DSBA/ECON 6112
Graduate Econometrics

Replication Project: Part 2

The following paper is posted to Canvas, along with the dataset:

Angrist, Joshua, Eric Bettinger, and Michael Kremer. 2006. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review*, 96 (3): 847-862.

Instructions:

- Read the paper.

- Replicate the results in Columns 1-4 of Table 2
  - The table (values) in your final write-up must match your do-file and log file.
  - Format the table as in the published paper (rounding, parenthesis around standard errors, etc.)
  - If there are any discrepancies between your results and those reported in the paper, you should highlight these and put a note explaining that this is the case. (These are generally due to typos, so there will be no more than a few in a single table.)

- Briefly (< 1 page) summarize the analysis and results presented in your table:
  - What estimator do you use?
  - What are the dependent and independent variables?
  - What do we learn from these results?
  - Interpret at least one regression coefficient estimate.

- You may consult with the same 1-2 other classmates as you did for Part 1, as long as you (1) include their name on your assignment, (2) submit your own *unique* answers and Stata files, and (3) do not consult any other classmates (but you may consult me).

You must submit the following files in Canvas:
- Do File
- Log File
- PDF or Word document that contains no more than 2 pages: one page for the formatted table and (no more than) one page for the written summary.

Note: The do-file and log file must show all work with the data, including opening the dataset, any data manipulations, and all subsequent code to produce the final output for your tables and responses.

Tips for Stata work:

- Label each line in your-do file with what number(s) you are replicating with the command. For example, put "*Table 2, column 1, Panel A" on one line, and then have your next line of code below this actually replicate the number(s).
- Remember to estimate "robust" standard errors: you need to add ", robust" (note the comma!) at the end of your regress command line. For example, "regress wage educ , robust" would give you the robust standard errors in our wage equation.
- To use only part of the sample, we use "if" statements with relational operators (=, <, >). For multiple restrictions, we use "&" between them.

Things to keep in mind:

- You are being asked to replicate part of a research paper *to practice doing research*, so treat the assignment as such.
- Your grade will be primarily based on the accuracy of all of the numbers in replicating the research paper, but formatting is also important and will count towards your grade. ***Your table formatting should exactly match the paper table, including typing the note below the table***.
- You must use Stata to calculate all of the numbers in the table, but you may use any software you'd like to actually put the numbers together in a table.

Grading rubric:

	Points Possible	**Evaluation criteria**
Do File	5	Complete, executes without error, produces log file
Log File	5	Complete, produced by do-file, clean (easy to read/find values), matches table values
Summary	5	Describe estimator, variables, findings, interpret coefficient
Table 2		
Formatting	5	Follow formatting instructions
Mean y, N	9	Reported numbers match paper and Stata files
Column 1 est, se	6	Reported numbers match paper and Stata files
Column 2 est, se	14	Reported numbers match paper and Stata files
Column 3 est, se	6	Reported numbers match paper and Stata files
Column 4 est, se	14	Reported numbers match paper and Stata files
TOTAL	**69**	

Note: "est" is short for "estimate" and "se" is short for "standard error"