

# README PACKAGE FOR: The Effects of Roads and Trade Migration: Evidence from a Planned Capital City

Authors:

Melanie Morten  
Jaqueline Oliveira

With research assistance from Betsenat Gebrewold and Cecilia Moreira.

1/22/2024

## OVERVIEW

The code in this replication package replicates all the tables and figures for “The effects of roads on trade and migration: Evidence from a planned capital city.”

All analysis files are provided.

There is a main Python file, `0_run_analysis_all.py`, that runs the analysis from start-to-finish. The python calls two files: a Stata file, `/Stata_code/replication_file.do` that runs all the empirical analysis and the Matlab code (Matlab is called from Stata). A R file, `/R_code/0_run_paper_figures.R` creates all of the maps in the paper. The two files can also be run separately.

## DATA AVAILABILITY AND PROVENANCE STATEMENTS

This project uses several data sources. The data sources and access information are summarized in the table below and then detailed in the following section. All secondary analysis datafiles are provided.

## Statement of rights

I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

## Summary of Availability

Some data cannot be made publicly available.

## Details on each data source

Name	Raw data provided	Analysis data provided	Citation
Brazilian Census microdata, 1970-2010	No	Yes	IBGE (1970-2010)
State-state migration flows, 1940-1980	Yes	Yes	IBGE (1987)
			IBGE (1941-1974), Vasconcelos (2001), Vasconcelos and Oliveira (2006)
State-state trade flows, 1942-1999 (intermittent)	Yes	Yes	
Consumer price index	Yes	Yes	IPEA (n.d.)
Divisão Territorial Brasileira	No	Yes	IBGE (1940-2010)
Ehrl municipality concordance	No	Yes	Ehrl (2017)
GIS data: Brazil administrative regions	Yes	Yes	IPEA (2020)
			Brazilian Ministry of Transport (n.d.), Ministério dos Transportes (1974), Rodrigues (2008)
GIS data: Road network	Yes	Yes	

### Brazilian Census microdata, 1970-2010

*IBGE, 1970-2010. "Censo Demográfico". Instituto Brasileiro de Geografia e Estatística.*

Access information. The raw census data need to be purchased. We do not provide the raw data files but provide the analysis files.

The 1970 micro-data can be purchased from IBGE through this link:

<https://loja.ibge.gov.br/censo-demografico-1970-micrdados-da-amostra-brasil.html>

The 1980 micro-data can be purchased from IBGE through these links:

- Sudeste and Sul: <https://loja.ibge.gov.br/censo-demografico-1980-micrdados-da-amostra-sudeste-e-sul.html>

- Norte, Nordeste and Centro-Oeste: <https://loja.ibge.gov.br/censo-demografico-1980-micrdados-da-amostra-norte-nordeste-e-centro-oeste.html>

The 1991 micro-data can be purchased from IBGE through these links:

- Sudeste and Sul: <https://loja.ibge.gov.br/censo-demografico-1991-micrdados-da-amostra-regi-o-sudeste-e-sul.html>

- Norte, Nordeste and Centro-Oeste: <https://loja.ibge.gov.br/censo-demografico-1991-micrdados-da-amostra-regioes-norte-nordeste-e-centro-oeste.html>

The 2000 and 2010 micro-data are publicly available from IBGE through this link:

<https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-2020-censo4.html?=&t=micrdados>

Access conditions: The IBGE allows for redistribution for the purpose of replication as long as data are properly cited.

### State-state migration flows, 1940-1980

*IBGE (1987). Estatísticas históricas do Brasil: séries econômicas, demográficas e sociais de 1550 a 1985. Série Estatísticas Retrospectivas.*

Data were digitized from a table provided in historical yearbooks.

Access conditions: The IBGE allows for redistribution for the purpose of replication as long as data are properly cited.

#### State-state trade flows

*IBGE (1941-1974), Anuários Estatísticos do Brasil. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.*

*Vasconcelos, J. R. D. (2001). Matriz do fluxo de comércio interestadual de bens e serviços no Brasil-1999.*

*Vasconcelos, J. R. D., & Oliveira, M. A. D. (2006). Análise da matriz por atividade econômica do comércio interestadual no Brasil-1999.*

We digitized the historical trade data.

1942-1974: Digitized from tables in Anuario Estatistico do Brasil

1998: Digitized from the table in de Vasconcelos (2001).

1999: Digitized from the table in de Vasconcelos and Oliveira (2006).

