Policy Analytics Challenge

Effects of climate shocks on educational outcomes, an application to Brazil.

Federico Veneri

October 2025

FOCUS OF THE PROJECT

OBJECTIVE

This project aims to explore the usefulness of public available dataset to:

- Quantify the effect of extreme climate condition on developmental outcomes (health/education).
- Provide some examples of modeling strategies that could be used to quantify these effects.

DATA

I focus on Brazil for this project, but LAC countries have similar data**.

- Mortality data.
- Hospital admission.
- Education attendance.
- Education outcomes.
- Extreme weather events

For Brazil, I explored:
- Mortality: Datasus-SIM
- Hospital admission
- Datasus-SIH
- **Educational outcomes:**
    - SAEB, pass rate

ANALYSIS & RESULTS

Data was aggregated to municipality level, climate shocks were defined

- **Education**: Extreme rainfall was associated with lower standardized test results on math and language, and lower pass rates at municipality level.
- **Health**: forthcoming.

** have!= public available

Agenda

1. **MUNICIPALITY FRAGILITY**
2. **DATA AND CHALLENGES**
3. **MODELING**
4. **RESULTS**
5. **POLICY DISCUSSION**

Climate challenges and education.

- Between 2022-2024, 400 million students globally have experience school closure due to extreme weather (WB,2024).
  - Higher impacts on lower income countries: 18 days lost per year (10% of academic year)
  - Even if a school remains open, data shows reduced attendance under climate shocks.
  - Lost schooling translate to learning deficits. Not only days lost but also interruptions in pedagogy plans.
    - Learning deficits affects human capital accumulation and ultimately growth.

- Evidence from Latin America and the Caribbean, provides additional evidence in same direction.
  Bernal & Vlaicu (2023): Colombia
  Rural households re allocate children time from school to domestic activities when faced with climate shocks.
  Lima et. al. (2024): Brazil
  Extreme weather events (Rainfall) results in adverse impact on standardized test results.

- The previous results showcase recent results for LAC associated with rainfall, but the literature is more richer:
  - Other extremeate weather events (heatwaves, drougths) have also been studied.
  - And other developmental outcomes (hospital admision, agricultural outcomes)

- My work is focus on rainfall shocks and educational challenges in Brazil
  - Tries to reproduce **Lima et. al. (2024)** based on public data.

Data and challenges

Municipal Mesh (IBGE).

- The **Brazilian Institute of Geography and Statistics** produce official shapefiles with Municipalities, UF, etc.

Acces: The shapifles are readily accesible, code implement to download the files. An R package is also avalible: geobr (IPEA)

Education microdata (INEP)

Sistema de Avaliação da Educação Básica (SAEB) and School Census (censo escolar).

- Provides infomation on educational outcomes (lenguage and math) at school and student level.
    - Data avalible for every two years (1995-2023)
    - Controls can be created at indivdual (SAEB) or school level (using school census). Potential geocoded for schools.

Acces: The microdata is accesible, code implement to download the files.

- **Challenges:** The municipality and school code is hashed, but previous author have aquired the unhased data.
    - Some public reports have municipality level aerages.

Climate

Copernicus ERA5-Land

- Provides information on key climate variables (e.g.: wind, precipitation, temperature).
    - Data is available hourly with a 9km resolution. Daily aggregations can be derived (sum , max, etc.)

Acces: Some R package allow to access the data, KrigR (Kusch and Davy,2022) on a predefined bounding box

- **Challenges:** The granularity of the data results in considerable large dataset.
    - Generating meaningfull agregations for municipalities is not straight foward.

Data and challenges

- While the previous slides provies the data used, aditional sources were identified in Brazil.

Health Indicators-Datasus (Ministry of health) (FORTHCOMING)
Sistema de Informações de Mortalidade (SIM) (Book, variables)
- Is one of the oldest health data sources in brazil (since 1975) unifing standarized data about death certificates.
- Indicator: mortality rates by age categories, causes and municipio (Admin2)

Sistema de Informações Hospitalares do SUS (SIH) (Book, variables)
- Provides administrative information (since 1981) about hospital admision, leght of stay and cost.
- Indicator: hospital admision by age categories, causes and municipio (Admin2).

