

The (fiscal) Dividend of Infrastructure: Roads and Revenues in Rwanda

This README was prepared following the WBG Reproducible Research Repository guidelines for World Bank Policy Research Working Paper submissions.

Overview

Field	Details
Paper Title	The (fiscal) Dividend of Infrastructure: Roads and Revenues in Rwanda
Authors	Abdou Musonera (University of Rwanda), Aimable Nsabimana (DECPM – World Bank Group), Daniel Overbeck (National University of Singapore)
Contact Author	Aimable Nsabimana : ansabimana1@ifc.org
Date	April 2026
Journal/Series	World Bank Policy Research Working Paper
JEL Codes	H20, H71, O18, R42
Keywords	Infrastructure, Economic Development, Taxation, Roads, Firms

Abstract

This paper demonstrates that infrastructure investments, specifically road upgrades in Rwanda, enhance fiscal capacity. Using a novel dataset of road upgrades and a decade of tax and establishment census records, the authors find significant increases in tax revenues in municipalities near upgraded roads. This growth is driven by firm entry and land value appreciation (captured via rental income taxes). While these revenues do not fully recover the central government's initial investment, local municipalities' revenues double within six years following the upgrade.

Data Availability and Origins

Statement on Rights: The data used in this paper were obtained under data-sharing agreements with Rwandan government institutions. The microdata are confidential and cannot be publicly released. Researchers wishing to access the data should contact the respective institutions directly.

Table 1: Data Sources

#	Dataset	Source Institution	Years Covered	Access data
1	Administrative Tax Records (central and local taxes: CIT, PIT, VAT, rental income tax, trade license fees)	Rwanda Revenue Authority (RRA): https://www.rra.gov.rw/en/home	2011–2022	June 2024-Apr2026
2	Rwandan Establishment Censuses	National Institute of Statistics of Rwanda (NISR): http://www.statistics.gov.rw/	2011, 2014, 2017, 2019	June 2024-Apr2026
3	Road Network and Construction Data (location, timing, nature, and costs of upgrades)	Rwanda Transport Development Agency (RTDA): https://www.rtda.gov.rw/home	2011–2024	June 2024-Apr2026
4	Road Construction Budget Documents	Ministry of Finance and Economic Planning (MINECOFIN): https://www.minecofin.gov.rw/	2011–2024	June 2024-Apr2026

Note: No data files are included in this repository due to confidentiality restrictions. The code is provided for transparency and reproducibility conditional on access to the underlying datasets.

Software Requirements

- **Stata** (version 17 or higher recommended). The following user-written packages are required. Install them before running the code:

```
ssc install eventstudyinteract, replace
ssc install event_plot, replace
ssc install reclink2, replace
ssc install reghdfe, replace
ssc install ftools, replace
ssc install estout, replace
```

- `eventstudyinteract` and `event_plot` are used for all staggered difference-in-differences event study estimations (Sun & Abraham 2021 interaction-weighted estimator).
- `reclink2` is used for probabilistic fuzzy matching of district/sector names between the tax records and the road sector ID dataset.
- `reghdfe` and `ftools` are used for high-dimensional fixed effects regressions.
- **GIS software (ArcGIS):** used upstream to produce the sector-level spatial buffer files (`sector_buffered.xlsx`). These files are inputs to the Stata code and are not regenerated by the `dofiles`.

Repository Structure

File/Folder	Description
roads_tax_returns/	Root folder (global path)
-- README.md	This file
-- data/	
-- roads/	
-- Buffer_sector_roads/	Raw data (NOT included -- see Data Availability)
-- sector_buffered.xlsx	GIS-generated sector buffers
-- multiple_treats.dta	Multiply-treated sector
flags	
-- elevation_pop.dta	Elevation and population
controls	
`-- noNR.dta	Highway-only sector
identifier	
-- 1_Road_Database_Excel_Nov_2024/	
`-- Roads_and_Sector_Code.xlsx	Road upgrade database (RTDA)
-- unique_sector_ID_final_2.dta	Sector ID crosswalk
-- Roads_and_Sector_Code_and_treatment.dta	Treatment assignment (paved roads)
-- Roads_and_Sector_Code_and_treatment_gravel.dta	Treatment (unpaved upgrades)
`-- Roads_and_Sector_Code_and_treatment_random_inter.dta	Placebo
treatment	
-- firms/	
`-- Rwanda_firm_sales_taxes2012_2024.dta	RRA firm-level tax records (NOT included)
-- use/	
-- final_Registration_with_entreprise_status.dta	Firm registration data (NOT included)

