Econ 270B: Problem Set 2
Does Working from Home Work? Evidence from a Chinese Experiment: A Computational Reproducibility Exercise

Part I

a. Created an account on SSRP.
b. Selected the article: "Does Working from Home Work? Evidence from a Chinese Experiment" published in the Quarterly Journal of Economics. This article is coauthored by Nick Bloom, James Liang, John Roberts and Zhichun Jenny Ying. All the reproduction materials are publicly available.
c. Scoping
    i. Initiated a "Minimum Reproduction" form on SSRP.
    ii. Main items reproduced: Tables 2, 3, 4, 7, 8, A5, A6; Figure X

Part II

d. Assessment
    i. All the analytic data sets and analysis code required for the computational reproducibility of the paper were made available by the authors. However, the raw data and cleaning code were not made available by the authors. Hence, the ACRe Reproducibility Score is "**Level 5: *Analytic data sets and analysis code are available and they produce the same results as presented in the paper (CRA).***"
    ii. Following is a detailed breakdown of the computational reproducibility for the main tables and figures selected in Part I c:
        - Table 2: Fully reproducible, accurate
        - Table 3: Fully reproducible, accurate
        - Table 4: Fully reproducible, accurate
        - Table 7: Fully reproducible, accurate
        - Table 8: Fully reproducible, accurate
        - Table A5: Fully reproducible, accurate
        - Table A6: Fully reproducible, accurate
        - Figure X: Not fully reproducible. Specifically, the STATA commands for the plot text boxes, bar colors, bar labels and y-axis labels are not included in the do-file
e. Improvement
    i. <u>Missing raw data and cleaning do-files</u>: The material for this paper does not contain the raw datasets that can be further processed to arrive at the analytic datasets. Additionally, the cleaning do-file is also not included in the material. Hence it would be ideal to not only include the analytic datasets and do-file, but also the raw data and cleaning code.
    ii. <u>Incomplete code for figures</u>: All the figures are not fully reproducible as in the paper as they mainly miss the y-axis titles, text boxes, line colors and value labels. This is not only the case with figure X but also for the other figures.

iii. A more detailed Readme file: Currently, the Readme file only contains high-level information that states the author's contact and suggests running the file in the same directory as all the data. However, a more detailed readme file that outlines the systematic steps required to reproduce all the material would be ideal. For example, the readme file could outline all the files in the folder and provide a brief description of each file.

iv. Folder organization: The replication folder can be organized in a standard way into sub-folders titled "data", "do", "tab" and "fig". In addition to this, if the authors also provided the raw data, the "data" subfolder could be further organized into sub-subfolders titled "raw" and "clean".

v. Incorrect numbering of figures in the do-file: The do-file "Tables2014.do" contains the code to create all the figures, however, the figure numbers do not correspond to the ones in the paper. The authors could double-check to make sure all figures correspond to the published version.

vi. Renaming the main do-file: The main do-file is currently named as "Tables2014.do" but can instead be renamed as "output.do" since it contains both tables and figures.

f. Robustness

i. None in particular.

Part III

Abstract

This comment revisits the analysis in Bloom et al. (2015). I do not identify any critical errors with the analytic dataset and code provided by the authors. Having said that, the authors have not provided the raw data and cleaning code, which I believe are important in order to do a thorough reproducibility check of the main findings of the paper. There are also other relatively small improvements that I have suggested in terms of providing a more complete code for the figures, organizing the folders, providing more information in the Readme file, and correcting the numbering of figures in the do-file. Finally, I could not point out any major robustness improvements to the paper. Overall, I do not find any critical errors or issues with the robustness of the estimates that challenge the main empirical claim of the paper suggesting the significant improvements to worker performance and satisfaction from working from home (WFH).