Interstate trade data for the period 1942-1974 were extracted from the Anuario Estatistico do Brasil (statistical yearbooks), available from this link:

<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=720>

Access conditions: The IBGE allows for redistribution for the purpose of replication as long as data are properly cited.

#### Consumer price index

*Institute for Applied Economic Research (n.d.). National Consumer Price Index (INPC). Available in <ipeadata.gov.br>. Accessed on 8/26/2021.*

#### INPC - Índice Nacional de Preços ao Consumidor

INPC data was obtained from the Institute for Applied Economic Research (IPEADATA).

Access conditions: IPEADATA does not provide the ability to store or reference custom extracts, but allows for redistribution for the purpose of replication as long as data are properly cited. The replication files folder provides the extracted data.

Instructions on how to obtain the data are presented below:

To obtain the series, (i) go to ipeadata.gov.br; (ii) click the green "regional" tab; (iii) go to "temas/precos"; (iv) and select the data series "INPC - capitals - indice (dez. 1993 = 100). Select "Municípios" as the geographical level and the desired start and finish dates for the series.

## List of official Brazilian geographical units, 1940-2010 (Diviso Territorial Brasileira)

*IBGE (1940-2010), Divisão Territorial Brasileira, available at:*

<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/23701-divisao-territorial-brasileira.html?edicao=23704&t=downloads>. Accessed August 2021.

These files list the underlying geographical units. We used this to construct the municipality concordance.

Raw data not provided; available to download from website

URL: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/23701-divisao-territorial-brasileira.html?edicao=23704&t=downloads>

## Crosswalk of minimum comparable areas

*Ehrl, P. (2017). Minimum comparable areas for the period 1872-2010: an aggregation of Brazilian municipalities. Estudos Econômicos (São Paulo), 47, 215-229.*

We use the file IBGE\_1872\_2010\_original.xls available from <https://sites.google.com/site/philippehrl/research> to build the concordance by merging together the IBGE (1940-2010) files.

URL: <https://sites.google.com/site/philippehrl/research>

Last accessed: August 2021

## Brazilian administrative regions (GIS shapefiles)

*IBGE (2020), Malhas Municipais 2020. Available at:*

[https://geoflp.ibge.gov.br/organizacao\\_do\\_territorio/malhas territoriais/malhas municipais/](https://geoflp.ibge.gov.br/organizacao_do_territorio/malhas territoriais/malhas municipais/). Accessed: 3/31/2021

We use the 2020 municipality shape file as the base shape file for our analysis.

Access conditions: The IBGE allows for redistribution for the purpose of replication as long as data are properly cited.

## Brazilian highway system, 1940-2010 (GIS shape files)

Ministério dos Transportes (n.d.), Evolução Rodoviária nos anos 60, 70, 80, 90, 2000, 2010.

URL: <https://www.gov.br/infraestrutura/pt-br/assuntos/dados-de-transportes/bit/bitmodosmapas#maprodo>. Accessed: 3/21/2021.

Ministério dos Transportes (1974), Planos de Viacao – Evolucao Historica 1808 – 1973, Rio de Janeiro.

Rodrigues, Paulo Roberto Ambrosio (2008). *Introdução aos sistemas de transporte no Brasil e à logística internacional*. Edições Aduaneiras.

We accessed shape files for 1960-2010 from the Brazilian Ministry of Transport (Ministério dos Transportes). Data was accessed from: <https://www.gov.br/infraestrutura/pt-br/assuntos/dados-de-transportes/bit/bitmodosmapas#maprodo> on 3/21/2021. Please note, as of 1/24/2024, this link appears to be broken.

To create the 1940 and 1950 road maps we took the 1960 road network and deleted the roads that we constructed between 1950 and 1960, and 1940 and 1950. To do this, we used the Brazilian National Transport plans (Ministerio dos Transportes, 1974) and Rodrigues (2008).