Cadastro Nacional de Estabelecimentos de Saúde(Book, variables)
- Provides aditional information to create control variables like instractruture, hospital personel. Geocoded Hospitals.

Acces:
- Several ways to acces DATASUS, a package currently in devemolpent (Microdatasus, Saldanha et at) has some functionalities.
- R function to dowload and perform primary cleaning impemented (Download_SIM_files, Download_SIH_files).

Challenges: server calls* are restricted for IP outside of brazil. A VPN and several tries can help.

Municipality level variables
Sistema IBGE de Recuperação Automática (SIDRA)
Acces: The IBGE has made avalible and API to downloade the data. The R package: sidrar, interfase to acces the data.
Challenges: Requires the table index. Few variables for municiapilites
- I use population to showcase some back of the envelope calculations.

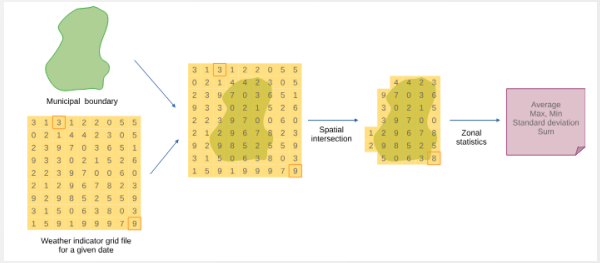
Feature engineering

Education outcome- Sistema de Avaliação da Educação Básica (SAEB) .

- I focus on yearly Municipality data for the 9th grade (last year of mandatory basic education)
- Considering:
  - Test scores (math and lenguage): a numerical score derived from a standarized test.
  - Passing rate: Number of student who pass / Number of students
- Note: Scores are “comparable” between year, but to ensure interpretation I standarized the yearly scores for SAEB in math and lenguage.

Rainfall shocks:

- I focused on zonal statistic indicators derived from RA5-Land (Saldanha et al,2024)
- To generate meaningful aggregations the authors suggest
  - Consider the areal overlap
  - Compute weighted operations based on percentage of area overlapping.
  - The authors made available daily municipality level data (1950-2022)



- For each m municipality, I aggregated rainfall to monthly precipitation (1980-2022).
- Computed the historical mean and standard deviation (μ_{jm}, σ_{jm})
  - For each j month, consider: $S^k_{mtj} = 1\{rainfall_{mjt} > \mu_{jm} + k\,\sigma_{jm}\}$
- We compute $NS^k_{mt} = \sum_j S^k_{mtj}$ and define that the **municipality experience a shock** if $S^k_{mt} = 1\{NS^k_{mt} \geq 1\}$

- A similar approach was previously used by Lima et al (2024).
  - Considering k standard derivations allows to examine the size of the shocks.

MODELING

Model 1:

- I exploited the exogenous climate shocks, their timing and geographical variation using the following specification:

$$y_{mt} = \beta_0 + \beta_1 S_{mt}^k + \boldsymbol{\theta_m} + \boldsymbol{\eta_t} + \epsilon_{mt}$$

Where: y_{mt} denotes the outcomes of interest in municipality m at time t (passing rate, standardized test results in math and lenguage)

S_{mt}^k denotes an indicator variable representing the rainfall shock, considering k standard deviations from the historical average.

$\boldsymbol{\theta_m}$ and $\boldsymbol{\eta_t}$ denotes municipality and year fixed effects.

Note: Future work could consider school or student level information to consider better control variables and more policy relevant question.

