

README for the Reproducibility Package for “Command and Can’t Control Assessing Central- ized Accountability in the Public Sector”

This is the replication package for the document “Command and Can’t Control Assessing Centralized Accountability in the Public Sector”, by:

- Saad Gulzar
- Juan Felipe Ladino
- Muhammad Zia Mehmood
- Daniel Rogger

Data statement

All data used are restricted and not included in the reproducibility package. Replicators interested in getting access to the data may contact the School Education Department of Punjab (<https://schools.punjab.gov.pk/>).

Datasets used:

- **Monthly performance reports on primary schools - Punjab, Pakistan:** Data collected by the Program Monitoring and Implementation Unit - School Education Department of Punjab, on a monthly basis in a random set of schools per Markaz. The data was collected between 2009 and 2018. Data was accessed by the team in July, 2019. Corresponding datasets in “data/”: “cc_mea_master.dta”, “cc_mea_master_finance.dta”, “cc_budget.dta”, “lnd_master.dta”, “satt_targets_dp3.dta”, “officers_punjab_pak.csv”.
- **District school performance ranking - Punjab, Pakistan:** Data collected by the Program Monitoring and Implementation Unit - School Education Department of Punjab, on a quarterly basis on the average performance of the schools in each District in the previous three months and the district ranking based on the average performance. The data was collected between 2011 and 2015. Data was accessed by the team in April, 2022. Corresponding datasets in “data/”: “ranking.dta”.
- **Annual school census - Punjab, Pakistan:** Data collected by the Program Monitoring and Implementation Unit - School Education Department of Punjab, on a yearly basis on all primary schools in Punjab. The data was collected between 2011 and 2018. Data was accessed by the

team in July, 2018. Corresponding datasets in “data/”: “anualschoolcensus.dta”, “census2012.xlsx”, “census2013.xlsx”, “census2014.xlsx”, “census2015.xlsx”, “census2016.xlsx”, “census2017.xlsx”.

- **Other datasets:** Supporting data for descriptive figures or robustness. Correspond to time trends on province outcomes between 2010 to 2016 from the Annual Status of Education Report, and voting results in Punjab elections for 2013 at the constituency level. Data was accessed by the team in April, 2023. This datasets are not from the School Education Department of Punjab. Corresponding datasets in “data/”: “asertime.dta”, “political_alignment.dta”.

Folders

1. **Code:** contains the do-files for replicating the results
 - **master:** initializes the folder, set the path and run the other do-files. Additional instructions in the do-file
 - **0_clean_data:** clean the raw dataset
 - **1_stack_data:** use the cleaned dataset to create the main data structures for analysis
 - **2_descriptive:** run the descriptive figures and tables
 - **3_iv_regressions:** run the regressions for the main analysis (section 4 and 5) and its appendix
 - **4_stacked_regressions:** run the regression for the naive analysis (section 6) and its appendix
2. **Processed:** in the original reproducibility package (RP), it contains the base dataset “cc_datapacks” and the resulting datasets from processing it (by running 1_stack_data). In the published version of the RP, this folder is empty due to the impossibility of making the project data public. However, the original folder structure is kept. the subfolder “stacked” is meant to store the stacked datasets for the naive analysis
3. **Data:** contains the raw, alternative and auxiliary data files to be used in the analysis
4. **Figures, Tables:** are used to keep the figures and tables resulting from running the respective codes. These folders contain the outputs produced by the replicators who reviewed the reproducibility package.

Outputs

- **Figures:** All figures are named in the form “FigureX”, where “X” correspond to its respective position in the paper (i.e. Figure5, FigureB6)

- Tables: All tables are named in the form “TableX”, where “X” correspond to its respective position in the paper (i.e. Table5, TableB6)

Any output not starting with “Figure” or “Table” is used as an intermediary output to build the final one

Outputs from the main text

- Figures
 - figures/descriptive
 - * Figure 3
 - * Figure 4
 - * Figure 5
 - Panel A - `_tpres`
 - Panel B - `_satt`
 - Panel C - `_fac`
 - Panel D - `_math`
 - Panel E - `_eng`
 - Panel F - `_urdu`
- Tables
 - tables/descriptive/
 - * Table 1
 - * Table 2
 - tables/causal/
 - * Table 3
 - * Table 4
 - * Table 5

Outputs from Appendix A

- figures:
 - figures/descriptive/
 - FigureA1
 - Panel A - `_tpres`
 - Panel B - `_satt`
 - Panel C - `_fac`
 - Panel D - `_math`

- Panel E - _eng
- Panel F - _urdu
- FigureA3
- Panel A - _tpres
- Panel B - _satt
- Panel C - _fac

- tables:

- tables/descriptive/ - TableA1

Outputs from Appendix B

- figures:

- figures/causal/

- * FigureB1

- Panel A - __tpres
- Panel B - __satt
- Panel C - __fac
- Panel D - __math
- Panel E - __eng
- Panel F - __urdu

- * FigureB2

- Panel A - __tpres
- Panel B - __satt
- Panel C - __fac
- Panel D - __math
- Panel E - __eng
- Panel F - __urdu

- * Figure B3

- Panel A - __tpres
- Panel B - __satt
- Panel C - __fac
- Panel D - __math
- Panel E - __eng
- Panel F - __urdu

- * Figure B4

- Panel A - __tpres
- Panel B - __satt
- Panel C - __fac
- Panel D - __math
- Panel E - __eng

· Panel F - __urdu

- tables:
 - tables/descriptive/
 - TableB1
 - tables/causal/
 - TableB2
 - TableB3
 - TableB4
 - TableB5
 - TableB6
 - TableB7
 - TableB8