obs     =      6,366
                                                LR chi2(15)       =     236.49
                                                Prob > chi2       =     0.0000
Log likelihood = -7808.2421                     Pseudo R2         =     0.0149
```

rating	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	-.0047552	.0047154	-1.01	0.313	-.0139971	.0044867
years	-.0070395	.0050613	-1.39	0.164	-.0169594	.0028804
children	-.0632364	.0153484	-4.12	0.000	-.0933187	-.0331541
religiosity	.1310093	.0161123	8.13	0.000	.0994298	.1625888
education	.0140007	.0081484	1.72	0.086	-.0019699	.0299713
occupation						
2	-.1317628	.1824797	-0.72	0.470	-.4894165	.2258909
3	-.2022679	.1794037	-1.13	0.260	-.5538926	.1493568
4	-.0632041	.1798682	-0.35	0.725	-.4157393	.2893311
5	-.1434889	.1826641	-0.79	0.432	-.5015038	.2145261
6	-.2022114	.2097835	-0.96	0.335	-.6133796	.2089568
husbandocc						
2	-.1708474	.0822583	-2.08	0.038	-.3320708	-.0096241
3	-.1705628	.0907884	-1.88	0.060	-.3485048	.0073792
4	-.0981205	.0797453	-1.23	0.219	-.2544184	.0581774
5	-.0671601	.0806878	-0.83	0.405	-.2253052	.090985
6	.0409246	.0912108	0.45	0.654	-.1378452	.2196944
/cut1	-2.221501	.2388776			-2.689692	-1.753309
/cut2	-1.522325	.2363926			-1.985646	-1.059004
/cut3	-.7805359	.2356443			-1.24239	-.3186815
/cut4	.1906019	.2355468			-.2710615	.6522652

Ordered Probit Model AMEs

```
Average marginal effects                        Number of obs     =      6,366
Model VCE    : OIM

dy/dx w.r.t. : children religiosity 2.husbandocc
1._predict   : Pr(rating==1), predict(pr outcome(1))
2._predict   : Pr(rating==2), predict(pr outcome(2))
3._predict   : Pr(rating==3), predict(pr outcome(3))
4._predict   : Pr(rating==4), predict(pr outcome(4))
5._predict   : Pr(rating==5), predict(pr outcome(5))

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
children     |
    _predict |
          1  |   .0023971   .0006187     3.87   0.000     .0011844    .0036098
          2  |   .0059013   .0014541     4.06   0.000     .0030514    .0087513
          3  |   .0103042   .0025046     4.11   0.000     .0053953     .015213
          4  |   .0055807    .001379     4.05   0.000      .002878    .0082835
          5  |  -.0241833   .0058533    -4.13   0.000    -.0356555   -.0127111
-------------+----------------------------------------------------------------
religiosity  |
    _predict |
          1  |  -.0049661   .0007522    -6.60   0.000    -.0064403   -.0034919
          2  |   -.012226   .0015932    -7.67   0.000    -.0153486   -.0091034
          3  |  -.0213475   .0026396    -8.09   0.000     -.026521   -.0161741
          4  |  -.0115618   .0015105    -7.65   0.000    -.0145223   -.0086013
          5  |   .0501015   .0060891     8.23   0.000     .0381671    .0620359
-------------+----------------------------------------------------------------
1.husbandocc |  (base outcome)
-------------+----------------------------------------------------------------
2.husbandocc |
    _predict |
          1  |   .0063213   .0027402     2.31   0.021     .0009505     .011692
          2  |   .0157578   .0070883     2.22   0.026      .001865    .0296506
          3  |   .0278462    .013252     2.10   0.036     .0018728    .0538196
          4  |   .0156119    .009041     1.73   0.084    -.0021083     .033332
          5  |  -.0655371    .031859    -2.06   0.040    -.1279795   -.0030947
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