		`-- elevation_pop.dta	Elevation and population
data (used in Fig A2)			
		-- temp/	
		`-- Rwanda_cpi_2009_2024.dta	CPI deflator
		-- final/	Analysis-ready datasets
		-- Roads_and_Tax_combined_cleaned.dta	Cleaned firm-level panel
		-- Roads_Tax_combined_cleaned_Aug.dta	Main analysis dataset (paved
roads)			
		-- Roads_Tax_combined_cleaned_Sep_gravel.dta	Analysis dataset (unpaved
upgrades)			
		-- Roads_Tax_combined_cleaned_Sep_random_inter.dta	Analysis dataset
(placebo)			
		-- liscence_tax_Aug2025.dta	Local trade license tax
		-- rental_tax_Aug2025.dta	Local rental income tax
		`-- collapsed_cleaned_panel_establishment_Aug2025.dta	Establishment
census panel			
		-- randomtime_placebos/	
		-- placebo_round1.dta	Stacked placebo results (200
rounds)			
		-- placebo_round_inter.dta	Intermediate placebo results
per iteration			
		-- WB WP submission/	Replication package folder
		-- 0_Master_dofile_May2026.do	Master dofile (runs all
scripts in sequence)			
		-- 1_Analysis_May2026.do	Main event study -- full
sample			
		-- 2_Analysis_robustness_May2026.do	Robustness checks (8
specifications, Fig A9)			
		-- 3_Analysis_Gravel_May2026.do	Unpaved upgrade placebo (Fig
A5)			
		-- 4_Analysis_Placebo_RandomTiming_May2026.do	Random timing placebo (200
iterations)			
		-- 5_Analysis_Fi_A2_description_May2026.do	Descriptive histograms (Fig
A2)			
		-- logfiles/	
		-- analysis.log	
		-- robustness.log	
		-- upgrated_without_paving.log	
		-- placeboRandom.log	
		-- FigA2_description.log	
		-- _results2/_results/_mainResults/	
		-- rentalSep25.pdf	Rental income tax
		-- licenseSep25.pdf	Trade license fees
		-- localSep25.pdf	Total local taxes
		-- centraltotalSep25.pdf	Total central taxes
		-- citSep25.pdf	Corporate Income Tax
		-- vatSep25.pdf	Value Added Tax
		-- pitSep25.pdf	Personal Income Tax
		-- entrySep25.pdf	Total firm entry
		-- formalentrySep25.pdf	Formal firm entry
		-- informalentrySep25.pdf	Informal firm entry
		-- femaleentrySep25.pdf	Female-owned firm entry
		-- maleentrySep25.pdf	Male-owned firm entry
		-- smallentrySep25.pdf	Small firm entry
		-- mediumentrySep25.pdf	Medium-size firm entry

	-- largeentrySep25.pdf	Large firm entry
	-- robustness_FigA9/	Robustness figures (8
panels)		
	-- local_a.pdf / central_b.pdf	Highway municipalities only
	-- local_c.pdf / central_d.pdf	All municipalities
	-- local_e.pdf / central_f.pdf	With population and
elevation controls		
	-- local_g.pdf / central_h.pdf	Excl. Kigali and multiply-
treated		
	-- without_paved_FigA5/	Unpaved upgrade placebo
	-- gravellocalSep25.pdf	
	-- gravelcentralSep25.pdf	
	-- Placebo/	Random timing placebo
	-- PlaceboRandomSep25.pdf	
	`-- FigA2/	Descriptive tax revenue
distributions		
	-- FigA2_before.pdf	Before road upgrade
	`-- FigA2_after.pdf	After road upgrade

