

DATA SUPPLEMENT
WHY HAS BRAZIL DEINDUSTRIALIZED SO MUCH?
AN EMPIRICAL INVESTIGATION

August 28, 2024

Abstract

The purpose of this file is to assist in the reproducibility of the paper *Why has Brazil deindustrialized so much? An empirical investigation* by Edmar L. Bacha, Victor S. Terziani, Claudio M. Considera, and Eduardo A. Guimarães. This file includes data sources, detailed explanations of procedures, methodology descriptions, and instructions on how to replicate the study. Additionally, 16 Excel files, 7 R files and 3 csv files are necessary companions to this file for reproducing the paper's results.

CONTENTS

1	Data	3
1.1	Brazilian Industrialization Rate and Relative Prices of Manufacturing . . .	3
1.2	OECD Industrialization Rate and Relative Prices of Manufacturing	4
1.3	Real Exchange Rate	5
1.4	Terms of Trade	6
1.5	Real Broad Dollar Index	6
1.6	Dummy Variables	6
1.7	Brazilian Manufacturing Relative Productivity	7
1.8	G7 Manufacturing Relative Productivity	7
2	Tables	9
2.1	Table 1: Regression Results & Table B1: REER Regression	9
2.2	Table A.1: Ljung-Box Tests	9
2.3	Table A.2: Engle-Granger Tests	9
3	Figures	10
3.1	Figures 1 through 8	10
3.2	Figures A.1 through A.4	10
4	Files	11

1 Data

This section provides a comprehensive guide on how to obtain and process the data in the paper, ensuring it is ready to reproduce the tables and figures.

The *data.xlsx* file contains all the final products described in the following sections, except for the productivity data.

This file has two sheets:

- The first is designed for use with statistical software.
- The second is intended for use by Excel users.

1.1 Brazilian Industrialization Rate and Relative Prices of Manufacturing

The Brazilian Industrialization Rate is measured by the ratio of manufacturing value added to the country's GDP. This ratio is presented in both constant prices and current prices.

Quarterly data for manufacturing value added and GDP are sourced from the file *Tab_Comp_CNT_1T24.xls*, which can be downloaded from IBGE's *Sistema de Contas Nacionais Trimestrais*¹ under the zipped folder *Tabelas_Completas*.² In the file *Tab_Comp_CNT_1T24.xls*, the sheet *Valores Correntes* provides current price data for manufacturing value added in column D and GDP in column R. The sheet *Valores Encadeados a Preços de 95* presents the same data in the same columns, but in constant prices of 1995.

Since the data is available only from 1996 onwards, we estimate the data for 1995 using information from FGV-IBRE found in the sheet *Série Encadeada Trimestral* of the file *fgv_ibre_data.xlsx*. We use the quarterly index of the value added of manufacturing volume and the GDP volume.

The file *br_ind_rate_final_data.xlsx* performs all the necessary calculations to derive the final industrialization rate series. In this file, the sheet *FGV IBRE Data* calculates

¹Brazilian expression for "Quarterly National Accounts System"

²Link to Tabelas_Completas, as of 03/07/2024: <https://www.ibge.gov.br/estatisticas/economicas/industria/9300-contas-nacionais-trimestrais.html?=&t=downloads>

manufacturing value added and GDP in constant prices by extending IBGE's series backward using FGV IBRE's volume index series. The sheet *IBGE Data* merges IBGE's data with the newly generated data and converts the series from constant prices of 1995 to constant prices of 2015Q1. Finally, the sheet *Industrialization Rate* calculates the industrialization rates for both current and constant prices and determines the relative prices by taking the ratios of the current prices series to the constant prices series.

1.2 OECD Industrialization Rate and Relative Prices of Manufacturing

The primary challenge in constructing these data was their unavailability in the original sources³. Therefore, we had to estimate them. We proceeded as follows: first, we extracted U.S. manufacturing value added (VA) and GDP data from the BEA website. However, these data were only available from 2005Q1. We used the data from 2005Q1 to 2007Q4 to estimate the seasonality and extended the series backward using the available annual data from BEA. Next, we combined these data with data from 19 OECD countries⁴ to create a proxy group for the OECD. Finally, we used the quarterly data from this group to convert the annual data available on the OECD website into quarterly data. The detailed process is described below.