Model 2:

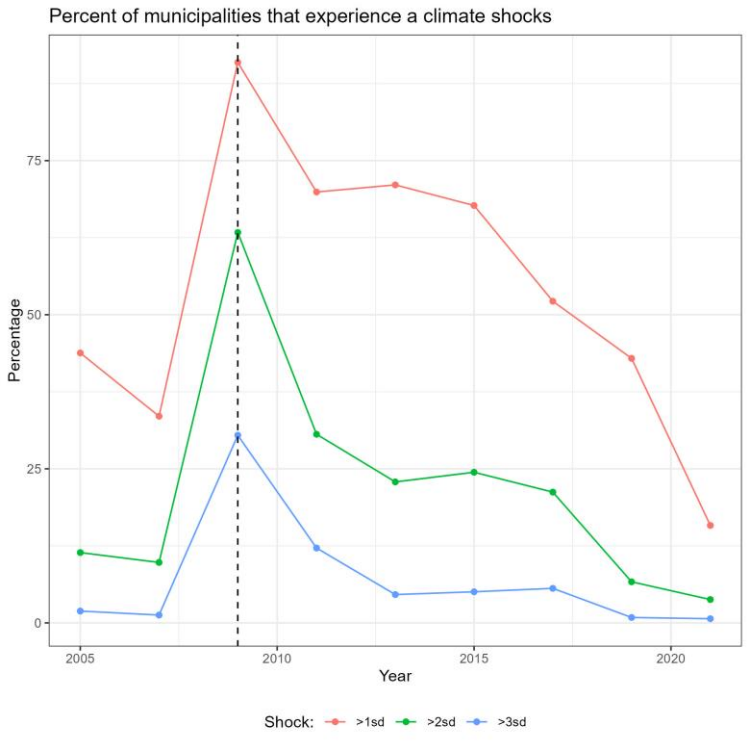
- Our secondary specification:

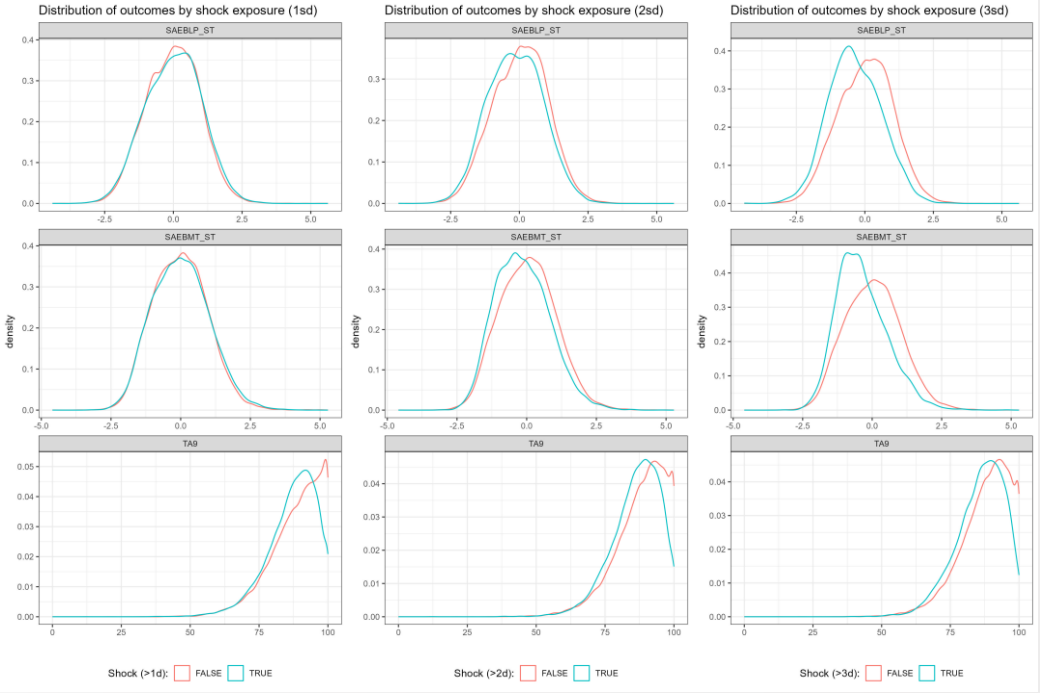
$$y_{mt} = \beta_0 + \beta_1 NS_{mt}^k + \boldsymbol{\theta_m} + \boldsymbol{\eta_t} + \epsilon_{mt}$$

Where: NS_{mt}^k denotes the number of months that can be consider extreme.

Exploiting the size of the shock and number of months that can be consider extreme allows to study heterogenous effects.

