

Read Me: Trade Finance Use by Heterogeneous Firms

1. Overview

This document guides the replicator on how to reproduce the results. The package includes the necessary data sets and code to reproduce the results in the paper. The replicator only needs to change the first line of **00_master.do** according to their file path name. All files can be run from **00_master.do**.

2. Data Availability

SOME DATA CANNOT BE MADE PUBLICLY AVAILABLE (IN RED)

All raw data data sets used to construct the final data set are listed below in the path “\$base/rawdata/”.

File Name	Citation	Description	Source/notes	Data & Availability
vnm_2007-2016.dta	Vietnam Statistical Yearbook. Hanoi: GSO, 2024.	Vietnamese firm-level panel data covering the period 2007-2016	GSO Vietnam	Accessed: September 2024 Confidential (Redistribution Prohibited) Data acquired through the (previously named) Industrial Statistics Department of GSO
_expimp.dta	Vietnam Customs Authority. "Trade Statistics Database." 2024.	Export-import activity indicators and flags for matching with firm data	Vietnam Customs Authority	Accessed: October 2024 Confidential (Redistribution Prohibited) Data acquired through the (previously named) Industrial Statistics Department of GSO
correspondence_m49_iso3.dta	United Nations. "Standard Country or Area Codes for Statistical Use." UN Statistics Division, 2024.	Country correspondence file mapping UN M49 codes to ISO3 country codes	UN Statistics Division	Accessed: November 2024 Public
import_annual_no_est.dta	Vietnam Customs Authority. "Import Trade Statistics." 2024.	Annual import data by firm and product (HS 6-digit level) with aggregated values and transaction counts	Vietnam Customs Authority	Accessed: October 2024 Confidential (Redistribution Prohibited) Data acquired through the (previously named) Industrial Statistics Department of GSO
export_annual_no_est.dta	Vietnam Customs Authority. "Export Trade Statistics." 2024.	Annual export data by firm and product (HS 6-digit level) with aggregated values and transaction counts	Vietnam Customs Authority	Accessed: October 2024 Confidential (Redistribution Prohibited) Data acquired through the (previously named) Industrial Statistics Department of GSO
firms_panel_premerge.dta	Vietnam Ministry of Finance. "Enterprise Database." Hanoi, 2024.	Firm-level panel data in pre-merge format, containing financial and operational characteristics prior to final	GSO Vietnam	Accessed: October 2024 Confidential (Redistribution Prohibited) Data acquired through the

		data integration, covering the period 2016-2022		(previously named) Industrial Statistics Department of GSO
WBank_rule_of_Law.xlsx	World Bank. "Worldwide Governance Indicators: Rule of Law." World Bank, 2024. https://www.worldbank.org/en/publication/worldwide-governance-indicators	Rule of Law index by country and year, measuring the extent to which agents have confidence in and abide by rules of society	World Bank Worldwide Governance Indicators	Accessed: November 2024 Public
survey_ready_2023-08-22	IFC. Trade Finance Firm Level Survey -- Viet Nam 2019-2021. Companion survey to the IFC and WTO. 2023 report. "Trade Finance in the Mekong region: A Study of Cambodia, the Lao People's Democratic Republic and Viet Nam."	Firm survey data collected by the IFC for the report "Trade Finance in the Mekong Region"	IFC	Accessed: August 2022 Confidential (Redistribution Prohibited)
unsd_migration_stock_raw.xlsx	United Nations Department of Economic and Social Affairs (UNSD). "International Migrant Stock." UN Statistics Division, 2024.	International migration stock by country and year (both origin and destination), used as proxy for diaspora networks and information frictions	UN Department of Economic and Social Affairs	Accessed: November 2024 Public
/prices¹				
Annual average CPI (Previous year = 100).csv	General Statistics Office of Vietnam. "Statistical Data: Consumer Price Index." Hanoi: GSO, 2024.	Annual average consumer price index, gold, USD price index and core inflation (Previous year = 100)	NSO prices	Accessed: November 2024 Public
Producer price index for agriculture, forestry and fishing (Year 2010 = 100) by Items and Year.csv	General Statistics Office of Vietnam. "Statistical Data: Producer Price Index." Hanoi: GSO, 2024.	Producer price index for agriculture, forestry and fishing (Year 2010 = 100)	NSO prices	Accessed: November 2024 Public
Producer price index for industry (Year 2010 = 100)	General Statistics Office of Vietnam. "Statistical Data: Producer Price Index." Hanoi: GSO, 2024.	Producer price index for industry (Year 2010 = 100)	NSO prices	Accessed: November 2024 Public