The data file *ValueAdded.xlsx*⁵ contains the quarterly BEA data for manufacturing value added and GDP. Current price values are in the sheet *TVA105-Q*, and constant price values are in the sheet *TVA106-Q*. The data file *usa_quarterly_data.xlsx* contains the data extracted from *ValueAdded.xlsx* in an ordered format. The code file *us.data.R* reads the manufacturing VA from 2005Q1 to 2007Q4 and then estimates quarterly dummies for the data. Using these quarterly dummies, we estimate the manufacturing VA, quarter by quarter, for the years 1995 to 2004, with the following steps. First, we estimate the first-quarter value for each year, using the formula:

$$Q_{1,t} = A_t \times \frac{4}{1 + (1 + D_2) + (1 + D_3) + (1 + D_4)}$$

where A_t is the annual value for year t , and D_i is the dummy for the i -th quarter. After estimating the first-quarter value, we used the dummies to multiply this value for each

³We couldn't find the data in the OECD website, as of 03/07/2024

⁴The full list of 19 countries plus the U.S. is: Austria, Belgium, Costa Rica, Czechia, Denmark, Finland, France, Germany, Hungary, Italy, Japan, Latvia, Luxembourg, Netherlands, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States.

⁵Link, as of 03/07/2024 <https://apps.bea.gov/histdatacore/fileStructDisplay.html?theID=11912&HMI=8&oldDiv=Industry%20Accounts&year=2023&quarter=,%20Q4&ReleaseDate=March-28-2024&Vintage=Quarterly>

quarter:

$$Q_{i,t} = Q_{1,t} \times (1 + D_i)$$

where $Q_{i,t}$ is the value for the i -th quarter in year t . The annual observations were obtained from various BEA sources and are in the file data *annual_us_manufacturing.xlsx*. This process allowed us to expand the U.S. series. Then, *us_data.R* reads U.S. GDP data for the period from the data file *GDP_USD_Luiz.csv*⁶ and merges it with the new manufacturing value-added data. Then it calculates the U.S. industrialization rate from 1995Q1 to 2004Q1.

The code file *oecd_industrialization_rate.R* reads data from the file *GDP_VA_Luiz.xlsx*, which includes industrialization rates for the list of countries, and merges this data with the newly generated U.S. data of up to 2005. Then, it calculates each country's contribution to the group's GDP and uses this contribution to compute the group's industrialization rate. It saves this data in file *final_group_data.xlsx*.

Finally, the file *oecd_conversion.xlsx* uses the calculated group's industrialization rate to convert annual OECD data from the OECD website into quarterly data. It first calculates the quarterly deviation of the country group's observations from the annual average. Then it applies those deviations into the annual OECD data. This process yields the final series, ready for use. The relative prices are given by the ratio of the current prices series to the constant prices series.

1.3 Real Exchange Rate

The file *rer_bcb.xlsx* contains the raw data for the Real/USD exchange rate, sourced from the Central Bank of Brazil (BCB) website. The file also performs the necessary transformations: it uses the BCB index monthly values to calculate the quarterly average and, finally, changes the base of the quarterly series to 2015Q1.

The file *reer_bcb.xlsx* contains the raw data for the Real/USD exchange rate, sourced from the Central Bank of Brazil (BCB) website. The file also performs the necessary transformations: it uses the BCB index monthly values to calculate the quarterly average and, finally, changes the base of the quarterly series to 2015Q1.

⁶We are indebted to Luiz de Mello from OECD for this data

1.4 Terms of Trade

The file *tot.xlsx* contains the raw data for Brazil's Terms of Trade, sourced from Funcex. The file also performs the necessary transformations: it uses the Funcex monthly index values to calculate the quarterly average and, finally, changes the base of the quarterly series to 2015Q1.