Instructions to Replicators

1. **Obtain data access.** Request access to the restricted datasets from the Rwanda Revenue Authority (RRA), the National Institute of Statistics of Rwanda (NISR), and the Rwanda Transport Development Agency (RTDA). The CPI deflator data should be obtained from the NISR or the National Bank of Rwanda.
2. **Install required Stata packages.** Run the installation block at the top of `0_Master_dofile_May2026.do` (or install individually as listed in the Software Requirements section above).
3. **Set file paths.** Open `0_Master_dofile_May2026.do` and update the `global root` path to point to the root folder on your local machine. All other paths are defined relative to `root` and do not need to be changed.
4. **Place raw data files** in the correct subfolders following the Repository Structure section above. File names must match exactly as referenced in the dofiles.
5. **Run `0_Master_dofile_May2026.do`** in Stata. It will execute all four dofiles in sequence: main analysis, robustness checks, unpaved upgrade placebo, and random timing placebo.
6. All figures (PDF) will be saved in `WB WP submission/_results2/_results/_mainResults/` and its subfolders. Placebo dataset results will be saved in `data/randomtime_placebos/`.

(i) On the fuzzy match step: The data cleaning dofile uses `reclink2` to match district and sector names between the tax records and the road sector ID crosswalk. The match threshold is set at `match_score > 0.9`. Replicators should inspect the match quality and may need to adjust this threshold depending on the exact version of the input files.

(ii) On the placebo loop: Dofile `4_Analysis_Placebo_RandomTiming_May2026.do` runs 200 iterations with randomly assigned treatment timing. Seeds are set as `set seed 12345` before iteration 1, and `set seed i` at the start of each subsequent iteration `i`, ensuring full reproducibility. An initial empty seed file (`placebo_round1.dta`) must exist before the loop starts — the master dofile handles this automatically.

(iii) Estimated runtime: The placebo loop (200 iterations) is computationally intensive. Estimated total runtime on a standard laptop: 1–2 hours.

List of Figures

All event study figures are produced using the `eventstudyinteract + event_plot` commands (Sun & Abraham 2021 interaction-weighted estimator) with sector and year fixed effects, and standard errors clustered at the district level.

Table 2: Main Results (1_Analysis_May2026.do)

Figure	Output File	Outcome Variable
Event study: Rental income tax	rentalSep25.pdf	log(total_rental)
Event study: Trade license fees	licenseSep25.pdf	log(total_lisence)
Event study: Total local taxes	localSep25.pdf	log(total_lisence + total_rental)
Event study: Total central taxes	centraltotalSep25.pdf	log(total_nat)
Event study: CIT	citSep25.pdf	log(total_CIT)
Event study: VAT	vatSep25.pdf	log(total_VAT)
Event study: PIT	pitSep25.pdf	log(total_PIT)
Event study: Total firm entry	entrySep25.pdf	log(no entrants)
Event study: Formal firm entry	formalentrySep25.pdf	log(formal entrants)
Event study: Informal firm entry	informalentrySep25.pdf	log(informal entrants)
Event study: Female-owned firm entry	femaleentrySep25.pdf	log(female entrants)
Event study: Male-owned firm entry	maleentrySep25.pdf	log(male entrants)
Event study: Small firm entry	smallentrySep25.pdf	log(small entrants)
Event study: Medium-size firm entry	mediumentrySep25.pdf	log(medium entrants)
Event study: Large firm entry	largeentrySep25.pdf	log(large entrants)

Robustness (2_Analysis_robustness_May2026.do -- Figure A9)

Eight specifications (panels a--h): highway municipalities only (a, b), all municipalities (c, d), with population and elevation controls (e, f), and excluding Kigali and multiply-treated sectors (g, h). Output saved to `robustness_FigA9/`.

Table 3: Unpaved Upgrade Placebo (3_Analysis_Gravel_May2026.do -- Figure A5)

Figure	Output File	Outcome Variable
Unpaved upgrade: Local taxes	gravellocalSep25.pdf	log(total_lisence + total_rental)
Unpaved upgrade: Central taxes	gravelcentralSep25.pdf	log(total_nat)

Table 4: Randomised Treatment Timing Placebo (4_Analysis_Placebo_RandomTiming_May2026.do)

Output	Description
placebo_round1.dta	Stacked results from 200 placebo iterations with randomly assigned treatment years (uniform draw from 2004-2024).
PlaceboRandomSep25.pdf	Scatter plot of placebo coefficients by relative time (<code>twoway scatter real_coef number</code>).