Ordered Logit Model

```
Ordered logistic regression                     Number of obs     =      6,366
                                                LR chi2(15)       =     224.52
                                                Prob > chi2       =     0.0000
Log likelihood = -7814.2247                     Pseudo R2         =     0.0142

------------------------------------------------------------------------------
      rating |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0052979   .0080033    -0.66   0.508    -.0209841    .0103883
       years |  -.0128822    .008628    -1.49   0.135    -.0297928    .0040283
    children |  -.1035099   .0262379    -3.95   0.000    -.1549352   -.0520846
 religiosity |   .2213074   .0272022     8.14   0.000     .1679921    .2746228
   education |   .0242361   .0138223     1.75   0.080    -.0028551    .0513273
             |
  occupation |
          2  |  -.1672703   .3016406    -0.55   0.579    -.7584751    .4239344
          3  |  -.3138057   .2962706    -1.06   0.290    -.8944855     .266874
          4  |  -.0808011   .2971164    -0.27   0.786    -.6631386    .5015363
          5  |  -.2308718   .3018664    -0.76   0.444    -.8225192    .3607755
          6  |  -.2984046   .3502441    -0.85   0.394    -.9848704    .3880612
             |
  husbandocc |
          2  |  -.2698151   .1363468    -1.98   0.048    -.5370499   -.0025803
          3  |  -.2649377   .1510062    -1.75   0.079    -.5609044     .031029
          4  |  -.1447067   .1319604    -1.10   0.273    -.4033443    .1139309
          5  |  -.0883686   .1335461    -0.66   0.508    -.3501141    .1733769
          6  |   .0656399   .1515847     0.43   0.665    -.2314607    .3627404
-------------+----------------------------------------------------------------
       /cut1 |  -4.095151   .4046165                     -4.888185   -3.302117
       /cut2 |  -2.521576   .3946938                     -3.295162    -1.74799
       /cut3 |  -1.145174   .3926742                     -1.914802    -.375547
       /cut4 |   .4424099   .3925007                     -.3268773    1.211697
------------------------------------------------------------------------------
```

Ordered Logit Model AMEs

```
Average marginal effects                        Number of obs     =      6,366
Model VCE    : OIM

dy/dx w.r.t. : children religiosity 2.husbandocc
1._predict   : Pr(rating==1), predict(pr outcome(1))
2._predict   : Pr(rating==2), predict(pr outcome(2))
3._predict   : Pr(rating==3), predict(pr outcome(3))
4._predict   : Pr(rating==4), predict(pr outcome(4))
5._predict   : Pr(rating==5), predict(pr outcome(5))

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
children     |
    _predict |
          1  |   .0015858   .0004318     3.67   0.000     .0007394    .0024322
          2  |   .0051409    .001327     3.87   0.000       .00254    .0077418
          3  |   .0110726   .0028081     3.94   0.000     .0055689    .0165764
          4  |   .0066989   .0017215     3.89   0.000     .0033249    .0100729
          5  |  -.0244983   .0061891    -3.96   0.000    -.0366287   -.0123678
-------------+----------------------------------------------------------------
religiosity  |
    _predict |
          1  |  -.0033905   .0005356    -6.33   0.000    -.0044402   -.0023408
          2  |  -.0109914   .0014541    -7.56   0.000    -.0138414   -.0081415
          3  |  -.0236737   .0029234    -8.10   0.000    -.0294035   -.0179438
          4  |  -.0143224    .001845    -7.76   0.000    -.0179385   -.0107064
          5  |   .0523781   .0063359     8.27   0.000     .0399599    .0647962
-------------+----------------------------------------------------------------
1.husbandocc |  (base outcome)
-------------+----------------------------------------------------------------
2.husbandocc |
    _predict |
          1  |   .0040976   .0019365     2.12   0.034     .0003022     .007893
          2  |   .0133108   .0062704     2.12   0.034      .001021    .0256006
          3  |   .0288335   .0141147     2.04   0.041     .0011691    .0564979
          4  |   .0177781    .010678     1.66   0.096    -.0031505    .0387066
          5  |    -.06402   .0327466    -1.96   0.051    -.1282021    .0001622
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level
```

