In this assignment, we work on a paper that analysis the educational effect of providing eyeglasses to students in China in an experimental setting. We discuss the second part of the paper that analyzes students' choices over accepting eyeglasses (Part 8 and table 8). The paper and the revised data (PS06-Glewwe2016) are on ELMS.
Glewwe, Paul; Albert Park and Meng Zhao, 2016, "A better vision for development: Eyeglasses and academic performance in rural primary schools in China." Journal of economic development. 122, 170-182.

The assignment has two parts. In part 1, you have to read the paper and answer the questions.
In part 2, you should use Stata. You need to write a do file that contains all the codes, and then generate the log file.
(In part 2, copy and paste the results from Stata. Use the font Lucida Console to keep the format of Stata results.)
You must upload this file (**HW3.doc**) with answers as well as **the log file** on ELMS.

Part 1:

Read the paper and answer the following questions:

a. What are the main questions the paper asks in part 8?

The main question the paper asks is how eyesight affects academic success, whether it improves or worsens a student's academic standing.

b. What is the variable of interest in part 8? Is it a binary variable or a continues variable? What category is one and what is base?

The main variable of interest in the article is how eye glasses correlate to the academic success of a student in China.

c) Explain what do you expect the sign of eye acuity to be and why?

When a student's vision is enhanced through the use of eyeglasses, it is reasonable to anticipate an accompanying improvement in their academic performance, a phenomenon attributed to several key factors. Firstly, the sharpened visual perception afforded by better vision enables students to see educational materials more clearly, which, in turn, facilitates improved comprehension and retention of information. Additionally, students with improved vision are more likely to actively engage in classroom activities, as they can participate without the hindrance of visual struggles, thus positively impacting their academic performance. Eyeglasses also play a role in reducing eye strain, making it easier for students to focus and concentrate on their studies. Moreover, the use of eyeglasses that correct vision problems can boost a student's self-esteem and confidence, ultimately influencing their overall academic performance and motivation in a positive manner. Consequently, in the context of rural primary schools in China, where improved eye acuity resulting from eyeglasses is a relevant concern, researchers would anticipate a positive relationship between eye acuity and academic performance.

Part 2:

The variables and their description is as follows:

ariable	efinition
ountycode towncode choolcode idcode	odes for county, township, school, and student

grade	Student's grade (1 to 5)
female	=1 for female, =0 for male
birthdate; examdate	Student's birth date and exam date
efteye; righteye	Student's eye exam results for left eye and right eye. It is a measure between 4 and 5.2, with 4 unable to read any line (out of 12 lines) in the eye exam board and 5.2 being able to read all the lines.
eight; weight	Student's height (centimeters) and weight (kilograms)
eadeduc; headocc	Household head's education (codes 1-8) and occupation (1-8) Education (=years): 1=16y; 2=14; 3,4,5=12y; 6=9y; 7=6y; 8=0y. Occupation: 1=farmer; 2=worker; 3=teacher; 6=village leader, 4,5,7,8=others.
lasses	=1 if the student had glasses before project started = if not
eceived	=1 if the student received glasses in project = otherwise
ligible	=1 if the student was considered as eligible to receive glasses =0 otherwise
ffer	=1 if the student was offered glasses in the project =0 otherwise
hinese04s2 math04s2 cience04s2	Test scores in Chinese, math, and science in 2004 (before the project started)
ownincpc	Township income per capita

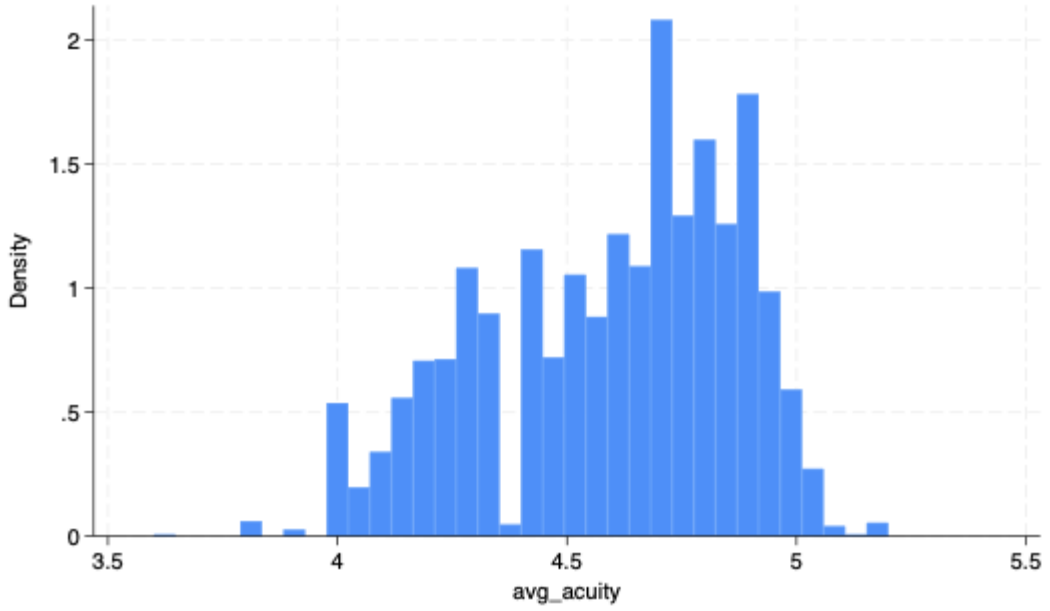
1. Generate the following new variables:

- age [use the formula: **(examdate-birthdate)/365**]
- average acuity of two eyes,
- **one variable** for years of education [You have to define a variable and replace it with 16 if the value in the variable headeduc is 1; replace it with 14 if the value in the variable headeduc is 2; replace it with 12 if the value in the variable headeduc is 3 or 4 or 5; replace it with 9 if the value in the variable headeduc is 6; replace it with 6 if the value in the variable headeduc is 7; replace it with 0 if the value in the variable headeduc is 8.]
- **four dummies** for head occupation being farmer, worker, teacher, or village leader, [Occupation codes are defined in the variable headocc as follows: headocc=1 represents farmer; headocc=2 represents worker; headocc=3 represents teacher; headocc=6 represents village leader, headocc=4 and 5 and 7 and 8 represent others.]
- a dummy variable for previously having glasses. This variable is one if the person had glasses before project, zero otherwise.
- z-score of the average of the three test scores (to do this part you first need to generate a new variable that is the average of three exam scores (lets say avetest). Then find the average and standard deviation of avetest, then generate a new variable (lets say zavetest) that is the avetest score minus the average of avetest divided by the standard deviation of avetest).

. summarize

Variable	Obs	Mean	Std. dev.	Min	Max
countycode	3,134	1.456286	.4981649	1	2
grade	3,134	4.246969	.8051157	3	5
towncode	3,134	10.98724	6.561149	1	22
schoolcode	3,134	79.5434	46.50742	1	164
idcode	3,134	24.1522	27.38272	1	325
female	3,134	.4853223	.4998643	0	1
birthdate	3,134	12297.89	796.1405	-12705	16501
examdate	3,130	16227.92	682.0467	-21911	16283
lefteye	3,126	4.581766	.3020573	3.5	5.2
righteye	3,130	4.606901	.3173848	3.6	5.3
height	3,130	140.4889	9.69358	102	173
weight	3,130	32.99827	6.86251	17	69
received	3,134	.6758137	.4681447	0	1
headeduc	3,133	6.077881	1.105686	1	8
headocc	3,134	1.370134	1.234498	1	8
glasses	3,134	1.958839	.1986952	1	2
chinese04s2	3,094	78.77054	10.46001	14	100
math04s2	3,094	78.72392	12.13549	11	100
science04s2	3,094	80.8092	9.84493	15.5	100
eligible	3,134	.7948309	.4038896	0	1
offer	3,134	1	0	1	1
townincpc	3,134	1542.971	514.8826	594	2516
age	3,130	10.76823	2.846645	-91.63287	79.29315
avg_acuity	3,126	4.594354	.2743058	3.6	5.2
years_of_e~n	3,133	8.352378	2.632745	0	16
farmer_dummy	3,134	.8972559	.3036728	0	1
worker_dummy	3,134	.0210593	.1436051	0	1
teacher_du~y	3,134	.0124442	.1108748	0	1
village_le~y	3,134	.0114869	.1065767	0	1
had_glasse~y	3,134	.6758137	.4681447	0	1
headeducy	67	14.71642	.9661854	14	16

2. Describe the main variable of interest, accepting the eyeglasses using numbers or graph (which is the variable “received” in table above).