¹ Link to specific series may require VPN VN connection to access:

<https://www.nso.gov.vn/px-web-2/?pxid=V1114&theme=Th%C6%B0%C6%A1ng%20m%E1%BA%A1i%2C%20gi%C3%A1%20c%E1%BA%A3>

<https://www.nso.gov.vn/px-web-2/?pxid=V1127&theme=Th%C6%B0%C6%A1ng%20m%E1%BA%A1i%2C%20gi%C3%A1%20c%E1%BA%A3>

<https://www.nso.gov.vn/px-web-2/?pxid=V1129&theme=Th%C6%B0%C6%A1ng%20m%E1%BA%A1i%2C%20gi%C3%A1%20c%E1%BA%A3>

<https://www.nso.gov.vn/px-web-2/?pxid=V1131&theme=Th%C6%B0%C6%A1ng%20m%E1%BA%A1i%2C%20gi%C3%A1%20c%E1%BA%A3>

<https://www.nso.gov.vn/px-web-2/?pxid=V1133&theme=Th%C6%B0%C6%A1ng%20m%E1%BA%A1i%2C%20gi%C3%A1%20c%E1%BA%A3>

<https://www.nso.gov.vn/px-web-2/?pxid=V1135&theme=Th%C6%B0%C6%A1ng%20m%E1%BA%A1i%2C%20gi%C3%A1%20c%E1%BA%A3>

by Industrial products and Year.csv				
Producer price index for services (Year 2010 = 100) by Items and Year.csv	General Statistics Office of Vietnam. “Statistical Data: Producer Price Index.” Hanoi: GSO, 2024.	Producer price index for services (Year 2010 = 100)	NSO prices	Accessed: November 2024 Public
export_price_index_2023.xlsx	General Statistics Office of Vietnam. “Statistical Data: Export Price Index.” Hanoi: GSO, 2024.	Merchandise Export Price Index (USD, year 2010 = 100)	NSO prices	Accessed: November 2024 Public
import_price_index_2023.xlsx	General Statistics Office of Vietnam. “Statistical Data: Import Price Index.” Hanoi: GSO, 2024.	Merchandise Import Price Index (USD, Year 2010 = 100)	NSO prices	Accessed: November 2024 Public
sector_correspondence.xlsx	Constructed by authors based on GSO sector classifications	Concordance between PPI sectors, IPI sectors, GDP sectors, and 2-digit VSIC codes	Authors	Accessed: November 2024 Public
/sector				
oecd_prod_valu_vnm.xlsx	OECD. “TiVA 2023 Edition: Principal Indicators.” OECD.Stat, 2023.	Production (gross output) and value added by economic activity for Vietnam, 2019–2020	OECD.Stat	Accessed: October 2024 Public
oecd_prod_valu_world.xlsx	OECD. “TiVA 2023 Edition: Principal Indicators.” OECD.Stat, 2023.	Production (gross output) and value added by economic activity for World, 2019–2020	OECD.Stat	Accessed: October 2024 Public
value_chain_length_ADB_2022.xlsx	RIGVC UIBE, 2016, UIBE GVC Indicators, http://rigvc.uibe.edu.cn/english/D_E/database_database/index.htm	Value chain length indicators (PLv, PLy, APL_f, APL_b) by sector and region for Vietnam, 2000–2022. Based on Asian Development Bank. “Multiregional Input-Output Database.”	UIBE GVC database	Accessed: October 2024 Public
/spatial				
VNM_adm3.shp, VNM_adm3.shx, VNM_adm3.dbf	Viet Nam administrative boundary, Geodatabase	Commune-level shapefile data (2,691 communes) including geographic coordinates and administrative identifiers; 6 columns with commune ID, names, and spatial references	GADM Project	Accessed: July 2024 Public
Admin_level3_codes_01_07_2024.xls	GSO, “Administrative units”	List of names and statistical codes of Vietnam’s administrative units as of 2024	https://www.nso.gov.vn/p-huong-phap-luan-thong-ke/danh-muc/don-vi-hanh-chinh/	Accessed: July 2024 Public
seaports.xlsx	Marine insights, 5 major ports of Vietnam	Vietnam seaport locations with geographic coordinates (latitude/longitude); includes primary and secondary port data with 5 columns (36 records across sheets)	https://www.marineinsight.com/knowledge/5-major-ports-of-vietnam/	Accessed: July 2024 Public
sim_dis_shp3_new.xlsx	Authors’ construction (see below)	District name matching between VN admin and GADM GIS files with similarity score		

