

Take Test: Homework 1

Test Information

Description In this exercise, you will analyze some data used by Wooldridge (2002). Put all of your commands in a **do-file**, and you will submit the **do-file** at the last question.

Instructions Your process for each question should be to

- 1. Put a command in your do- file.
- 2. Run your do- file in Stata.
- 3. Check that your output produces reasonable results, find and submit the corresponding answers or output files among them.
- 4. Go to the next question.

Click Save and Submit to save and submit your answers.

Multiple Attempts This test allows 5 attempts. This is attempt number 1.

Click Force Save All Answers to save all answers. This test can be saved and resumed later.

QUESTION 1

5 points

Saved

The data is in nls_woold_wide.dta. It is in wide format. Use the **reshape** command and obtain the resulting long-form data. Use the command **label** with syntax, **label variable variablename "your label"**, to label the variables year, wage, exper, and inlf. That is, you have to run the following commands:

```
label variable year "year"
label variable wage "wage"
label variable exper "experience"
label variable manuf "individual belongs to manufacturing industry"
label variable inlf "in the labor force"
```

Save the new long formed dataset with updated labels using **save nls_woold_long.dta**. Upload the new long-formed data set.

Selected Answer: nls_woold_long.dta Remove

Attach File Browse My Computer Browse Content Collection

QUESTION 2


5 points

Saved

Use the command to **describe**, and **xtsum** answer the following questions.
The total number of observations including all time-series across all individuals is , and the panel-data consists of panels (individuals) with time periods (except wage and manuf). The number of total variables is .

QUESTION 3

5 points

 Saved

Use a command **generate** the new variable lwage containing the log of wage. Label the new variable as "log of wage". Use **xtsum**, and enter the standard deviation of lwage across individuals. (You can round up to two decimals points.)

QUESTION 4

5 points

 Saved

Use a command **generate** the new variable exper2 containing squared values of exper. Label the new variable as "squared value of experience". Use **xtsum**, and enter the *within* standard deviation of the new variable exper2. (You can round up to two decimals points.)

QUESTION 5

5 points

 Saved

xtsum all the variables in your dataset. The number of variables that have non-zero within standard deviations, and thus time-varying, is . Also, the number of the time-invariant variables within zero within standard deviations . (Note: **DO NOT count the variables id and year in both blanks**)

QUESTION 6

5 points

Save Answer

The next 11 problems run a series of regressions. In these regressions, we will use covariates are exper exper2 manuf, black, and educ. Note that the first three covariates are time-varying and the last two are time-invariant.
Use **regress** to perform Poled OLS of lwage on the covariates defined in the above paragraph, and time-dummies. For the time-dummies, use the command **i.year** to include year indicators variables. Request a cluster-robust estimator of the VCE, clustering on id. Also, use the command **estimates store POLS** to save the regression output.
What is the p-value of estimated coefficient for manuf?

QUESTION 7

5 points

Save Answer

Calculate the predicted values of your regression by using the command **predict**, with syntax **predict hatlw**, where hatlw are the fitted values from the regression. Draw a scatter plot of

haltw and experience, where haltw is on y-axis and experience is on x-axis. Save and upload the scatter plot.

Attach File

Browse My Computer

Browse Content Collection

QUESTION 8

5 points

Save Answer

Use **xtreg**, **re** to perform *a random-effects* regression of lwage on the same set of covariates in the previous question 7. Here, you must request a default estimator of the standard error or VCE (which is equivalent to not using , **cluster(id)** option), so that we can perform a Hausman test in the later question. Use **estimates store RE** to save the results. Lastly, save the estimation sample in a new variable, in other words **generate variable = e(sample)**, you will need it to answer a subsequent question. What is the standard error for the covariate exper? (You can round up to four decimal points.)

QUESTION 9

6 points

Save Answer

Use **xtreg**, **fe** to perform a fixed-effects regression of lwage on the time-varying covariates, exper exper2 manuf, and the year indicator variables. (Use i.year to include year indicators variables.) Here, request an IID estimator of the VCE, so that we can perform a Hausman test in the later question. Use **estimates store FE** to save the results. What is the value of the estimated coefficient for exper2? (You can round up to four decimal points.)

QUESTION 10

6 points

Save Answer

Use **estimates table** to produce a table of the coefficient estimates obtained thus far. Is the value of the estimated coefficient for "exper" in Random Effect Model is smaller than the one in the Fixed Effect Model?

- ☐ Yes
- ☐ No

QUESTION 11

6 points

Save Answer

Stata does not allow you to type **D.i.year**. (The reason is that D.i.year and i.D.year are different and this difference violates the order invariance property.) We solve this problem by using a command below:

tabulate year, gen(iyear)

What is the cumulative percentage of the observation in year 85?

QUESTION 12

6 points

Save Answer

Use a command **regress** to obtain the FD estimator of the coefficients on the time-varying covariates, exper exper2 manuf, and the year indicator variables. (Use the iyear indicators previously generated.) Request a *cluster-robust estimator* of the VCE. Use **estimates store fd** to save the results. What is the estimated cluster robust standard error for the coefficient of the first-differenced exper ? You can round up to four decimal points.

QUESTION 13

6 points Save Answer

Use **estimates table** to produce a table of the coefficient estimates from fixed effect and first differenced models. The value of the estimated effect for "manuf" in First Differenced Model is [smaller/larger] than the one in the Fixed Effect Model.

QUESTION 14

10 points Save Answer

Use **hausman** to perform a Hausman test of the random-effects assumptions. Use the option *sigmamore* to use the more efficient estimate of sigma square of epsilon.

The null hypothesis of the Hausman test states that there [is/isn't] a systematic difference in coefficients between RE/FE estimators because the unobserved time-invariant component alpha_i is [exogenous/endogenous] to time-varying regressors X_it. And the output of Hausman test that you just performed reports that the p-value of the test statistic under the null hypothesis is [number], and this result suggests that we [cannot/can] reject the null hypothesis.

Choose a right combination of words inside the brackets []:

- ☐ isn't, exogenous, 0.7426, cannot
- ☐ is, endogenous, 0.6426, cannot
- ☐ isn't endogenous 0.7426 can
- ☐ [is, exogenous, 0.8426, cannot]

QUESTION 15

10 points Save Answer

Use **egen mean()** with the by prefix and the for each command to generate the panel-level means of the time-varying variables. Hints: Include the mean effects of the year dummies, remember that the time-indicators vary over time. Thus, you have to enter the following commands:

```
local explain = "exper exper2 manuf iyear81 iyear82 iyear83 iyear84 iyear85 iyear86 iyear87"
foreach var of varlist `explain' {
    bysort id: egen m_`var' = mean(`var') if samp_re
-}
```

Now, use xtreg, re vce(cluster id) to perform random-effects with the same variables as in the previous random-effects and the Mundlak variables. Use estimates store Mundlak to save the results.

According to the result in Mundlak regression, the panel-level means of the time-varying variables are all statistically insignificant. True or False?

- ☐ True
- ☐ False

QUESTION 16

QUESTION 10

10 points Save Answer

Use **estimates table** to produce a table of the estimates from Random Effect, Fixed Effect, and Mundlak models. Upload your do file that includes all of your work.

Attach File

Browse My Computer

Browse Content Collection
