The Elusive Impact of Corporate Tax Incentives

Code Repository Documentation

Massimiliano Cali*
Giorgio Presidente†
Thiago Scot‡

August 2024

*World Bank
†Bocconi University
‡World Bank

Contents

1 Overview

The codes in this repository replicate the tables and figures from "The Elusive Impact of Corporate Tax Incentives", by Cali, Presidenti, and Scot. The replication folder contains the codes to go from the raw administrative data to the results in the paper.

This documentation is structured as follows. Section 2 describes the data sources and their availability. Section 3 describes the datasets used in the analysis. Section 4 provides details on the computational requirements. Section 5 provides instructions to replicators. Section 6 provides a mapping between the codes and the tables and figures of the paper.

2 Data Availability and Provenance Statements

2.1 Statement about Rights

- ☑ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

2.2 Summary of Availability

- ☐ All data **are** publicly available.
- ☑ Some data **cannot be made** publicly available.
- ☐ **No data can be made** publicly available.

The administrative data for this project is owned by the Tunisian National Institute of Statistics (INS). The data were made available to us exclusively for the purpose of this research project through collaboration agreements between the corresponding author and the government agency.

Researchers interested in obtaining the data for their own analyses can directly contact the National Institute of Statistics (INS)[1].

2.3 Details on each Data Source

The analysis is based on administrative records on Tunisian registred firms, the Repertoire National des Entreprises (RNE) provided by the Tunisian National Institute of Statistics (INS). The RNE micro data such as age, location, legal form, and economic activity has been merged with with tax returns to include yearly measures of declared revenue before tax, profits, and revenue from exports; Social security data on the total yearly employment and wage bill; and customs data on total value and quantity of exports and imports. All the merging has been done by the INS using identified data; the dataset provided to the researchers is already anonymized and contains information from all sources at the firm-year level. Data is available mainly from 1995 to 2022.

The *EU_ TN_ WLD.dta* dataset contains information on Tunisia's trade with EU countries by industry-codes and was generated using data from the World Integrated Trade Dataset (WITS).Researchers

[1] Ramzi Chanouf (chanouf.ramzi@ins.tn).

interested in accessing the data can access the WITS website and perform an advanced query on COMTRADE data, selecting the total imports from the European Union from the World (*EU_ WLD* variable in our code) and from Tunisia (*EU_ TUN*), for the period 2008-2020, by NAT code. The *Naccounts_ tunis.dta* dataset contains macroeconomic aggregates for Tunisia and it was generated using data from IMF Data Portal (consumption, GDP, government consumption and fixed capital formation) and Eurostat (exchange rate).

3 Datasets

Dataset	Description	Notes	Provided
dw2022v.dta	Combined anonymized administrative data with tax records, Social security and customs data (RNE)	Confidential	No
EU_TN_WLD.dta	EU Imports from Tunisia as well as from the world	Public	Yes
Naccounts_tunis.dta	Tunisia national accounts data	Public	Yes

4 Computational Requirements

- **Government Computer**
  - OS: Windows 11 Enterprise
  - Processor: Intel(R) Core(TM) i5-1145G7 CPU @ 2.60GHz
  - Memory available: 15.7 GB
  - Software version: Stata version 14.2
- **Replicators' computer: only script** `8_CIT_HonestDID.do`
  - OS: Windows 11 Enterprise
  - Processor: Intel(R) Core(TM) i5-1145G7 CPU @ 2.60GHz
  - Memory available: 15.7 GB
  - Software version: Stata version 18 MP

5 Instructions to replicators

5.1 Folder structure

For full replication of the project, we suggest the following, complete folder structure.

1_data
- 1_raw
- 2_intermediate
- 3_final

2_code

3_output
- 1_logs
- 2_figures
- 3_tables

This folder structure is reflected in the do-file `0_CIT_master.do`, which we provide in the code repository. We provide the do-files to create the datasets used in the project and the subsequent different analysis. The `0_CIT_master.do` do-file define *globals* in Stata, which correspond to the above folder tree. The code rely on these folder paths. Each data creation and analysis code calls `0_CIT_master.do` in the beginning of the do-file.

6 List of tables, figures and programs

The provided code reproduces:

☐ All numbers provided in text in the paper

☑ All tables and figures in the paper

☐ Selected tables and figures in the paper, as explained and justified below.

6.1 Mapping of tables and analysis code

Table 1 provides a mapping between all tables of the paper (including the online appendix) and the codes producing these results. The codes generate *.tex* files containing the content of the tables.