sim_dis_shp3_cor.xlsx	Authors' construction (see below)	District name matching between VN admin and GADM GIS files with similarity score and manual review decision for a subset of districts (whose names don't match exactly)	Manually review decision for districts that don't match exactly: accept or reject (0/1) name matching based on subjective judgement	
to_merge_unique_wtgeo_2_new.xlsx	Authors' construction (see below)	List of communes with GIS codes and estimated distance to the closest major port		

3. Instructions for Replicators

Python processing (produces inputs into 0_2_prep_spatial.do below)

i) *_seaport_1_match.py*

The Python script matches district from the official VN admin file to GADM commune polygons (VNM_adm3.shp). For districts that don't match exactly, writes a fuzzy similarity table (match candidates with highest similarity scores to "sim_dis_shp3_new.xlsx" for human review. The manually reviewed file should be saved under a new name ("sim_dis_shp3_cor.xlsx" in the current pipeline)

ii) *_seaport_2_compute.py*

The Python script takes all exact match district names plus manually reviewed matches from the previous step ("sim_dis_shp3_cor.xlsx") to match commune names. Communes are matched within districts and include a mix of exact and fuzzy matches. Then calculate distances from each matched commune's centroid to seaports.

Stata processing

Stata cleaning

iii) *0_master*

This sets the global file paths and calls all cleaning and analysis do files in sequence. The 2 exceptions are the python scripts that will need to be run separately before *0_2_prep_spatial.do*

iv) *0_master_cleaning*

This runs the full cleaning pipeline from raw data through to the two final analysis datasets (census_with_custom.dta and customs_census_analysis_no_est.dta). The 2 exceptions are the python scripts that will need to be run separately before *0_2_prep_spatial.do*

v) *0_0_prep_supplementary_data*

This imports raw price, sector, and geographic correspondence files and saves them as Stata intermediates (CPI, PPI, OECD TiVA, value chain length) for use by the firm cleaning script.

vi) *0_2_prep_spatial*

This imports the Python-generated commune-level distance output and saves a cleaned Stata dataset with each commune's minimum distance to the five major Vietnamese seaports.

vii) *1_0_cleaning_firms*

This cleans the firm-level panel by removing outliers, deflating financial variables, and deriving key analysis variables (age, FDI status, productivity, borrowings), producing the cleaned firm–year datasets.

viii) 1_1_cleaning_past

This builds the pre-2016 import and export activity dataset used to calculate cumulative years of trading experience for each firm.

ix) _productivity

This is a helper script called by 1_0_cleaning_firms.do that computes labor productivity (log revenue per employee) and ACF total factor productivity by 2-digit sector.

