

Learning when schools shutdown: impacts of H1N1 outbreak on learning loss and learning gaps

2024-04-24

Introduction

School shutdowns were adopted worldwide to contain the transmission of the Covid-19. To benchmark the impacts of the current outbreak on learning, this paper explores the 2009 H1N1 pandemic in the state of São Paulo, Brazil, where all the schools managed by the state government had their winter break extended. Thirteen most hit municipalities adopted the same measure in their local schools, while the other municipal authorities maintained the school calendar as usual. We take advantage of the variation between municipalities to employ a difference-in-differences approach on the set of locally-managed schools. We also explore the variation within municipalities to employ a triple difference-in-differences that compares schools run by state and local authorities.

Unit of observation: schools. Source of the data: Prova Brasil and School Census. Grade level: fifth grade. Period: 2005, 2007 (pre-treatment) and 2009 (post-treatment). Dependent variables: proficiency of students in Portuguese, proficiency of students in Math, standardized performance in Portuguese and Math, percentage of students with insufficient performance in Portuguese and percentage of students with insufficient performance in Math. Independent variables: the vector of control variables was chosen based on a Lasso regression, which selects the variables that better predict the variation in the proficiency score. The selected variables are: if the school has science and computer lab, sport court, library and access to the internet; instruction hours per day; students per class; approval rates; GDP per capita of the municipality where the school is located; and socioeconomic characteristics of fifth graders. The vector of socioeconomic variables includes the percentage of mothers with a high school diploma; the percentage of students that already repeated or dropped out school; the percentage of whites and girls; the percentage of students that already work, that previously studied in a private school and that have a computer at home; the percentage of students whose parent's incentive them to study, to do the homework, to read, to not miss classes, and that talk about what happens in the school.

To reproduce the code, modify the path in the section Preparing Folder Paths in the Master do file.

Data Availability Statement

The data supporting the findings of this study are available from the following sources:

1. GDP at Municipal Level

- **Source:** Instituto Brasileiro de Geografia e Estatística (IBGE)
- **Files:**
- 'base_de_dados_2010_2014.xls' for GDP at municipal level from 2010-2014. Download here (Choose: Estatísticas -> Por Pesquisa e Estudo -> Produto Interno Bruto dos Municípios -> Downloads -> 2014 -> base -> base_de_dados_2010_2014.xls)
- 'base.xls' for GDP at municipal level from 1999-2009. Download here (Choose: Estatísticas -> Por Pesquisa e Estudo -> Produto Interno Bruto dos Municípios -> Downloads -> 2005-2009 -> banco_dados.zip -> base.xls)
- 'popmunic2007layoutTCU14112007.xls' for Population by municipality in 2007. Download here (Choose: Estatísticas -> Por Tema -> População -> Contagem da População -> Downloads -> Contagem_da_Populacao_2007 -> Populacao_Enviada_TCU_2007_11_14 -> popmunic2007layoutTCU14112007.xls)

- **Located:** Raw/IBGE
2. **H1N1 Hospitalizations**
 - **Source:** DataSUS
 - **File:** ‘A161451189_28_143_208.csv’ for the number of hospitalizations due to H1N1. Download here (Choose Informações de Saúde (TABNET) -> Epidemiológicas e Morbidade -> Doenças e Agravos de Notificação - 2007 em diante (SINAN) -> INFLUENZA PANDÊMICA, then follow the selections for 2009 and confirmed cases, and click on ‘COPIA COMO .CSV’)
 - **Located:** Raw/SUS
 3. **Prova Brasil Microdata**
 - **Source:** Instituto Nacional de Pesquisas Educacionais (INEP)
 - **Files:** Microdata for Prova Brasil for the years 2007 and 2009. Download here (Choose: Acesso à informação -> Dados Abertos -> Microdados -> Prova Brasil)
 - **Located:** Raw/Prova Brasil
 4. **Flow Indicators**
 - **Source:** Instituto Nacional de Pesquisas Educacionais (INEP)
 - **Files:** Download the ‘TX RENDIMENTO ESCOLAS’ files for 2007, 2008, and 2009. Download here (Choose: Acesso à informação -> Dados Abertos -> Indicadores Educacionais -> Taxas de Rendimento)
 - **Located:** Raw/Rendimiento
 5. **Censo Escolar Microdata**
 - **Source:** Instituto Nacional de Pesquisas Educacionais (INEP)
 - **Files:** Microdata for Censo Escolar for the years 2005, 2007, 2008, and 2009. Download here (Choose: Acesso à informação -> Dados Abertos -> Microdados -> Censo Escolar)
 - **Note:** Recent changes by the Brazilian Government may affect the availability and structure of these data.
 - **Located:** Raw/Censo Escolar/200X/DADOS/
 - **Direct download:** link to directly download is now here, however this is subject to change by the Brazilian Government.
 6. **IDEB (Education Development Index)**
 - **Source:** Instituto Nacional de Pesquisas Educacionais (INEP)
 - **Files:** Different files for regions, states, municipalities, and schools for the years considered, focusing on Ensino Fundamental Regular and Médio. Download here (Choose: Áreas de atuação -> Pesquisas Estatísticas e Indicadores Educacionais -> IDEB -> Resultados)
 - **Located:** Raw/IDEB
 7. **IDESP por Escola**
 - **Source:** Secretaria da Educação do Estado de São Paulo, Governo de São Paulo.
 - **Files:** IDESP by school. Download here. (Choose Idesp - Escolas Série histórica de 2007 a 2019 - Ensino Fundamental - Anos Iniciais.csv)
 - **Located:** Raw/IDESP