DESCRIPTIVE STATISTICS





- Municipalities facing larger shocks are associated with:
  - Lower passing rate
  - Lower average test results in math
  - Lower average test results in language

MODELING: RESULTS

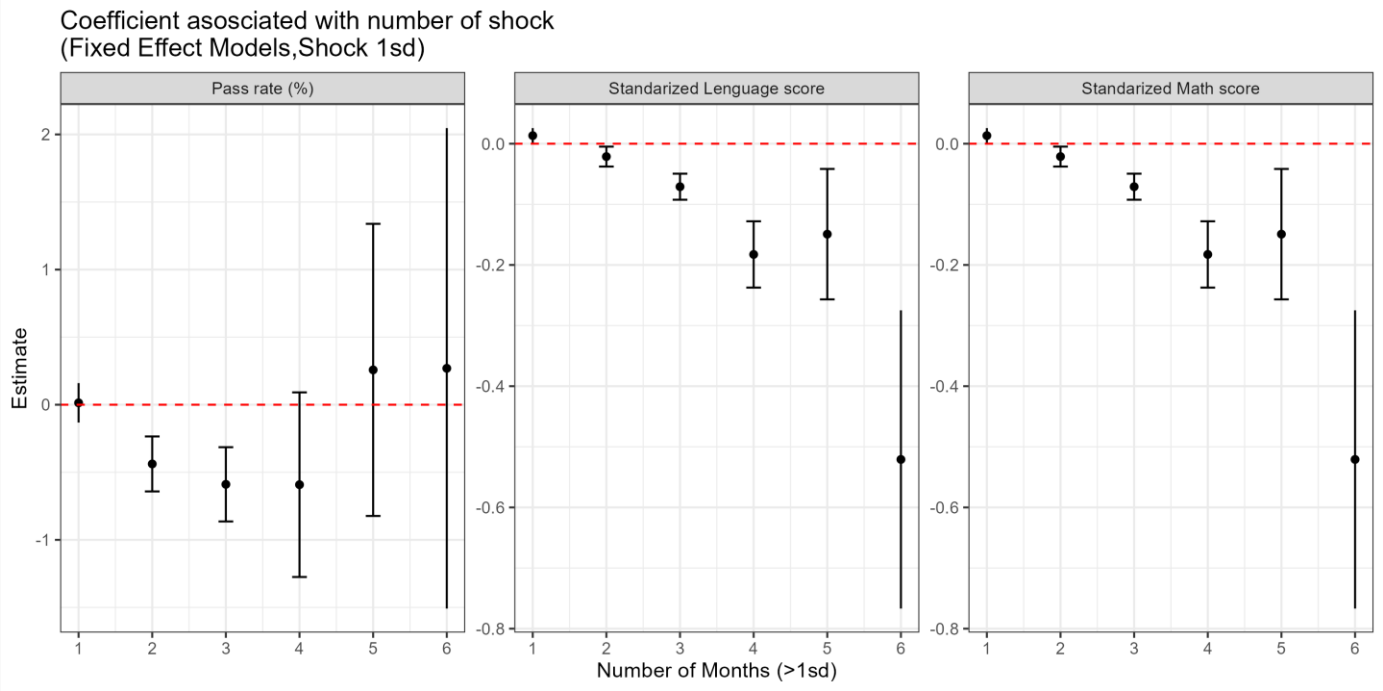
	Rainfall Shock (1sd)			Rainfall Shock (2sd)			Rainfall Shock (3sd)		
	SAEB Math std	SAEB Lenguage std	Pass (%)	SAEB Math	SAEB Lenguage std	Pass (%)	SAEB Math std	SAEB Lenguage std	Pass (%)
Rainfall Shock	-0.004	-0.001	-0.141*	-0.046***	-0.030***	-0.383***	-0.091***	-0.068***	-0.284*
	(0.006)	(0.006)	(0.068)	(0.007)	(0.007)	(0.083)	(0.010)	(0.011)	(0.127)
FE: Year	Y	Y	Y	Y	Y	Y	Y	Y	Y
FE: Municipality	Y	Y	Y	Y	Y	Y	Y	Y	Y
Num.Obs.	47645	47645	49180	47645	47645	49180	47645	47645	49180
R2	0.737	0.701	0.546	0.738	0.701	0.546	0.738	0.701	0.546

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

- Faced with a Rainfall shock, we can expect (statistical significantly):
  - Lower passing grades, even for lower shocks.
  - Lower average test results in both math and language.