1.5 Real Broad Dollar Index

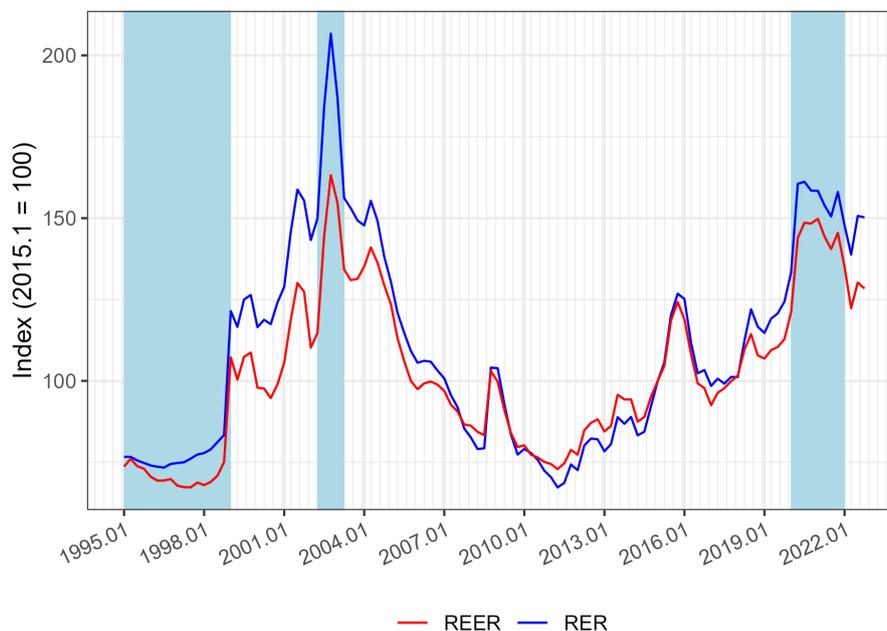
The file *dollar_index_fed.xlsx* contains the raw data for two monthly Real Broad Dollar Index time series, referred to as the Old Series and the New Series. Both series were sourced from the U.S. Fed. The file converts the monthly data to quarterly averages by calculating the mean of the monthly observations. It then calculates the average difference between the two series for the four quarters of 2006. Finally, it merges the series by using the New Series from 2006 onwards and subtracting the average difference from the Old Series to extend it backward. The series base is changed by dividing the series by its 2015Q1 value.

1.6 Dummy Variables

Binary variables are determined by historical facts and events. However, the precise timing of when they are set to 1 or 0 is based on the behavior of the Real Exchange Rate. In Figure 1, below, we highlight the timing of the binary variables and illustrate how the level of the Real Exchange Rate changes during these periods. The Real Effective Exchange Rate is also included for comparison. The timings are:

- Managed: 1995Q1 - 1998Q4
- Fear of Lula: 2002Q3 - 2003Q1
- COVID: 2020Q2 - 2021Q4

Figure 1: Exchange Rates and binary variables



1.7 Brazilian Manufacturing Relative Productivity

The file *productivity_br.xlsx* contains data on Brazilian employment and value added for both the manufacturing sector and the total economy, sourced from FGV-IBRE. The file also performs the necessary transformations: it first calculates the value added per person employed - referred to as employment productivity - for both the manufacturing sector and the total economy. Then, it calculates the relative manufacturing productivity by taking the ratio of manufacturing employment productivity to total employment productivity. This data is used in Figure 7 and Figure 8 of the paper.

The same process is done for the Agriculture data in *agro_data* sheet.

1.8 G7 Manufacturing Relative Productivity

The G7 is composed by the United States (US), the United Kingdom (UK), Japan, Germany, France, Italy and Canada.

The file *employment_data.R* processes the following employment data:

- Employment data for the European countries and Japan, from 1995 to 2020, sourced

from the EUKLEMS section of the Luiss Lab of European Economics website⁷ and stored in *national_accounts.Rds*.

- Employment data for the European countries and Japan, from 2021 to 2022, sourced from the OECD website and stored in *oecd_g7_employment.csv*.
- Employment data for the US, from 1995 to 2022, sourced from the BLS website⁸ and stored in *bls_employment.xlsx*
- Employment data for Canada, from 1995 to 2022, sourced from the Statistics Canada website⁹ and stored in *canada_employment.xlsx*.

It also reads and processes file *un_value_added.xlsx*, sourced from the UN website¹⁰ which contains manufacturing Value Added and Total Value Added for all countries in the G7, from 1995 to 2022.