x) 1_2_cleaning_rule_of_law

This imports and cleans the World Bank Worldwide Governance Indicators – Rule of Law, producing a country–year panel with RoL estimates and within-year quartile splits for merging into the analysis dataset.

xi) 1_3_migration_stock

This builds the Vietnamese diaspora share by destination country using UN migration stock data, producing a country-level dataset used as a proxy for information frictions in export markets.

xii) 1_4_build_census_with_custom

This builds the firm–year panel (census_with_custom.dta) by merging firm characteristics with customs summary statistics and port distance, used by the descriptives and interest rate scripts.

xiii) 1_5_mills_ratio

This estimates Heckman selection models for import and export LC use and saves the Inverse Mills Ratios used for the selection correction in Tables 3 and 4, as well as a convenience file with state ownership status.

xiv) 1_6_build_customs_analysis

This builds the main transaction-level analysis dataset (customs_census_analysis_no_est.dta) by merging annual customs records with firm characteristics, deflators, past trade experience, Rule of Law, diaspora shares, and Inverse Mills Ratios.

xv) _exchange_rate

This is a helper script called by 1_6_build_customs_analysis.do that assigns annual average VND/USD exchange rates for 2016–2023, sourced from the State Bank of Vietnam.

Stata analysis

xvi) 0_master_analysis

This sets the global file paths and calls all analysis do files in sequence.

xvii) tab1_descriptives

This produces Table 1, presenting descriptive statistics for all firms, importers, importers using LCs, exporters, and exporters using LCs.

xviii) fig2_lc_shares

This produces Figure 2, showing the share of trade flows settled in local currency over time, separately for imports, exports, and total trade.

xix) fig3_kernels

This produces Figure 3, displaying kernel density plots comparing the distribution of firm size, age, and prior trade experience across non-traders, traders, and LC users, separately for imports and exports.

xx) tab2_anova

This produces Table 2, decomposing the variance in LC use across firm, product, partner, and year fixed effects using three ANOVA model specifications.

xxi) tab3_imports

This produces Table 3, estimating the determinants of LC use in imports via linear probability models with product-partner-year, sector-year, and province-year fixed effects.

xxii) tab4_exports

This produces Table 4, estimating the determinants of LC use in exports via linear probability models with product-partner-year, sector-year, and province-year fixed effects.

xxiii) tab5_standardized

This produces Table 5, re-estimating the import and export LC use regressions with all continuous regressors standardized to allow direct comparison of coefficient magnitudes.

xxiv) tab6_fig4_rol_diaspora

This produces Figure 4 and Table 6, examining how rule of law and diaspora networks in the partner country interact with firm characteristics to affect LC use in imports and exports.

xxv) tab7_lc_share

This produces Table 7, providing a robustness check by replacing the binary LC use indicator with the continuous LC share of trade value as the dependent variable in the main import and export regressions.

xxvi) fig56_interest_rate

This produces Figures 5 and 6, analyzing the relationship between firm characteristics and the interest rate proxy (interest expense over average borrowings), with binscatter plots and coefficient plots from regressions on importers.

Mapping which do file reproduces the figures and tables in the manuscript.

Do file	Output in the manuscript
fig2_lc_shares	Figure 2
fig3_kernels	Figure 3
tab6_fig4_rol_diaspora	Figure 4
fig56_interest_rate	Figure 5 Figure 6
tab1_descriptives	Table 1
tab2_anova	Table 2
tab3_imports	Table 3
tab4_exports	Table 4
tab5_standardized	Table 5

tab6_fig4_rol_diaspora	Table 6
tab7_lc_share	Table 7

4. Requirements

The package was run on the following:

- Dell Pro 14 Premium
- OS: Windows
- STATA MP/19

In STATA, ensure the following packages are installed:

- binscatter2
- center
- estout
- ftools
- ivreg2
- ivreghdfe
- ranktest
- reghdfe

Run time: 3 hours (cleaning + analysis)

Storage requirements: 6 GB