Table 5: Descriptive Figure -- Tax Revenue Distributions**(5_Analysis_Fi_A2_description_May2026.do -- Figure A2)**

This dofile produces overlapping histograms of log(total tax revenues) comparing treated and control municipalities separately for the pre-treatment and post-treatment periods. Total taxes are defined as the sum of central taxes, local trade license fees, and local rental income taxes.

Figure	Output File	Description
Fig A2 -- Before upgrade	FigA2_before.pdf	Distribution of log(total tax revenues) for treated vs. control municipalities before road upgrade
Fig A2 -- After upgrade	FigA2_after.pdf	Distribution of log(total tax revenues) for treated vs. control municipalities after road upgrade

Master Dofile (0_Master_dofile_May2026.do)

The master dofile is calibrated to the exact structure and globals used in the four dofiles. Replicators only need to update **global root**.

```

/*=====
PROJECT:  The (fiscal) Dividend of Infrastructure:
          Roads and Revenues in Rwanda
AUTHORS:  Abdou Musonera, Aimable Nsabimana, Daniel Overbeck
DATE:     April 2026
PURPOSE:  Master dofile

INSTRUCTIONS:
1. The data sharing agreements are with the authors and institutions.
2. The authors do not have any right to share the data used in this study.
3. However, authors can share temporary access for any verification.
=====*/

clear all
set more off
set maxvar 10000
version 17

global root      "C:/Users/[your_username]/roads_tax_returns"

* Derived globals
global path      "${root}"
global data      "${root}/data/roads/Buffer_sector_roads"
global logf      "${root}/WB WP submission/logfiles"
global code      "${root}/WB WP submission"
global analysis  "${root}/WB WP submission/_results2/_results/_mainResults"
global robustn   "${root}/WB WP submission/_results2/_results/_mainResults"
global wtht_paved "${root}/WB WP submission/_results2/_results/_mainResults/without_paved_FigA5"
global placeboRdm "${root}/WB WP submission/_results2/_results/_mainResults/Placebo"
global figA2     "${root}/WB WP submission/_results2/_results/_mainResults/FigA2"

* Create output folders if they do not already exist
cap mkdir "${root}/WB WP submission/logfiles"
cap mkdir "${root}/WB WP submission/_results2"
cap mkdir "${root}/WB WP submission/_results2/_results"
cap mkdir "${root}/WB WP submission/_results2/_results/_mainResults"
cap mkdir "${root}/WB WP submission/_results2/_results/_mainResults/robustness_FigA9"
cap mkdir "${root}/WB WP submission/_results2/_results/_mainResults/without_paved_FigA5"

```

```

cap mkdir "${root}/WB WP submission/_results2/_results/_mainResults/Placebo"
cap mkdir "${root}/WB WP submission/_results2/_results/_mainResults/FigA2"
cap mkdir "${root}/data/randomtime_placebos"

*-----
* 2. INSTALL REQUIRED PACKAGES (run once -- comment out after first install)
*-----

* ssc install eventstudyinteract, replace // Sun & Abraham (2021) estimator
* ssc install event_plot, replace // Event study plots
* ssc install reclink2, replace // Probabilistic fuzzy matching
* ssc install reghdfe, replace // High-dimensional FE regressions
* ssc install ftools, replace // Dependency for reghdfe
* ssc install estout, replace // Table export
* ArcGIS was used for maps

*-----
* 3. MAIN EVENT STUDY ANALYSIS -- FULL SAMPLE
* Dofile : 1_Analysis_May2026.do
* Input : data/final/Roads_Tax_combined_cleaned_Aug.dta
* Output : _results2/_results/_mainResults/ (15 figures)
* Log : logfiles/analysis.log
*-----

do "${code}/1_Analysis_May2026.do"

di "Step 3 complete: Main event study figures produced."

*-----
* 4. ROBUSTNESS CHECKS
* Dofile : 2_Analysis_robustness_May2026.do
* Inputs : Roads_Tax_combined_cleaned_Aug.dta
* multiple_treatments.dta / elevation_pop.dta / noNR.dta
* Output : robustness_FigA9/ (8 panels, a-h)
* Log : logfiles/robustness.log
*-----

do "${code}/2_Analysis_robustness_May2026.do"

di "Step 4 complete: Robustness figures produced."

*-----
* 5. UNPAVED UPGRADE PLACEBO
* Dofile : 3_Analysis_Gravel_May2026.do
* Input : Roads_and_Tax_combined_cleaned.dta
* Output : without_paved_FigA5/ (2 figures)
* Log : logfiles/upgraded_without_paving.log
*-----

do "${code}/3_Analysis_Gravel_May2026.do"

di "Step 5 complete: Unpaved upgrade placebo figures produced."

*-----
* 6. PLACEBO TEST -- RANDOMISED TREATMENT TIMING (200 iterations)
* Dofile : 4_Analysis_Placebo_RandomTiming_May2026.do
* Seeds : set seed 12345 (iteration 1); set seed i (iterations 2-200)
* Output : data/randomtime_placebos/placebo_round1.dta
* Placebo/PlaceboRandomSep25.pdf

