README for "Upping the Ante: The Equilibrium Effects of Unconditional Grants to Private Schools"
Authors: Tahir Andrabi, Jishnu Das, Asim Khwaja, Selcuk Ozyurt, and Niharika Singh
Date: April 2020
AEA Data Code and Repository Project ID: openicpsr-118805

This document guides users interested in replicating results from the above paper. We describe each folder and its contents below.

Data: This folder contains the data required to replicate the analysis in the paper. There are 3 files:
- data_school_panel.dta: This file contains school level variables used in analysis.
  - Unique identifier: mauza_code-school_code-round
- data_teacher.dta: This file contains teacher level variables and school level variables generated from teacher roster data.
  - Unique identifier: mauza_code-school_code-teacher_id-round
- data_school_LassoCleaned: This file contains all variables used in the post double selection Lasso results.
  - Unique identifie: mauza_code-school_code

Do: Contains do-files for replicating the analyses in the paper. There are three files that produce all the analysis in the paper. The list of do-files is given below:
- 0_programs: This do-file contains all the macros and program needed to run the analysis do files.
- 1_tables_main.do: This do-file produces Tables 1-8 in the paper.
- 2a_tables_appendix.do: This do-file produces Tables B1-B20 in the online appendix. Tables that include panels where post double selection Lasso technique is used are produced using file 2b described below.
- 2b_tables_appendix _pdslasso.do: This do-file creates a dataset that can be used to implement the post-double selection Lasso procedure (data_school_LassoCleaned.data) and then uses the dataset to run this procedure to check the robustness of our main results.
- lassoClean.ado: used in file 2b. This file is part of the package provided with NBER Duflo Machine Learning lecture.

Output: This folder contains the results from running the do-files above. Each generated file corresponds to a table number in the main paper (Tables 1-8) or the online appendix (Tables B1-B20). These raw table files are either in .xml or .tex formats; we stitch together panels in TeX for tables that have multiple panels and do some additional formatting in Excel for some tables. After processing, these are all compiled in "AER_Tables.tex" or the "excel_tables_AppendixB.xlsx" to get the final formatted tables available in the paper. The folder also has an excel file with the IRR calculations shown in the paper ("IRR_calculations.xlsx").