## DATASET LIST

Name	Files	Location	Provided	Source
Bartik shock	bartik_shock_longdiff_meso.dta	/SecData/Analysis_data	Yes	Constructed from non-public raw census files, IBGE (1970-2010)
Bartik shock	bartik_shock_trade_meso.dta	/SecData/Analysis_data	Yes	Constructed from non-public raw census files, IBGE (1970-2010)
Collapsed census data	census_data_for_structural_dataset_meso.dta	/SecData/Analysis_data	Yes	Constructed from non-public raw census files, IBGE (1970-2010)
State-state migration flows, 1940-1980, merged with travel time	mig_UF1940.dta	/SecData/Analysis_data	Yes	Digitized from historical source. IBGE (1987)
Meso-meso migration flows, 1970-2010, merged with travel time	N_od_meso.dta	/SecData/Analysis_data	Yes	Constructed from non-public raw census files, IBGE (1970-2010)
State-state trade flows merged with travel time	trade_UF1940.dta	/SecData/Analysis_data	Yes	Digitized from various historical sources. IBGE (n.d.)
State-state trade flows	bilateral_interstate.dta	/SecData/Digitized_PDF	Yes	Digitized from various historical sources. IBGE (n.d.)
Geographical concordances	UF_concordance.dta	/SecData/Geo_concordance	Yes	Created by authors
	municipality_concordance_2010.dta	/SecData/Geo_concordance	Yes	Created by authors, using IBGE (2010) and Ehrl (2017)
	municipality_concordance_with_ehrl.dta	/SecData/Geo_concordance	Yes	Created by authors, using IBGE (2010) and Ehrl (2017)
	raw_UF_codes.dta	/SecData/Geo_concordance	Yes	Constructed from non-public raw census files, IBGE (1970-2010)
Travel time data	travel_times_google_all_cities.dta	/SecData/GIS_output	Yes	Created by authors using google server
	tt_all_citiesspeed_10.dta	/SecData/GIS_output	Yes	Created by authors using fast marching algorithm
	tt_mesospeed_10.dta	/SecData/GIS_output	Yes	Created by authors using fast marching algorithm
	tt_uf1940speed_10.dta	/SecData/GIS_output	Yes	Created by authors using fast marching algorithm
Road length and distances	dist_centroid_meso.dta	/SecData/GIS_distances	Yes	Constructed from GIS road files
	total_road_length_by_year.dta	/SecData/GIS_distances	Yes	Constructed from GIS road files
Price index	price_index_brasil.dta	/SecData/IPEA/CPI	Yes	Institute for Applied Economic Research (IPEA). National Consumer Price Index (INPC)
	price_index_muni.dta	/SecData/IPEA/CPI	Yes	Institute for Applied Economic Research (IPEA). National Consumer Price Index (INPC)
GIS data	brasilia_prj.shp	/Public/GIS_data/cities	Yes	IBGE (2020)
	brazil_capital_cities_prj.shp	/Public/GIS_data/cities	Yes	Created from IBGE (2020)
	meso_prj.shp	/Public/GIS_data/meso	Yes	IBGE (2020)
	8_regions_prj.shp	/Public/GIS_data/mst	Yes	Created by authors in ArcGIS
	mst_all_prj.shp	/Public/GIS_data/mst	Yes	Created by authors using minimum spanning tool in ArcGIS
	mst_all_rio_prj.shp	/Public/GIS_data/mst	Yes	Created by authors using minimum spanning tool in ArcGIS
	mst_pie_prj.shp	/Public/GIS_data/mst	Yes	Created by authors using minimum spanning tool in ArcGIS
	uf27_prj.shp	/Public/GIS_data/uf	Yes	Created by authors using concordance and IBGE (2020)
	uf1940_prj.shp	/Public/GIS_data/uf1940	Yes	Created by authors using concordance and IBGE (2020)
Road network	highways_1940_prj.shp	/Public/GIS_data/roads/1940	Yes	Created by deleting roads from 1960 road network (Ministério dos Transportes, n.d., accessed 2021) based on Rodrigues (2008) and Ministério dos Transportes (1974).
	highways_1950_prj.shp	/Public/GIS_data/roads/1950	Yes	Created by deleting roads from 1960 road network (Ministério dos Transportes, n.d., accessed 2021) based on Rodrigues (2008) and Ministério dos Transportes (1974).
	highways_1960_prj.shp	/Public/GIS_data/roads/1960	Yes	Ministério dos Transportes (n.d.). Accessed 2021.
	highways_1970_prj.shp	/Public/GIS_data/roads/1970	Yes	Ministério dos Transportes (n.d.). Accessed 2021.
	highways_1980_prj.shp	/Public/GIS_data/roads/1980	Yes	Ministério dos Transportes (n.d.). Accessed 2021.
	highways_1990_prj.shp	/Public/GIS_data/roads/1990	Yes	Ministério dos Transportes (n.d.). Accessed 2021.
	highways_2000_prj.shp	/Public/GIS_data/roads/2000	Yes	Ministério dos Transportes (n.d.). Accessed 2021.
	highways_2010_prj.shp	/Public/GIS_data/roads/2010	Yes	Ministério dos Transportes (n.d.). Accessed 2021.