Explanation of the code

Background

In 2009, amid the Swine flu, part of the schools in the state had their winter-break extended:

-> All schools managed by the state government across the 645 municipalities of the state.

-> Locally-managed schools of thirteen municipalities: Sao Paulo (the capital of the state), Campinas, Diadema, Embu das Artes, Indaiatuba, Mairiporã, Osasco, São Bernardo do Campo, Santo André, São Caetano do Sul, Sumaré, Ribeirao Preto, Taboao da Serra.

This natural experiment allows us to estimate the impact of the school shutdowns on children’s learning.

Our analysis:

-> Unit of observation: schools

- > Grade level: fifth grade
- > Source of the data: Prova Brasil, IDEB, IBGE, School Census, DataSus. Details below.
- > Period: 2005, 2007 (pre-treatment) and 2009 (post-treatment)
- > Treatment Groups:
 - Schools managed by local authorities in the 13 municipalities that opted to extended the winter break
 - Schools managed by the state government across the 645 municipalities of the state
- > Comparison Group:
 - Schools managed by the local authorities in the 632 municipalities that did not opt to extend the winter break
 - The treatment dummy is whether the school was closed in 2009
 - Dependent variables: proficiency of students in Portuguese, proficiency of students in math, standardized performance in Portuguese and Math, percentage of students with insufficient performance in Portuguese, percentage of students with insufficient performance in math, approval, repetition and dropout.
 - Independent variables: treatment status (variable treated = 1 for Treatment Group; and 0 for Comparison Group); year dummy (variable year); municipality in each the school is located at (variable codmunic).

Codes

1. GDP at municipal level

-> How long it takes to run?
A few seconds

-> What it does?
The code imports GDP and number of inhabitants at municipality level

We download and work on 3 .xls files from Instituto Brasileiro de Geografia e Estatística (IBGE)

-> GDP at municipal level from 2010-2014

The name of the file is 'base_de_dados_2010_2014.xls'.

To download: <https://www.ibge.gov.br>. Choose: Estatísticas -> Por Pesquisa e Estudo -> Produto Interno Bruto dos Municípios -> Downloads -> 2014 -> base -> base_de_dados_2010_2014.xls

-> GDP at municipal level from 1999-2009

The name of the file is 'base.xls'

To download: <https://www.ibge.gov.br> Choose: Estatísticas -> Por Pesquisa e Estudo -> Produto Interno Bruto dos Municípios -> Downloads -> 2005-2009-> banco_dados.zip -> base.xls

-> Population by municipality in 2007

We only have the Demographic Census from 10 to 10 years. However, IBGE discloses population estimates every once in a while. Our best option considering our period of analysis is to look at 2007 numbers

The name of the file file is 'popmunic2007layoutTCU14112007.xls' To download: <https://www.ibge.gov.br> Choose: Estatísticas -> Por Tema -> População -> Contagem da População -> Downloads -> Contagem_da_Populacao_2007 -> Populacao_Enviada_TCU_2007_11_14 -> popmunic2007layoutTCU14112007.xls

We save these 3 .xls files in our project folder: h1n1-school-closures-sp-2009/DataWork/Datasets/Raw/IBGE

-> What it creates? One .dta file named 'GDP per capita.dta'

saved in: h1n1-school-closures-sp-2009/DataWork/Datasets/Intermediate

2. H1N1.do

-> How long does it take to run?

A few seconds

-> What it does?