3. Run a regression (Just OLS. Do not use logit or probit models) of receiving eyeglasses on eye acuity and some explanatory variables that explain acceptance of eye glasses. (You have to start with a simple regression of accepting on eye acuity and add/drop explanatory variables until you have a good model. You should be able to "theoretically defend" your model. You do not need to report all the steps in here. Try them but **report only the final model you prefe**r. Copy and paste the regression result of the preferred model in here.)

regress avg_acuity weight grade years_of_education

```
      Source |       SS           df       MS      Number of obs   =     3,125
-------------+----------------------------------   F(3, 3121)      =     25.20
       Model |  5.55974695         3  1.85324898   Prob > F        =    0.0000
    Residual |  229.483292     3,121   .07352877   R-squared       =    0.0237
-------------+----------------------------------   Adj R-squared   =    0.0227
       Total |  235.043039     3,124  .075237849   Root MSE        =    .27116

------------------------------------------------------------------------------
       avg_acuity | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
------------------+-----------------------------------------------------------------
           weight |  -.0004654      .0008    -0.58   0.561     -.002034    .0011033
            grade |  -.0495871    .006801    -7.29   0.000     -.062922   -.0362521
years_of_education |   -.003416   .0018464    -1.85   0.064    -.0070364    .0002043
            _cons |   4.848754   .0336485   144.10   0.000     4.782778    4.914729
------------------------------------------------------------------------------
```

4. (**This question is the main part. It explains your findings**) Now run a regression of receiving eyeglasses on the following explanatory variables: average visual acuity, a dummy for female, a dummy for having glasses before program began, multiple dummies for head household occupation, household head's years of schooling, township

per capita income, and z-score of average test scores. Copy and paste the regression result here and write a few paragraphs explaining the results. (You should explain the effect of each variable, its significance, and a short explanation for why the effect is the way it is in your results.)

regress received avg_acuity farmer_dummy worker_dummy had_glasses_dummy headocc years_of_education townincpc

```
      Source |       SS           df       MS      Number of obs   =     3,125
-------------+----------------------------------   F(7, 3117)      =         .
       Model |   684.27392         7  97.7534171   Prob > F        =         .
    Residual |          0     3,117           0   R-squared       =    1.0000
-------------+----------------------------------   Adj R-squared   =    1.0000
       Total |   684.27392     3,124  .219037746   Root MSE        =         0

------------------------------------------------------------------------------
    received | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   avg_acuity |   2.33e-16          .        .        .           .          .
 farmer_dummy |  -3.15e-16          .        .        .           .          .
 worker_dummy |  -2.34e-16          .        .        .           .          .
had_glasses_dummy |          1          .        .        .           .          .
     headocc |  -6.20e-17          .        .        .           .          .
years_of_education |  -1.89e-18          .        .        .           .          .
    townincpc |  -1.72e-21          .        .        .           .          .
        _cons |  -7.77e-16          .        .        .           .          .
------------------------------------------------------------------------------
```

5. Analyze the effect of eye acuity on probability of accepting the offer for male versus female students. (Generate a new variable for the interaction between eye acuity and the dummy for gender. Add this variable to the model described in question 4) Is there any evidence that the effect of eye acuity on probability of accepting the offer depends on gender? Copy and paste the regression result here. Explain.

```
      Source |       SS           df       MS      Number of obs   =     3,125
-------------+----------------------------------   F(8, 3116)      =         .
       Model |   684.27392         8    85.53424   Prob > F        =         .
    Residual |          0     3,116           0   R-squared       =    1.0000
-------------+----------------------------------   Adj R-squared   =    1.0000
       Total |   684.27392     3,124  .219037746   Root MSE        =         0

------------------------------------------------------------------------------
    received | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   avg_acuity |   2.30e-16          .        .        .           .          .
 farmer_dummy |  -3.12e-16          .        .        .           .          .
 worker_dummy |  -2.30e-16          .        .        .           .          .
had_glasses_dummy |          1          .        .        .           .          .
     headocc |  -6.13e-17          .        .        .           .          .
years_of_education |  -1.89e-18          .        .        .           .          .
```

```
       townincpc |  -1.84e-21          .       .       .           .          .
eye_acuity_gender |   4.61e-18          .       .       .           .          .
           _cons |  -7.77e-16          .       .       .           .          .
```

6. We want to analyze the effect of test score on probability of accepting the offer for male versus female students (or equivalently, the gender effect for students with different test scores). It is similar to the previous question. Copy and paste the regression result here. Explain.

```
 Source |       SS           df       MS      Number of obs   =     3,094
-------------+----------------------------------   F(2, 3091)      =    292.81
       Model |  72553.2202         2  36276.6101   Prob > F        =    0.0000
    Residual |   382953.39     3,091  123.893041   R-squared       =    0.1593
-------------+----------------------------------   Adj R-squared   =    0.1587
       Total |   455506.61     3,093  147.270162   Root MSE        =    11.131

------------------------------------------------------------------------------
   math04s2 | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
 science04s2 |   .4913628   .0203292    24.17   0.000     .4515027     .531223
     female |   .4654589   .4004161     1.16   0.245    -.3196496    1.250567
      testz |          0  (omitted)
      _cons |   38.79192   1.666049    23.28   0.000     35.52525     42.0586
```