```

```

* Log      : logfiles/placeboRandom.log
* NOTE     : Computationally intensive (~1-2 hours). Comment out if only
*            replicating main results.
*-----

clear
gen number      = .
gen real_coef   = .
gen ci_low      = .
gen ci_up       = .
gen significant = .
save "${root}/data/randomtime_placebos/placebo_round1.dta", replace

do "${code}/4_Analysis_Placebo_RandomTiming_May2026.do"

di "Step 6 complete: Random timing placebo done."

*-----
* 7. DESCRIPTIVE FIGURE -- TAX REVENUE DISTRIBUTIONS (Figure A2)
* Dofile   : 5_Analysis_Fi_A2_description_May2026.do
* Inputs   : data/final/Roads_Tax_combined_cleaned_Aug.dta
*           data/use/elevation_pop.dta
* Output   : FigA2/FigA2_before.pdf -- Pre-upgrade tax distribution
*           FigA2/FigA2_after.pdf  -- Post-upgrade tax distribution
*-----

do "${code}/5_Analysis_Fi_A2_description_May2026.do"

di "Step 7 complete: Descriptive Figure A2 produced."

*-----
* ALL DONE
*-----

di " "
di "===== "
di " REPLICATION COMPLETE."
di " Main figures   : ${analysis}/"
di " Robustness     : ${analysis}/robustness_FigA9/"
di " Gravel placebo : ${wht_paved}/"
di " Random placebo : ${placeboRdm}/"
di " Figure A2      : ${figA2}/"
di " Log files      : ${logf}/"
di "===== "

```

Maps and GIS Figures (ArcGIS)

The following figures were produced using **ArcGIS** and are not generated by the Stata dofiles. They rely on shapefiles provided by the National Institute of Statistics of Rwanda (NISR) and road network data from the Rwanda Transport Development Agency (RTDA). These figures are included in the paper as static outputs and cannot be reproduced from the Stata code alone.

Table 5: Maps produced in this study

Figure	Title	Description	Data Sources
Figure 1	National roads and treated municipalities in Rwanda	Map showing the spatial distribution of national paved roads across treated municipalities versus those that did not (control) across Rwanda.	Road network and upgrade data: RTDA. Administrative boundaries and municipality shapefiles: NISR.

Figure A6	Total number of firms across municipalities	Map showing the number of all establishments across municipalities in Rwanda in 2011 and 2019 , illustrating the spatial distribution of economic activities	Establishment census data (2011 and 2019): NISR. Administrative municipality shapefiles: NISR.
Figure A8	National upgraded and non-upgraded roads	Map of national roads that were upgraded (paved) during the study period and those that remained unpaved, overlaid on the Rwanda road network.	Road network and upgrade status data: RTDA. Administrative boundaries: NISR.

Shapefiles: The administrative boundary shapefiles used to produce these maps were obtained from the National Institute of Statistics of Rwanda (NISR). These shapefiles are not publicly available and were shared under a data-sharing agreement. Researchers wishing to reproduce these maps should contact NISR directly to request access. **Software:** ArcGIS (8th version). No ArcGIS project files/map documents are included in this repository. The maps were exported as image files and included directly in the paper.

