

Data Exercise 2
Econ 322: Econometrics
Fall 2019
Yutian Yang

Upload this data exercise to Sakai by Monday, Oct 23rd at 12:00 p.m.

Your document should include responses to the question. The answers themselves should be attached to the Sakai assignment as *Yourname_p2.pdf* file (export to .pdf from whatever software you use). Make sure your name is at the top. You may work together in small groups of 2–3 if you would like, but the write-up must be your own; you may not turn in identical write-ups. Please write the names of your collaborators on your homework as well. Also, include your R script as a separate upload, with *yournamen_p2.R* as the filename. You can also upload a zip file with file name “*yourname_p2.zip*”.

Note: For all the answers based on regressions, create a regression table with:

- the mean of the estimated coefficients
- the standard errors of coefficients,
- Indicate the p -value (** indicates $p < 0.01$, * indicates $p < 0.05$, and $p < 0.10$).
- the 95% confidence intervals,
- when talking about the significant level of an estimated coefficient, always say it is **significant at a certain level (1%, 5% or 10%)** or not;
- Do not just copy the R output and paste it into the table. Instead, create your own professional-looking table that reports the R output in a way that is easy to read and understand.
- The R tutorial has instructions on how to do these things. You should also Google around for R help!

Your table should be of “publication quality” with appropriate titles, labels, and table notes.

Programming Resources

- R Videos and Workshop Materials libguides.rutgers.edu/data/data_R
- Tutorials for Learning R r-bloggers.com/how-to-learn-r-2
- Introduction to R Seminar stats.idre.ucla.edu/r/seminars/intro
-

Additional Notes

- Your code should have comments at the top. At a minimum, include your name and that the code is for Econ 322, data assignment 1.
- Download the R script from Sakai. Fill in the blanks of the codes.
- First, set your working directory to the folder where your data is located (Session > Set Working Directory > Choose Directory).
- **The project requires a package “AER”. Install the package before running your codes.**
- Answer the following questions based on regressions and summarized statistics:
Hint:

1. When running a quadratic regression, use `lm(A~B+I(B^2))` to represent the quadratic form.
2. When running a log regression, simply use `lm(log(A)~log(B))` to represent the log-log form.
3. There are some new commands such as “`ivreg()`” and “`ar.ols()`” needed to solve the problems. Google and learn how to use them.

Question 1: Nonlinear Regression “How Does Income Affect Test Score?”

1. Run the prior commands to import the test score data from AER. Create two variables “size” and “score”.
2. Run an OLS regression with test score as the dependent variable, income as the independent variable. Test if there is a quadratic impact of income on test score (Robust standard errors). Write the answer in your report and provide evidence.
3. Run an OLS regression to estimate how income change affects the percentage change of test score (Robust standard errors). Write the answer in your report and provide evidence.
4. Run an OLS regression to estimate the elasticity of test score over income (Robust standard errors). Write the answer in your report and provide evidence.

Question 2: Instrumental Variables “How Does the Price Affect Cigarette Consumption?”

1. Run the prior commands to import the test score data from AER. Create two variables “rprice” and “salestax”. Only pick the data from year 1995 by creating a subset called c1995. The following estimations should all base on this subset.
2. Run an OLS regression with “log(packs)” as the dependent variable, “log(rprice)” as the independent variable (Robust standard errors). Find the estimated elasticity and write it in the report.
3. Run an TSLS regression to estimate the price elasticity of cigarette consumption. Use “salestax” as the instrumental variable (Robust standard errors). Compare the outcomes of the TSLS regression and the OLS regressions above and discuss whether there is endogeneity problem of the OLS regression. Write your answer in the report.

Question 3: Time Series “How to predict GDP growth?”

1. Download the “us_macro_quarterly.xlsx” data from Sakai. Import the data into Rstudio before you run any codes.
2. Run the prior commands to format date column and adjust the column names.
3. Run the prior commands to generate time series data set.
4. Run the prior commands to generate variable “GDPgrowth”
5. Run a auto-regressive regression with only 1 period lag using “ar.ols()” command. Show the coefficients in your report.