Poisson Model

```
Poisson regression                              Number of obs     =     27,308
                                                LR chi2(7)        =    2123.85
                                                Prob > chi2       =     0.0000
Log likelihood = -12362.865                     Pseudo R2         =     0.0791

------------------------------------------------------------------------------
    hospvis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        age |  -.0071362   .0015424    -4.63   0.000    -.0101592   -.0041132
       hsat |   -.186052   .0069948   -26.60   0.000    -.1997616   -.1723424
       educ |   -.045944   .0084623    -5.43   0.000    -.0625299   -.0293582
     docvis |   .0306762   .0011552    26.55   0.000      .028412    .0329404
  1.handdum |   .2691112   .0439668     6.12   0.000     .1829378    .3552846
    1.addon |   .3218622   .1077161     2.99   0.003     .1107425    .5329819
    1.bluec |  -.1103086   .0405875    -2.72   0.007    -.1898586   -.0307585
      _cons |  -.1971156   .1345566    -1.46   0.143    -.4608416    .0666105
------------------------------------------------------------------------------
```

Poisson Model AMEs

```
Average marginal effects                        Number of obs     =     27,308
Model VCE    : OIM

Expression   : Predicted number of events, predict()
dy/dx w.r.t. : age hsat educ docvis 1.handdum 1.addon 1.bluec

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        age |  -.0009865   .0002138    -4.61   0.000    -.0014056   -.0005674
       hsat |  -.0257194   .0010537   -24.41   0.000    -.0277846   -.0236543
       educ |  -.0063512   .0011744    -5.41   0.000    -.0086529   -.0040495
     docvis |   .0042406    .000174    24.38   0.000     .0038996    .0045816
  1.handdum |   .0403667   .0071613     5.64   0.000     .0263309    .0544026
    1.addon |   .0521476   .0202966     2.57   0.010      .012367    .0919282
    1.bluec |  -.0148143   .0052965    -2.80   0.005    -.0251954   -.0044333
------------------------------------------------------------------------------
```

OLS Regression

Source	SS	df	MS			
				Number of obs	=	27,308
				F(7, 27300)	=	91.40
Model	489.306593	7	69.9009419	Prob > F	=	0.0000
Residual	20877.8454	27,300	.764756242	R-squared	=	0.0229
				Adj R-squared	=	0.0226
Total	21367.152	27,307	.782478925	Root MSE	=	.8745

hospvis	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-.0015269	.0005016	-3.04	0.002	-.0025101	-.0005436
hsat	-.0240788	.0026008	-9.26	0.000	-.0291765	-.0189811
educ	-.0049511	.0024185	-2.05	0.041	-.0096916	-.0002106
docvis	.0170985	.0010107	16.92	0.000	.0151175	.0190795
1.handdum	.0580785	.0183262	3.17	0.002	.0221582	.0939988
1.addon	.0443904	.0390841	1.14	0.256	-.0322164	.1209971
1.bluec	-.0159247	.0128963	-1.23	0.217	-.0412021	.0093527
_cons	.3663938	.0439651	8.33	0.000	.28022	.4525676

References

[1] Regina T Riphahn, Achim Wambach, and Andreas Million. Incentive effects in the demand for health care: a bivariate panel count data estimation. *Journal of applied econometrics*, 18(4):387–405, 2003.

Negative Binomial Model

```
Negative binomial regression                    Number of obs     =     27,326
                                                LR chi2(3)        =     676.70
Dispersion     = mean                           Prob > chi2       =     0.0000
Log likelihood =  -10037.84                     Pseudo R2         =     0.0326

------------------------------------------------------------------------------
     hospvis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   -.006069   .0022079    -2.75   0.006    -.0103964   -.0017417
        hsat |  -.2193225   .0099992   -21.93   0.000    -.2389206   -.1997244
     handdum |   .4574462   .0707898     6.46   0.000     .3187007    .5961916
       _cons |  -.4580444   .1258712    -3.64   0.000    -.7047473   -.2113414
-------------+----------------------------------------------------------------
    /lnalpha |   1.896108   .0403598                      1.817004    1.975212
-------------+----------------------------------------------------------------
       alpha |   6.659923   .2687934                      6.153396    7.208146
------------------------------------------------------------------------------
LR test of alpha=0: chibar2(01) = 5175.62              Prob >= chibar2 = 0.000
```

References

[1] Regina T Riphahn, Achim Wambach, and Andreas Million. Incentive effects in the demand for health care: a bivariate panel count data estimation. *Journal of applied econometrics*, 18(4):387–405, 2003.