The code imports the number of hospitalizations per 1,000 inhabitants at the municipality level

We download and work on one .csv file from DataSUS

-> Number of hospitalizations due to H1N1

The name of the file is 'A161451189_28_143_208.csv'

To download: <http://www2.datasus.gov.br/DATASUS/index.php>.

Choose Informações de Saúde (TABNET) ->

Epidemiológicas e Morbidade -> Doenças e Agravos de Notificação - 2007 em diante (SINAN) ->
INFLUENZA PANDÊMICA

On the screen: INFLUENZA PANDÊMICA - SINAN - BRASIL

In 'Linha', choose: 'Município de residência'

in 'Coluna', choose: 'Mês 1o sintoma(s)'

On the screen: INFLUENZA PANDÊMICA - SINAN - BRASIL

Choose 2009

On the screen: SELEÇÕES DISPONÍVEIS

Choose 'Município de residência' and select all

Choose 'Classif. Final' and select 'confirmado'

On the screen: Click on 'MOSTRA' after all sections. Go to the end of the page and click:

'COPIA COMO .CSV'

We save this .csv file in our project folder: h1n1-school-closures-sp-2009/DataWork/Datasets/Raw/SUS

-> What it creates? One .dta file named 'H1N1.dta' saved in child-labor-ban-brazil/DataWork/Datasets/Intermediate

3. Importing & Harmonizing Prova Brasil -> How long does it take to run? 7 minutes

-> What it does? The code imports the microdata from Prova Brasil. Prova Brasil is a proficiency test applied to all students from public schools in Brazil.

Students from 5th and 9th grades take Portuguese and Math exams every 2 years and answer a socioeconomic questionnaire.

Teachers and principals also answer a questionnaire on school teaching methodologies, infrastructure, sources of student deficit, etc.

We download and work on .txt files from Instituto Nacional de Pesquisas Educativas (INEP).

-> Prova Brasil Microdata To download: <https://www.gov.br/inep/pt-br> Choose: Acesso à informação -> Dados Abertos -> Microdados -> Prova Brasil

Download: Microdados Prova Brasil 2007

Microdados Prova Brasil 2009

We save both folders downloaded in: h1n1-school-closures-sp-2009/DataWork/Datasets/Raw/Prova Brasil
The folder with 2007 Microdata, we named: 2007 The folder with 2009 Microdata, we named: 2009

-> What it creates? 3 .dta files named 'Students - Prova Brasil.dta', 'Teachers - Prova Brasil', and 'Principals - Prova Brasil' saved in child-labor-ban-brazil/DataWork/Datasets/Intermediate

4. Importing & Harmonizing grade promotion, repetition, and dropout

-> How long it takes to run?

One minute

-> What it does?

The code imports the microdata of approval, repetition and dropout rates

We download and work on .txt files from Instituto Nacional de Pesquisas Educationais (INEP)

-> Flow Indicators

To download: <https://www.gov.br/inep/pt-br>

Choose: Acesso à informação -> Dados Abertos -> Indicadores Educacionais -> Taxas de Rendimento

Download: 2007, Escolas. File 'TX RENDIMENTO ESCOLAS 2007'
+
2008, Escolas. File 'TX RENDIMENTO POR ESCOLA 2008'
+
2009, Escolas. File 'TX RENDIMENTO POR ESCOLA 2009'

We save the excel files we downloaded in: h1n1-school-closures-sp-2009/DataWork/Datasets/Raw/Rendimento

-> What it creates? One .dta file named 'Flow Indicators.dta' saved in: h1n1-school-closures-sp-2009/DataWork/Datasets/Intermediate

5. Importing & Harmonizing Censo Escolar -> How long it takes to run? Half an hour.

OBS-> Starting in 2023, the Federal Government stopped allowing the download the Census of Education as it used to do when we started working paper. If you try to download the microdata, it will not be the same data organization and variables.

-> What it does? The code imports the microdata from School Census. School Census is conducted annually and with enrollments, teachers, school infrastructure and student's classrooms.

We download and work on .txt files from Instituto Nacional de Pesquisas Educationais (INEP).

-> Prova Brasil Microdata

To download: <https://www.gov.br/inep/pt-br>

Choose: Acesso à informação -> Dados Abertos -> Microdados -> Censo Escolar

Download:
Microdados Censo Escolar 2005
+
Microdados Censo Escolar 2007

```
+
Microdados Censo Escolar 2008
+
Microdados Censo Escolar 2009
```

*Recently, the Brazilian Government removed the microdata from its website.

