The Effects of Transit Trade Systems on International Trade

Data Document

This document first describes the data used and how to get access to the data. Then, we describe the data package that goes along the paper, including: (i) the files included, (ii) the statistical softwares used, (iii) how to construct the estimating sample for the estimations, (iv) how to employ the codes to replicate the results, and (v) the variables definitions.

1. Data Sources and Access

1.1. Datasets

We employ the following datasets:

- Salvadorian Customs Data: Data provided by the official Salvadorian customs office (DGA). It consists of the entire universe of export transactions that originated in El Salvador from 2007 to 2013.
- TIM transits: Data provided by the Inter-American Development Bank (IADB). It consists of transaction-level records from the TIM information system and covers all shipments processed under TIM starting in El Salvador since its inception in 2011 to 2013.
- Firms registry: Data provided by the Inter-American Development Bank (IADB). It contains the municipality code where each firm is located.

1.2. Access

While, unfortunately, we cannot directly share the data used in our econometric analysis with third parties due to confidentiality agreement we had to sign with El Salvador's customs agency, we would like to highlight that similar transaction-level export data could have been obtained for the same country by other colleagues (see, for instance, the World Bank's Exporter Dynamic Database). In order to obtain these export data for El Salvador, you can send a formal request to the country's customs agency, Dirección General de Aduanas (DGA), to the following e-mail address: oficialinfo.dga@mh.gob.sv. As for the TIM data, you can contact The Central American Economic Integration Secretariat (SIECA) by filling in this web form.

2. Data Package

2.1. Files

The files included in the data package are:

- `data_generation.do`: It generates the proxy data to illustrate how to merge the different datasets, calculate the descriptive statistics, and run the regressions.
- `restat_submission_codes_match.do`: It carries out the match between the different datasets.
- `restat_desc_figures_tables.do`: It calculates and generates the tables and figures of descriptive statistics.

- `Tables_TIM_ReStat_Submission.R`: It calculates and generates the tables of OLS and IV regressions.

- `Event_Studies_ReStat_Submission.R`: It calculates and generates the figures of the event studies.

- `mun_ij_distances_long.dta`: It contains the distance among Salvadorian municipalities to calculate figures.

- `customs_municip_codes.dta`: It contains municipality codes for each internal custom in El Salvador.

2.2. Software

In order to run the previous scripts, we employed:

- Stata x64 17.0 for the do-files.

- R x64 4.1.2 for the R files.

2.3. Data Construction

The datasets share information that make possible to merge them.

The TIM transits dataset contains all shipments records through land that originated in El Salvador and finished in other Central American countries sharing the same information technology.[1] In particular, each record is a "Single Transport Document" (DUT for its Spanish acronym). A DUT includes the shipping company, the product codes (Harmonized

[1]Guatemala, Honduras, Nicaragua, Costa Rica and Panama

System at 8-digit level) of each product, the names of the firms exporting the products, the name of the foreign buyers, the country where the transit ends, the export value and weight, and the date of the transit, among other information.

The customs dataset includes all export transactions of Salvadorian firms. In particular, each record is a "Single Customs Document" (DUA for its Spanish acronym) . A DUA includes the exporting firm id and name, the product code (Harmonized System at 8-digit level), the importing country, the name of the foreign buyer, the export value and weight, and the date of the transaction, among other information. The registry data can be merged to this data by using the firm id.

The do-file `restat_submission_codes_match.do` undertakes the matching between the datasets using the shared fields as received. To do so, the do-file has a series of macros that have to be defined by using the variable names that identify the different variables. We provide the variable definitions for these macros below.

We also provide the do-file `data_generation.do` that generates proxy datasets with similar features as the original data to test the code.

Finally, the datasets `mun_ij_distances_long.dta` and `customs_municip_codes.dta` have to be in the same folder.

2.4. Replication of Results

To replicate descriptive and regression results, the dataset generated in the previous step needs to be in the same folder as defined on the scripts `restat_desc_figures_tables.do`, `Tables_TIM_ReStat_Submission.R`, and `Event_Studies_ReStat_Submission.R`.

Below there is a list of the tables and figures generated by each of them:

- restat_desc_figures_tables.do: Tables 1, and appendix table A1. Figure 3 and appendix figure A1.
- Tables_TIM_ReStat_Submission.R: Tables 2, 3, 4, 5, 6, 7 and appendix tables A2, A3, A4, A5 and A6. calculates and generates the tables of OLS and IV regressions.
- Event_Studies_ReStat_Submission.R: Figure 4 and appendix figures A2, A3 and A4.

2.5. Variable Definitions

- tim_v_dut: DUT code (TIM original variable).
- tim_vr_cou: starting country route section (TIM original variable).
- tim_vr_cou_end: country where transit ends (TIM original variable).
- tim_vr_cust: customs (TIM original variable).
- tim_vr_cust_int: internal custom were transit starts (TIM original variable).
- tim_vd_date: date in DD/MM/YYYY format (TIM original variable).
- tim_vd_name: firm names (TIM original variable).
- tim_vd_buyer: buyer names (TIM original variable).
- tim_vp_prod: product codes at 8-digit HS level (TIM original variable).