Table 1: Mapping of tables and analysis code

Table	**Code**	**File Name**
Table 1	3_CIT_descriptive_stats.do	table_descriptive1_updated.tex
Table 2	4_CIT_Industry_level.do	did_ind_lnw_prepost.tex
Table A1	*3*_CIT_descriptive_stats.do	*t*able_descriptive_balanced_updated.tex
Table A2	4_CIT_Industry_level.do	did_ind_lnw_prepost_contr.tex
Table A3	3_CIT_descriptive_stats.do	merged_aggregates.tex

6.2 Mapping of figures and analysis code

Table 3 provides a mapping between all figures of the paper (including the online appendix) and the code producing these results. The codes generate *.pdf* or *.png* files containing the graphs.

Table 2: Mapping of figures and analysis code

Figure	**Panels**	**Code**	**File Name**
Figure 1		Not generated by code	
Figure 2	A,B,C,D	3_CIT_descriptive_stats.do	**A**: nfirms_share.png **B**: employees_share.png **C**: revenue_share.png **D**: exportvalue_share.png
Figure 3	A,B,C	3_CIT_descriptive_stats.do	**A**: sector_comp_firms_offshore.png **B**: sector_comp_revenue_offshore.png **C**: sector_comp_emp_offshore.png
Figure 4	A,B,C,D,E,F	3_CIT_descriptive_stats.do	**A**: nfirms.png **B**: entry_d.png **C**: exit_d.png.png **D**: employees.png **E**: wagebill.png **F**: revenue_beforetax.png
Figure 5	A,B,C,D,E,F	4_CIT_Industry_level.do	**A**: did_ind_lnw_nfirms.png **B**: did_ind_lnw_entry_d.png **C**: did_ind_lnw_exit_d.png **D**: did_ind_lnw_employees.png **E**: did_ind_lnw_wagebill.png **F**: did_ind_lnw_revenue_beforetax.png
Figure 6	A,B,C,D	6_CIT_firm_level_regs	**A**: did_firm_ln_employees.png **B**: did_firm_employees_2_or_more.png **C**: did_firm_employees_5_or_more.png **D**: did_firm_employees_10_or_more.png
Figure 7	A,B,C	6_CIT_firm_level_regs	**A**: did_firm_ln_wb.png **B**: did_firm_ln_output.png **C**: did_firm_ln_prof_w.png

Table 3: Mapping of annexed figures and analysis code

Figure	Panels	Code	File Name
Figure A1		Not generated by code	
Figure A2	A,B	3_CIT_descriptive_stats.do	**A**: levels_active_firm3.png **B**: active_firm3.png
Figure A3	A,B,C,D,E,F	3_CIT_descriptive_stats.do	**A**: levels_nfirms.png **B**: levels_entry_d.png **C**: levels_exit_d.png **D**: levels_employees.png **E**: levels_wagebill.png **F**: levels_revenue_beforetax.png
Figure A4	A,B,C,D,E,F	4_CIT_Industry_level.do	**A**: did_ind_nfirms.png **B**: did_ind_entry_d.png **C**: did_ind_exit_d.png **D**: did_ind_employees.png **E**: did_ind_wagebill.png **F**: did_ind_revenue_beforetax.png
Figure A5	A,B,C,D,E,F	4_CIT_Industry_level.do	**A**: did_ind_w_nfirms.png **B**: did_ind_w_entry_d.png **C**: did_ind_w_exit_d.png **D**: did_ind_w_employees.png **E**: did_ind_w_wagebill.png **F**: did_ind_w_revenue_beforetax.png
Figure A6	A,B,C,D,E,F	4_CIT_Industry_level.do	**A**: did_ind_EUimp_nfirms.png **B**: did_ind_EUimp_entry_d.png **C**: did_ind_EUimp_exit_d.png **D**: did_ind_EUimp_employees.png **E**: did_ind_EUimp_wagebill.png **F**: did_ind_EUimp_revenue_beforetax.png
Figure A7	A,B,C,D,E,F	4_CIT_Industry_level.do	**A**: did_ind_dom_nfirms.png **B**: did_ind_dom_entry_d.png **C**: did_ind_dom_exit_d.png **D**: did_ind_dom_employees.png **E**: did_ind_dom_wagebill.png **F**: did_ind_dom_revenue_beforetax.png
Figure A8	A,B	4_CIT_Industry_level.do 6_CIT_firm_level_regs	**A**: did_ind_importvalue (1).png **B**: did_firm_ln_impv.png
Figure A9	A,B,C,D	6_CIT_firm_level_regs	**A**: did_firm_w_ln_employees.png **B**: did_firm_w_employees_2_or_more.png **C**: did_firm_w_employees_5_or_more.png **D**: did_firm_w_employees_10_or_more.png
Figure A10	A,B,C	6_CIT_firm_level_regs	**A**: did_firm_w_ln_output.png **B**: did_firm_w_ln_wb.png **C**: did_firm_w_ln_prof_w.png
Figure A11	A,B,C,D	6_CIT_firm_level_regs	**A**: did_firm_onforeign_ln_employees.png **B**: did_firm_onforeign_employees_2_or_more.png **C**: did_firm_onforeign_employees_5_or_more.png **D**: did_firm_onforeign_employees_10_or_more.png
Figure A12	A,B,C	6_CIT_firm_level_regs	**A**: did_firm_onforeign_ln_output.png **B**: did_firm_onforeign_ln_wb.png **C**: did_firm_onforeign_ln_prof_w.png
Figure A13	A,B,C,D	6_CIT_firm_level_regs	**A**: did_firm_exportonly_ln_employees.png **B**: did_firm_exportonly_employees_2_or_more.png **C**: did_firm_exportonly_employees_5_or_more.png **D**: did_firm_exportonly_employees_10_or_more.png