We save the 4 folders downloaded in: h1n1-school-closures-sp-2009/DataWork/Datasets/Raw/Censo Escolar. The folder with 2005 Microdada, we named: 2007 The folder with 2007 Microdada, we named: 2007 The folder with 2008 Microdada, we named: 2008 The folder with 2009 Microdata, we named: 2009

```
-> What it creates?
    5 .dta named 'School Infrastructure'
```

'Class-Hours' 'Class-Size' 'Enrollments' 'Students per teacher' saved in: h1n1-school-closures-sp-2009/DataWork/Datasets/Intermediate

6. Importing & Harmonizing IDEB -> How long it takes to run? One minute.

-> What it does? The code imports the microdata of Education Development Index (IDEB).

We download and work on .xlsx files from Instituto Nacional de Pesquisas Educationais (INEP).

-> IDEB DE CAREFUL!!!!!!! INEP ALWAYS UPLOAD NEW DATA SO YOU CAN DOWNLOAD A FILE WITH MORE COLUMNS THAN WE DID. CHECK THE DO FILE '6. Importing & Harmonizing IDEB' TO SEE HOW MANY COLUMNS THE EXCEL FILES WE WORK WITH HAVE. YOU CAN ALSO WRITE TO US AND WE CAN SHARE THE VERSION WE DOWNLOAD FROM THE WEBSITE.

To download: <https://www.gov.br/inep/pt-br> Choose: Áreas de atuação -> Pesquisas Estatísticas e Indicadores Educacionais -> IDEB -> Resultados

Then you have two options 'Municípios' and 'Escolas'

Download: Regiões e estados Ensino Fundamental Regular e Médio. File 'divulgacao_regioes_ufs_ideb_2019'

Download: Municípios

Ensino Fundamental Regular - Anos Iniciais. File 'divulgacao_anos_iniciais_municipios_2019' + Ensino Fundamental Regular - Anos Finais. File 'divulgacao_anos_finais_municipios_2019'

Download: Escolas Ensino Fundamental Regular - Anos Iniciais. File 'divulgacao_anos_iniciais_escolas_2019' + Ensino Fundamental Regular - Anos Finais. File 'divulgacao_anos_finais_escolas_2019'

We save the excel files we downloaded in: h1n1-school-closures-sp-2009/DataWork/Datasets/Raw/IDEB

-> What it creates? 2 .dta files named 'IDEB by school.dta' 'IDEB by municipality.dta'

saved in: h1n1-school-closures-sp-2009/DataWork/Datasets/Intermediate

7. Setting up dataset -> How long it takes to run?

-> What it does? The code creates the dataset we use in our analysis by merging all data we imported and harmonized.

-> What does it create? One .dta file named 'h1n1-school-closures-sp-2009' saved in h1n1-school-closures-sp-2009/DataWork/Datasets/Final

8. Descriptives -> How long it takes to run? Few minutes

-> What it does? The code runs the descriptive statistics.

-> What does it create? Balance test tables saved in h1n1-school-closures-sp-2009/DataWork/Output/Tables Time-trend figures saved in h1n1-school-closures-sp-2009/DataWork/Output/Figures

9. Regressions -> How long it takes to run? At least 12 hours (in my computer)

-> What it does? The code runs our regressions on the impact of the school shutdowns on learning.

-> What does it create? Regression tables saved in h1n1-school-closures-sp-2009/DataWork/Output/Tables

****Não testamos se monitores ajudam porque vi por ex que em 2009 não havia monitores em SP nas turmas de 5o ano. **Também não testamos se**

Globals.

The code sets globals with:

- control variables used in balance test and regressions
- codes of the treated municipalities

Folder structure It is possible to reproduce all the analysis from scratch using only the codes available in this repository.

Set up the folder structure:

- Create the folder “h1n1-school-closures-sp-2009”.
- inside this folder, create “DataWork”
 - inside “DataWork”, create “Datasets”
 - inside "Datasets", create 3 folders: "Raw", "Intermediate" and "Final"
 - inside "Raw", create the folders: IBGE, SUS, Prova Brasil, Censo Escolar, IDEB, Rendimento, RAIS
 - inside each of the folders created in Raw save the .txt files, .csv and xls files we instructed you to download (in the explanation of each do file, we show how to download all the microdata you need to run the analysis.
 - inside “DataWork”, create “Output”
 - inside “Output”, create 2 folders: “Tables” and “Figures”.

For questions please message Vivian Amorim (vivianamorim5@gmail.com)