The script first merges the OECD employment data with the EUKLEMS employment data for Japan and the European countries. Then it merges the results with employment data files of the US and Canada and the UN file of Value Added. Finally, it sums the Manufacturing Value Added of all countries, as well as Total Value Added and Employment to arrive to the G7 figures. It uses the G7 aggregated data to calculate the relative manufacturing productivity by taking the ratio

$$\frac{\left(\frac{\text{Manufacturing Value Added}}{\text{Total Value Added}} \right)}{\left(\frac{\text{Manufacturing Employment}}{\text{Total Employment}} \right)}$$

This merged data is combined with Brazilian data and stored in the file *productivity_final_data.xlsx*, which is used to generate Figure 7 and Figure 8 in the paper.

⁷Link, as of 03/07/2024: <https://euklems-intanprod-1lee.luiss.it/download/>

⁸Link, as of 03/07/2024: <https://www.bls.gov/productivity/tables/hours-employment-detailed-industries.xlsx>

⁹Link, as of 03/07/2024: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410002301&pickMembers%5B0%5D=1.1&pickMembers%5B1%5D=2.2&pickMembers%5B2%5D=4.1&pickMembers%5B3%5D=5.1&cubeTimeFrame.startYear=1995&cubeTimeFrame.endYear=2022&referencePeriods=19950101%2C20220101>

¹⁰Link, as of 03/07/2024: <https://unstats.un.org/unsd/snaama/Basic>

2 Tables

2.1 Table 1: Regression Results & Table B1: REER Regression

Table 1 contains all the estimated regressions presented in the paper. The code to generate Table 1 is contained in the file *regression_and_tests.R*. This code performs the following steps:

- Estimates the first stage using the standard *lm* R function.
- Estimates the second regression using the standard *lm* R function.
- Estimates the third regression using the *iv_robust* function, as the standard *lm* function does not properly estimate the standard errors for a IV second stage regression.

The file also sources code from *helper_functions.R* to assist in producing better tables.

Table B1, which uses the Real Effective Exchange Rate for regressions, is generated with code similar to that of Table 1. This code is also found in the file *regression_and_tests.R*.

2.2 Table A.1: Ljung-Box Tests

Table A.1 contains all the estimated Ljung-Box tests presented in the appendix. The code to generate Table A.1 is contained in the file *regression_and_tests.R*. It performs the Ljung-Box tests using function *Box.test* from the *stats* package. The file also sources code from *helper_functions.R* to assist in producing better tables.

2.3 Table A.2: Engle-Granger Tests

Table A.2 contains all the estimated Engle-Granger tests presented in the appendix. The code to generate Table A.2 is contained in the file *regression_and_tests.R*. The code performs the Engle-Granger tests by using the *ur.df* function from the *urca* package, which conducts an Augmented Dickey-Fuller test on the residuals of all regressions. The file also sources code from *helper_functions.R* to assist in producing better tables.

3 Figures

3.1 Figures 1 through 8

Code for figures 1 through 8 are contained in file *figures.R*. It uses the *ggplot2* R package to make plots of the desired variables.

3.2 Figures A.1 through A.4

Code for figures A.1 through A.4 are contained in file *regression_and_tests.R*. It uses *ggAcf* and *ggPacf* functions from the *forecast* package to make Autocorrelation Function and Partial Autocorrelation Function plots for the second stage regressions.

4 Files

This section enlists all files contained in the Data Supplement.

annual_us_manufacturing.xlsx

bls_employment.xlsx

br_ind_rate_final_data.xlsx

canada_employment.csv

data.xlsx

dollar_index_fed.xlsx

employment_data.R

fgv_ibre_data.xlsx

figures.R

final_group_data.xlsx

GDP_USD_Luiz.csv

GDP_VA_Luiz.xlsx

helper_functions.R

national_accounts.Rds

oecd_conversion.xlsx

oecd_g7_employment.csv

oecd_industrialization_rate.R

productivity_br.xlsx

productivity_final_data.xlsx

reer_bcb.xlsx

regression_and_tests.R

rer_bcb.xlsx

Tab_Comp_CNT_1T24.xls

tot.xls

un_value_added.xlsx

us_data.R

usa_quarterly_data.xlsx

ValueAdded.xlsx