## **COMPUTATIONAL REQUIREMENTS**

## Software requirements

Stata was last run with version 17. The Stata codes require that the following packages be installed:

- blindschemes,
- did\_imputation,
- estout,
- ftools,
- ivreg2,
- ranktest,
- ppmlhdfe, and
- reghdfe.

Matlab code was run using Matlab\_2022a. The Matlab codes requires Matlab's Optimization Toolbox.

R code was run using RStudio 2022.02.0.

The code takes approximately 3 hours to run.

## Controlled Randomness

N/A

## Memory, Runtime, and Storage Requirements

The replication archive is 1.6 GB. The code was tested on a Macbook Pro with 32 GB memory and a 1000 GB hard drive.

## **DESCRIPTION OF PROGRAMS/CODE**

We provide a masterfile, 0\_run\_analysis\_all.py that runs all the analysis in the paper. This python code calls two sub-files: a file that runs all the analysis in Stata and Matlab (/Stata\_code/replication\_file.do) and a second file that runs all the analysis in R (/R\_code/0\_run\_paper\_figures.R). We provide all secondary data analysis data in the paper.

Once the code has run, the replicator can compile the tex file in /Outpath/tables\_figure.tex that will compile all the tables and figures.

## **INSTRUCTIONS TO REPLICATORS**

The replicator needs to set the local folder and set paths to Stata and Matlab.

If running the Python code, do this by editing:

Line 15: MY\_PATH = "ADD PATH TO FOLDER HERE"  
e.g., MY\_PATH = "/Users/melaniemorten/Downloads/183316"  
Line 18: # Edit MY\_STATA to the edition of Stata on your computer  
e.g., MY\_STATA = "stata-mp"

Complete the instructions to update the zsh/bash profile given on line 26 of the script to set the path for Matlab:

```
# need to have updated paths
# can type 'path' inside matlab to see where matlab path is. Should be e.g.,
# Applications/MATLAB_R2022a.app
# if the path needs to be edited: from terminal, run: nano ~/.zshrc (or nano ~/.bash_profile) then
# paste the following lines into the file
# export PATH=/Applications/MATLAB_R2022a.app:$PATH
# export PATH=/Applications/StataMP.app/Contents/MacOS:$PATH
# confirm the filepaths are correct by typing 'which stata-mp' in the terminal and 'which matlab'
# if have problems: make sure that the files paths are consistent across both the bash profile and
# the zshrc profile. You can verify which shell is being used by entering "echo $SHELL" inside
# the terminal.
```

If the replicator prefers, the two files can be run separately in Stata and R instead of through the Python code. The Stata code should be run first and the R code should be run second.

To do this, edit /Stata\_code/replication\_file as follows:

```
line 17 (local replicationpath) to the file path where folder is downloaded e.g., local
replicationpath /Users/melaniemorten/Downloads/183316
line 18 (local matlabLoc) to where matlab is located (likely:
"/Applications/MATLAB_R2022a.app/bin/matlab")
```

Edit /R\_code/0\_run\_paper\_figures.R as follows

line 47 (MY\_PATH) in 0\_run\_paper\_figures.R to set the folder

## LIST OF TABLES AND PROGRAMS

Each table and figure presented in the paper, with its applicable program, is listed below. Once the replication files have run, the created tables and figures can be easily compiled by compiling the tex file /Output/tables\_figures.tex.

Figure 1: Map of straight-line instrument and radial highways  
- Created in: fig\_instrument.R

Figure 2: Time-varying effects of the predicted road network

- Created in: makeEventStudies.do

Table 1: Migration and Trade elasticities to travel time on roads

- Created in: makeIVEstimates.do

Table 2: Migration and housing elasticity

- Created in: estMigHouseElast.ado

Table 3: Counterfactual migration and trade costs

- Created in: table\_cf\_costs.ado

Table 4: GE counterfactuals

- Created in: GE\_cf\_2022\_10\_22

Figure 3: Spatial heterogeneity in welfare impacts of roads

- Created in: plot\_GE\_cf.R

List of tables and figures in appendix

Appendix Figure 1: Evolution of Brazil's federal highway system, 1940-2010

- Created in: fig\_road\_evolution.R

Appendix Figure 2: The growth of the road network

- Created in: fig\_length\_road.R

Appendix Figure 3: Time-varying effects of the predicted road network (State vs Meso geo-units)

- Created in: appendix\_figures.do

Appendix Figure 4: Time-varying effects of the predicted road network (Imputation Estimator)

- Created in: appendix\_figures.do

Appendix Table 1: Correlation of MST-induced reduction in travel time with geographic variables

- Created in: show\_corr\_traveltimes.ado

Appendix Table 2: Time-varying effects of the predicted road network.