Continued on next page

Table 3 – continued from previous page

Figure	Panels	Code	File Name
Figure A14	A,B,C	6_CIT_firm_level_regs	**A**: did_firm_exportonly_ln_output.png **B**: did_firm_exportonly_ln_wb.png **C**: did_firm_exportonly_ln_prof_w.png
Figure A15		3_CIT_descriptive_stats.do	hist_enteryear.png
Figure A16	A,B,C,D	6_CIT_firm_level_regs	**A**: did_firm_immediate_ln_employees.png **B**: did_firm_immediate_employees_2_or_more.png **C**: did_firm_immediate_employees_5_or_more.png **D**: did_firm_immediate_employees_10_or_more.png
Figure A17	A,B,C	6_CIT_firm_level_regs	**A**: did_firm_immediate_ln_output.png **B**: did_firm_immediate_ln_wb.png **C**: did_firm_immediate_ln_prof_w.png
Figure A18	A,B,C,D	6_CIT_firm_level_regs	**A**: did_firm_ln_employees_regional.png **B**: did_firm_employees_2_or_more_regional.png **C**: did_firm_employees_5_or_more_regional.png **D**: did_firm_employees_10_or_more_regional.png
Figure A19	A,B,C	6_CIT_firm_level_regs	**A**: did_firm_ln_output_regional.png **B**: did_firm_ln_wb_regional.png **C**: did_firm_ln_prof_w_regional.png
Figure A20	A,B	7_CIT_firm_level_regs_PSM	**A**: hist_phat.png **B**: hist_phat_restrict.png
Figure A21	A,B,C,D,E	8_CIT_HonestDID	**A**: did_ind_lnw_nfirms.png **B**: honest_manuf_did_ind_nfirms.png **C**: honestlast_manuf_did_ind_nfirms.png **D**: honest_did_ind_nfirms.png **E**: honestlast_did_ind_nfirms.png
Figure A22	A,B,C,D,E	8_CIT_HonestDID	**A**: honest_did_ind_entry_d.png **B**: honest_did_ind_exit_d.png **C**: honest_did_ind_employees.png **D**: honest_did_ind_wagebill.png **E**: honest_did_ind_revenue_beforetax.png
Figure A23	A,B,C,D,E	8_CIT_HonestDID	**A**: honest_manuf_did_ind_entry_d.png **B**: honest_manuf_did_ind_exit_d.png **C**: honest_manuf_did_ind_employees.png **D**: honest_manuf_did_ind_wagebill.png **E**: honest_manuf_did_ind_revenue_beforetax.png

Note on Figure Generation: Figures A21, A22, and A23 were generated separately from the other exercises due to differences in the Stata versions and server restrictions at the INS. We have provided aggregate industry-level data, generated by 4_CIT_Industry_level.do do file, and the corresponding Stata code used to generate these figures in the replication package 8_CIT_HonestDID, ensuring they can be replicated independently outside government facilities.