4th Annual Econometrics Game

Hosted by OEconomica

April 7, 2018

1 Objective

Welcome to the 4th Annual Econometrics Game! **Your task is to use the 2008-2009 fracking boom to identify a causal effect.**

The fracking boom, a large increase in oil and natural gas production resulting from new extraction technology, likely had direct impacts on local health, economic, and social outcomes. In addition, the fracking boom has been used as an exogenous increase in income to explore more fundamental theories of behavior. As inspiration, consider the recent work of Melissa S. Kearney and Riley Wilson: "Male Earnings, Marriageable Men, and Non-Marital Fertility: Evidence From the Fracking Boom," (2017). Their paper is provided. Note, the amount of oil and natural gas actually extracted is endogenous to local circumstances. As such, the authors' identification strategy is to instrument for production with the underlying geography of the region (where the shale deposits are).

We've provided you with two unique datasets. The first is Kearney and Wilson's dataset. The key variable here is the estimate from their first stage. You should take this variable to be the exogenous component of oil and natural gas production (the raw production values and the shale deposit locations are private and not shared). The second is Gallup respondent level polling data. Gallup surveys people for their views on a variety of economic, political and well-being questions. The Gallup data is presented for you to use as an outcome variable. You are welcome to use other outcome variables instead. The focus should be on identifying the causal effect of fracking on some outcome (not necessarily on the Gallup responses). **Prior to using the Gallup data, you must sign a Data Use Agreement if you are not from the University of Chicago.** If you have not signed a Data Use Agreement, please let a member of Oeconomica know.

2 Data

1. **Fracking Data**: A panel dataset consisting of first-stage estimates of fracking production (the results of Kearney and Wilson's regression number 1), income, and other demographics at the PUMA level from 2000 to 2012. A data dictionary is provided.
2. **Gallup Data**: The U.S. Daily Tracking Poll is a daily survey asking 1,000 U.S. adults about various political, economic, and well-being topics. The available data is from 2008-2016, and the smallest geographic identifier provided is Zip Code. A data dictionary and methodology write-up is provided. The Gallup Data is provided in both SPSS and STATA formats.
3. To join the datasets, you will need to use a crosswalk from Zip Codes to PUMAs. We recommend using the Missouri Data Center for this: http://mcdc.missouri.edu/websas/geocorr14.html

3 Submission and Judging

At **12pm** tonight you must submit an academic paper (i.e. in the format of introduction, literature review, data description, model, results, conclusion, and appendices), as well as a short slide deck presentation of your results. The top five teams will be selected based on the extent to which their paper and presentation:

1. Showcases understanding of economic, econometric, and statistical techniques
2. Is complete, well-argued, internally consistent, and readable
3. Sheds new light on an important economic question
4. Is unique and interesting

At 10am the morning of Sunday, April 8th, we will have a breakfast with all the teams, as well as representatives from our corporate sponsors, Brattle Group and Cornerstone Research. Following breakfast, at 11am the top five teams will be announced. The top five teams will have have half an hour to rehearse their presentation. Presentations will start at 11:30am, and each team will have 15 minutes to present (the slide deck they prepared the evening before). The panel of judges will award first, second, and third place prizes.

The judges are: Professor Lars Peter Hansen (UChicago, Economics), Professor James Heckman (Uchicago, Economics), Professor Stéphane Bonhomme (UChicago, Econometrics), Eliot Abrams (PhD candidate in Economics at the University of Chicago Booth School of Business).

- **Side Mission**: In addition to first, second and third place, a prize will be awarded for the best visualization.

4 Other Remarks

The competition is from 10:00am to 12:00am (midnight) on April 7th in Rooms C01, C02, C03 in Booth Harper Center, and 10:00am to 1:30pm on April 8th in Room 147 in Saieh Hall for Economics. You are not to consult any individuals outside your team during the competition. If you have any additional questions, please let the organizers Lindsey Currier (lcurrier@uchicago.edu) and Johnny Ma (johnnyma@uchicago.edu) know.

Best of luck!

4.1 Data tips

1. If you are reading the Gallup data into R, we recommend using the function read_spss for the SPSS version, or read_dta for the STATA version, both from the package "haven". If you convert the files to CSV and then read them into R, some of the values for the variables might not show up.
2. If you are reading the Gallup data into STATA, note that even if there is a value label, the values of the variables may not show up with –browse– or –tabulate– for a few variables. The following command will let you know if there is a value label: labelbook varname. To convert the value labels back to their original values: –strip–label varname. Note: you can still conduct analyses with value label; even if values appear blank, they are not treated as missing.
3. Note that PUMA numbers are not unique, they are unique to a given state.

Sponsored by:

BECKER FRIEDMAN INSTITUTE FOR RESEARCH IN ECONOMICS THE UNIVERSITY OF CHICAGO

CORNERSTONE RESEARCH

Brattle THE POWER OF ECONOMICS