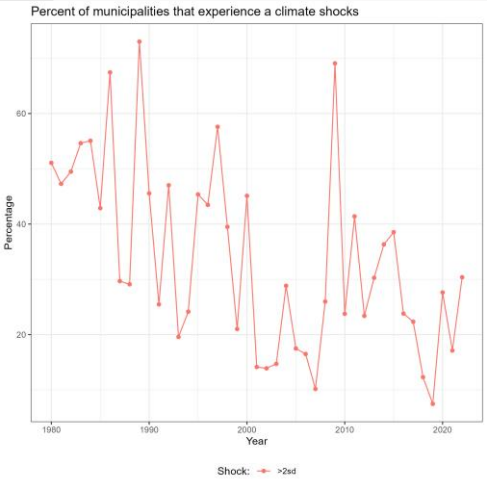
Note: as a robustness check, a model without 2009 was also estimated. Results were qualitatively similar.

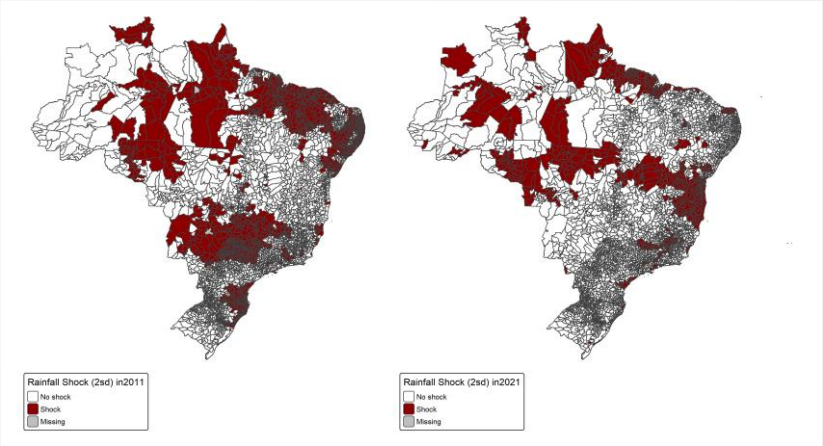
MODELING: RESULTS



- Faced with more extreme weather month:
  - We observe a larger decrease in test average test result.
  - There seems to be a u-shape for passing rate.

SHOCKS IN CONTEXT





- Rainfall shocks are not a one-time events for Brazilian municipalities.
- Municipalities that experiences this shocks are not always the same.

- The size of the shocks in terms of population is relevant, if we consider the cases of 2021:

Shocks (2Sd)	Population	% Population	Municipalities	% Municipalities
N	185168185	87%	4614	83%
Y	28100999	13%	953	17%

POLICY IMPLICATIONS

Previous results suggest that a rainfall shocks are associated with:

- A decrease in pass rate and learning outcomes.
- Larger impact when more extreme (and sustained) rainfall is observed.
- Climate shocks affect human capital accumulation and can hinder future economic growth.

Study of the mechanism and heterogenous effects can help design effective policy.

- Are shock associated with lower attendance rate? Yes, (WB,2024).
  - If yes, remote learning plans can reduce learning losses and mitigate climate challenges.
  - Developing digital solutions for education is on the IADB's focus area.
  - Before implementing any policy more information about access to remote learning is needed.

- Are shocks larger on communities/school with less infrastructure?
  - If yes, investment in infrastructure can help mitigate rainfall shocks.
  - However, this are more expensive and may require larger investments.

Study limitations:

- The current study uses municipality aggregated data, but micro data is available to inform our policies.
- Only one outcome was considered, study can be further expanded (heatwaves-hospital admission).
  - The study of other climate challenges, require more specialized modeling: Distributed lag (non-linear) models.

Policy Analytics Challenge

Effects of climate shocks on educational outcomes, an application to Brazil.

Federico Veneri

October 2025