Output – Exhibit Linkage Table

Exhibit label in manuscript	Output file name	Output file location
Figure 1: National roads and treated municipalities in Rwanda	all_paved_final_apr26.png	WB WP submission_results2\ArcGIS
Figure 2(a): Road upgrading and tax collection	localSep25.pdf;	WB WP submission_results2_results_mainResults
Figure 2(b): Road upgrading and tax collection	centraltotalSep25.pdf;	WB WP submission_results2_results_mainResults
Figure 2(c): Road upgrading and tax collection	rentalSep25.pdf;	WB WP submission_results2_results_mainResults
Figure 2(d): Road upgrading and tax collection	licenseSep25.pdf;	WB WP submission_results2_results_mainResults
Figure 3 (a): Road upgrading and firm entry by size	entrySep25.pdf;	WB WP submission_results2_results_mainResults
Figure 3 (b): Road upgrading and firm entry by size	smallentrySep25.pdf;	WB WP submission_results2_results_mainResults
Figure 3 (c): Road upgrading and firm entry by size	mediumentrySep25.pdf;	WB WP submission_results2_results_mainResults
Figure 3 (d): Road upgrading and firm entry by size	largeentrySep25.pdf	WB WP submission_results2_results_mainResults
Figure A.1: Unpaved vs. paved road example	A1.a: road2.jpeg	WB WP submission_results2\Fotoes
	A1.b: road1.jpeg	
Figure A.2: Tax revenues from treated and control municipalities	FigA2_before.pdf; FigA2_after.pdf;	WB WP submission_results2_results\WB WP submission_results2_results_mainResults\Fig A2
Figure A.3: Trends in the sources of district budgets (2012-2024)	Distribution in the district budget.png	WB WP submission_results2_results\Descriptive
Figure A.4 (a): Central tax collections	centraltotalSep25.pdf;	WB WP submission_results2_results_mainResults
Figure A.4 (b): Central tax collections	citSep25.pdf;	WB WP submission_results2_results_mainResults
Figure A.4 (c): Central tax collections	pitSep25.pdf;	WB WP submission_results2_results_mainResults
Figure A.4 (d): Central tax collections	vatSep25.pdf	WB WP submission_results2_results_mainResults
Figure A.5 (a): Road upgrading without paving and tax collections	gravellocalSep25.pdf;	WB WP submission_results2_results_mainResults\without paved FigA5

Figure A.5 (b): Road upgrading without paving and tax collections	gravelcentralSep25.pdf;	WB WP submission_results2_results_mainResults\without paved FigA5
Figure A.6: Total number of firms across municipalities	final_number_firms2011.png, final_number_firms2019.png	WB WP submission_results2\ArcGIS
Figure A.7 (a): Firm entry by status and gender	informalentrySep25.pdf;	WB WP submission_results2_results_mainResults
Figure A.7 (b): Firm entry by status and gender	formalentrySep25.pdf;	WB WP submission_results2_results_mainResults
Figure A.7 (c): Firm entry by status and gender	femaleentrySep25.pdf;	WB WP submission_results2_results_mainResults
Figure A.7 (d): Firm entry by status and gender	maleentrySep25.pdf	WB WP submission_results2_results_mainResults
Figure A.8: National Upgraded and Non-upgraded roads	National_roads3.png	WB WP submission_results2\ArcGIS\National_roads3.png
Figure A.9: Robustness of main results	central_b, central_d, central_f, central_h; local_a, local_c, local_e, local_g	WB WP submission_results2_results_mainResults\robustness_FigA9
Figure A.10: Placebo: randomizing treatment timing	Placebo results for main.pdf	WB WP submission_results2_results_mainResults\Placebo

Acknowledgements

The authors thank Anne Brockmeyer, Arti Grover, Marcio Cruz, Byamukama Bernis, Habiyaemye Gaspard, Justice Mensah, Laura Montenbruck, and Paolo Mauro, as well as participants at seminars at IFC, the African Development Bank, and the 2025 CSAE Africa Economic Development Workshop.

Disclosure

The authors declare no conflicts of interest. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily represent the views of the World Bank Group, IFC, or their other affiliated institutions.