- Created in: makeEventStudies.do

Appendix Table 3: Migration elasticities to travel time on roads - three-way clustered standard errors

- Created in: makeIVEstimates.do

Appendix Table 4: Migration and Trade elasticities to travel time on roads: First Stage

- Created in: makeIVEstimates.do

Appendix Table 5: Elasticities to travel time on roads: PPML robustness to higher-order first-stage residual polynomials

- Created in: makeIVEstimates.do

Appendix Table 6: Elasticity of trade flows to distance

- Created in: trade\_dist.ado

Appendix Table 7: Migration and Trade elasticities to travel time on roads - robustness to dropping Goias

- Created in: makeIVEstimates.do

Appendix Table 8: Estimating trade elasticity

- Created in: estimate\_trade\_theta.ado

Appendix Table 9: Comparing response across geography and stock/flow

- Created in: compare\_units.ado

Appendix Table 10: Estimating distance coefficient (meso level)

- Created in: convert\_state\_costs\_meso.ado

Appendix Table 11: Estimating distance coefficient (state-level)

- Created in: convert\_state\_costs\_meso.ado

Appendix Table 12: First stage: migration and housing elasticity

- Created in: estMigHouseElast.ado

Appendix Table 13: Imputed prices: migration and housing elasticity

- Created in: estMigHouseElast.ado

Appendix Table 14: Robustness of welfare to elasticity and distance elasticity

- Created in: GE\_cf\_2022\_10\_22

Appendix Figure 5: Geographical units

- Created in: fig\_geo\_regions.R

Appendix Figure 6: Construction of MST networks

- Created in: plot\_mst\_pie\_uf27.R

Appendix Table 15: Validation of travel time measure with google travel times

- Created in: check\_google\_rad\_mst.ado

Appendix Table 16: Classification of trade data by year

- Created in: TradeDataAvailability

Appendix Table 17: Census summary stats

- Created in: summary\_stats.ado [uses non-public data]

Appendix Table 18: Summary statistics of hedonic variables

- Created in hedonic\_reg.ado [uses non-public data]

Appendix Table 19: Hedonic regressions

- Created in hedonic\_reg.ado [uses non-public data]

## REFERENCES

Ehrl, Philipp. "Minimum comparable areas for the period 1872-2010: an aggregation of Brazilian municipalities." *Estudos Econômicos (São Paulo)* 47 (2017): 215-229.

IBGE. 1940-2010. Divisão Territorial Brasileira. Instituto Brasileiro de Geografia e Estatística.

<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/23701-divisao-territorial-brasileira.html?edicao=23704&t=downloads> (accessed August 2021).

IBGE. 1941-1974. Anuários estatístico do Brasil. Instituto Brasileiro de Geografia e Estatística. Rio de Janeiro.

IBGE, 1970-2010. "Censo Demográfico". Instituto Brasileiro de Geografia e Estatística.

IBGE. "Estatísticas históricas do Brasil: séries econômicas, demográficas e sociais de 1550 a 1985." *Séries Estatísticas Retrospectivas* (1987).

IBGE. 2020. Malhas Municipais 2020.

[https://geoftp.ibge.gov.br/organizacao\\_do\\_territorio/malhas territoriais/malhas\\_municipais/](https://geoftp.ibge.gov.br/organizacao_do_territorio/malhas territoriais/malhas_municipais/). Accessed: 3/31/2021

Institute for Applied Economic Research (IPEA), n.d. National Consumer Price Index (INPC). IPEADATA: Database from the Institute for Applied Economic Research. Regional. 2021. ipeadata.gov.br (accessed 8/26/2021).

Ministério dos Transportes (n.d.), Evolução Rodoviária nos anos 60, 70, 80, 90, 2000, 2010.  
<https://www.gov.br/infraestrutura/pt-br/assuntos/dados-de-transportes/bit/bitmodosmapas#maprodo>  
(accessed 3/21/2021).

Ministério dos Transporte. Planos de Viacao – Evolucao Historica 1808 – 1973, Rio de Janeiro. 1974.

Rodrigues, Paulo Roberto Ambrosio. *Introdução aos sistemas de transporte no Brasil e à logística internacional*. Edições Aduaneiras, 2008.

Vasconcelos, José Romeu de. "Matriz do fluxo de comércio interestadual de bens e serviços no Brasil-1999." (2001).

Vasconcelos, José Romeu de, and Márcio Augusto de Oliveira. "Análise da matriz por atividade econômica do comércio interestadual no Brasil-1999." (2006).