**PB4A7 Summative coursework submission**

**I. Data analysis replication report and extension (2500 words)**

Write a data analysis replication report of [“*Benjamin Hansen, 2015. Punishment and Deterrence: Evidence from Drunk Driving American Economic Review”*](https://www.aeaweb.org/articles?id=10.1257/aer.20130189)

**Part 1: Replication:** In your report you will have to briefly summarise the study and hypothesis, and then replicate, or at least attempt to replicate, the main result(s) with at least one accompanying graph. To replicate the results, you should write an original Stata program in a do-file that re-estimates at least the main model from the paper.

**Part 2: Critique and extension:** Critique the study (comment on the motivation, theory, model, estimation, interpretation, data, any issues you find with the data, any threats to internal and external validity). You can reference other published papers as appropriate. Give a brief description of how you would extend the results of the paper given the chance. What would you do different? What could you do better? What would you do next? What are the potential issues of RD designs that we need to be aware of? Propose an alternative technique that can address the limitations pointed out and how this could be conducted in a new study.

**Main instruction:** You will receive a cleaned dataset, so you do not have to clean and merge the datasets yourself – you will focus on the data analysis only. You must write an original program to replicate the main result(s) in the paper.[[1]](#footnote-1) The STATA programmes provided by the authors are available for your reference and for your learning, do not use these programs in your replication but rather the do file provided in moodle. The idea is that you should do your best to replicate the main finding by estimating the models. If you can replicate the results, report your results. If you come close, report that and explain your reasoning. If you can’t replicate the results, explain why (is it because you can’t figure it out, or can assumptions made in the paper be challenged, or some combination of those?).

The replication will require 6 questions (4 of them will be discussed on the seminar and an additional 2 will require you to produce your own code):

1. Create the variables for Driving Under the Influence, and quadratic term for Blood alcohol level and produce two histograms of BAC1 (one as a discrete variable, one as a continuous). Explain the differences of the histograms
2. Running regressions on covariates (white, male, age and accident) to see if there is a jump in average values for each of these at the cutoff and explain the results
3. Produce main recidivism results of the paper (with our dataset) using recidivism as dependent variable as well as with a changed bandwith of the RDD to 0.055 to 0.105
4. Replicate Qe by running "donut hole regressions". Explain why one might need a donut hole regression? How do I run a donut hole regression?
5. Run local polunomials and explain why local polynomials might be needed after a donut hole
6. Produce an RDplot and a cmogram with bandwidths of 0.060 & 0.11. Explain what the respective graphs show

**Submission template:**

1. Cover page
   * Note: use only your candidate number, and avoid name and other details– this is not your LSE ID number.
2. Part 1: Replication (1000 words)
   * Briefly summarize the study, its hypothesis, data, and method
   * Replicate the main results and present your replication results in a table.
   * Use at least one graph to visually communicate your arguments.
3. Part 2: Critique and extension (1500 words)
   * critique of the study including the strengths and limitations (both of the paper and in wake of your replication)
   * how can a future study to build on, and improve, the results
4. References
5. Appendix (with any additional methodological notes, additional analysis, **copy of the do file**).

**II. Poster of the data analysis replication report**

Prepare a poster summarising both the replication and the critique and extension. You can use an A0 (841 x 1189 mm or 33.1 x 46.8 inches) poster. The poster should adopt a clear and logical layout and should be informative to a person unfamiliar with the original study or replication.

**Formatting and further considerations:**

**Submission:**

* Submit your assignment as two separate .pdf files (one for the report and another for the poster) via the PB4A7 Moodle submission portal by the deadline. The submission portal will be open on (27/11/2022). You can edit/modify/change your submission until the deadline. You cannot do so afterwards.
* If you are seeking any extension, please first consult LSE’s extension policy: <https://info.lse.ac.uk/current-students/services/assessment-and-results/exceptional-circumstances/extension-policy>

**For the “Data analysis replication report and extension”:**

* You can access the paper from the American Economic Review journal website <https://www.aeaweb.org/articles?id=10.1257/aer.20130189>. The do file and cleaned dataset are available on Moodle.
* The word limit excludes: Cover page, references in the bibliography, text in tables and figures, table and figure titles and notes and the appendix. It includes: All in-text citations, text in heading and sub-headings and footnotes.
* Use this formatting and font:
  + Times New Roman Font, font size 11 and 1.5 spacing, justified.
  + Number sections and subsections sequentially: 2, 2.1, 2.1.1, and so on.
  + Number your appendices sequentially and refer to them in the text where appropriate.
* References: Use either APA6th style of references.
  + All references mentioned in the Reference List must be cited in the text, and vice versa.
  + Please make sure you check this before submitting.
* Other presentation considerations:
  + All identifying information has been removed from the manuscript, including the author name and student ID number (e.g. 2022XXXX). Only your candidate number should be put in the cover sheet (you can find this on LSE For You, under the Student banner).
  + The submission has been 'spell checked' and 'grammar checked'.
  + **Tables should not be a picture of Stata output.**
* The assessment criteria will include considerations like:
  + Clarity and precision of presentation
  + Understanding of topic, theories and methods, and wider literature
  + Critical insight and ability to appropriately evaluate different ideas
  + Practical applications and use of statistical packages and tools
  + Coherence of structure and argument
  + You can find more details in the PBS MSc Handbook: <https://www.lse.ac.uk/PBS/Study>

**For the poster:**

* You can follow the advice here to prepare the poster: <https://blogs.lse.ac.uk/impactofsocialsciences/2018/05/11/how-to-design-an-award-winning-conference-poster/>
* No restrictions on formatting etc., so use your imagination!
* The assessment criteria for the poster will include considerations like:
  + Clear and accurate communication of all key scientific information
  + Logically laid-out and easy to follow
  + Eye-catching and visually appealing
  + Well-researched, appropriately referenced and well-prepared
  + Analytical, critical and synthetic
  + Poster shows understanding of the topic

1. For more on “what is replication” please read this article: <https://doi.org/10.1371/journal.pbio.3000691> [↑](#footnote-ref-1)