- `tim_vp_val`: transaction value in dollars (TIM original variable).
- `cus_v_idfirm`: firm id (customs original variable).
- `cus_v_name`: firm names (customs original variable).
- `cus_v_buyer`: buyer names (customs original variable).
- `cus_v_cou`: importing country (customs original variable).
- `cus_v_prod`: product names (customs original variable).
- `cus_v_val`: export value (customs original variable).
- `cus_v_wei`: export net weight (customs original variable).
- `cus_v_cust`: customs names (customs original variable).
- `cus_v_date`: date in DD/MM/YYYY format (customs original variable).
- `cus_v_trans`: transport mode (customs original variable).
- `reg_v_idfirm`: firm id (registry original variable).
- `reg_v_mun` : municipality (registry original variable).
- `exp_tfpca`: export value at firm-product-custom-destination level.
- `lexp_tfpca`: log export value at firm-product-custom-destination level.
- `lexp_tfca`: log export value at firm-custom-destination level.

- `wei_tfpca`: weight value at firm-product-custom-destination level.
- `lwei_tfpca`: log weight value at firm-product-custom-destination level.
- `nb_tfpca`: number of foreign buyers at firm-product-custom-destination level.
- `lnb_tfpca`: log number of foreign buyers at firm-product-custom-destination level.
- `nt_tfpca`: number of transactions at firm-product-custom-destination level.
- `lnt_tfpca`: log number of transactions at firm-product-custom-destination level.
- `p_tim_tfpca`: binary variable assigning a one if the firm used TIM at firm-product-custom-destination level.
- `p_tim_tfca`: binary variable assigning a one if the firm used TIM at firm-custom-destination level.
- `tim_tf`: binary variable assigning a one if the firm used TIM at firm level.
- `q_tim_tmca`: binary variable assigning a one if at least a firm in a municipality used TIM at custom-destination level.
- `q_tim_tdca`: binary variable assigning a one if at least a firm in a department used TIM at custom-destination level.
- `q_tim_tf`: binary variable assigning a one if at least a firm in a municipality used TIM.
- `munic`: code that identifies a municipality.

- `cod_dep`: code that identifies a department.
- `X`: log export value at region-custom-destination level, where region could either be municipality or department.
- `N_F`: log number of exporting firm at region-custom-destination level, where region could either be municipality or department.
- `N_FE`: log number of exporting firm-product combinations at region-custom-destination level, where region could either be municipality or department.
- `T`: binary variable assigning a one if the region used TIM at region-custom-destination level, where region could either be municipality or department.
- `NFExNF`: log number of exporting firm-product combinations per exporting firm at region-custom-destination level, where region could either be municipality or department.
- `XxNFE`: log average export value per firm-product combinations at region-custom-destination level, where region could either be municipality or department.
- `XxNF`: log average export value per firm at region-custom-destination level, where region could either be municipality or department.
- `MCA`: Municipality-custom-destination group variable.
- `DCA`: Department-custom-destination group variable.
- `FPCA`: Firm-product-custom-destination group variable.

- `PCT`: Product-destination-year group variable.
- `CT`: Destination-year group variable.
- `AT`: Custom-year group variable.
- `FT`: Firm-year group variable.
- `FP`: Firm-product group variable.
- `MT`: Municipality-year group variable.
- `MS`: Municipality-semester-year group variable.
- `seq`: year-semester group variable.
- `F1T`: One period forward value of the binary variable assigning a one if the region used TIM at region-custom-destination level, where region could either be municipality or department.
- `F2T`: Two periods forward value of the binary variable assigning a one if the region used TIM at region-custom-destination level, where region could either be municipality or department.
- `Xgr`: log export growth at region-custom-destination level, where region could either be municipality or department.
- `L1Xgr`: One period lagged value of log export growth at region-custom-destination level, where region could either be municipality or department.

- `L2Xgr`: Two period lagged value of log export growth at region-custom-destination level, where region could either be municipality or department.

- `F1p_tim_tfpca`: One period forward value of the binary variable assigning a one if the region used TIM at region-custom-destination level, where region could either be municipality or department.

- `F2p_tim_tfpca`: Two period forward value of the binary variable assigning a one if the region used TIM at region-custom-destination level, where region could either be municipality or department.

- `F1q_tim_tmca`: One period forward lag of the binary variable assigning a one if at least a firm in a municipality used TIM at custom-destination level.

- `F1q_tim_tdca`: One period forward lag of the binary variable assigning a one if at least a firm in a department used TIM at custom-destination level.

- `F2q_tim_tmca`: Two periods forward lag of the binary variable assigning a one if at least a firm in a municipality used TIM at custom-destination level.

- `F2q_tim_tdca`: Two periods forward lag of the binary variable assigning a one if at least a firm in a department used TIM at custom-destination level.

- `rfpc_a_tfp`: A binary indicator whether the firm-product-destination combination has positive exports through more than one custom.

- `rfpa_c_tfp`: A binary indicator whether the firm-product-custom combination has positive exports to more than one destination.

- `lrfpc_a_tfp`: One period lag of the binary indicator whether the firm-product-destination combination has positive exports through more than one custom.

- `lrfpa_c_tfp`: One period lag of the binary indicator whether the firm-product-custom combination has positive exports to more than one destination.

- `time_to_treatment`: Number of periods until TIM is available at the municipality-custom-destination level.

- `MaxTim`: A binary indicator capturing whether TIM is available at some point at the municipality-custom-destination level.