Exploring Education in Brazil

University of Mannheim Department of Political Science

Research Design

Filippo Panfoli 2202955

filippo.panfoli@students.uni-mannheim.de | PanfoliF |
0009-0000-7861-8036 | https://github.com/PanfoliF

Replication materials / LaTeX sources: PanfoliF/Research-Proposal-RD-NEW

Version: v0.0 | **Last updated:** January 9, 2026

Keywords: Education Policy; Teacher Quality; Class Size; Student Schievement; School Failure; Panel Data; Educational Inequality; Brazil

JEL: C72; D82

Prepared for: Research Design, taught by Prof. Dr. Sean Carey

January 9, 2026

Abstract

This research proposal investigates the determinants of educational outcomes in Brazil thanks to a causal-inference framework. Motivated by the persistent empirical uncertainty surrounding what works in education, the study exploits Brazil's federal structure and rich subnational variation to estimate the effect of school inputs on student performance and progression.

Empirically, the project relies on the Brazilian Education Panel Databases (Huberts et al., 2025). The analysis builds on a causal directed acyclic graph (DAG) and a two-margin conceptualization of outcomes. Teacher quality is operationalized using measurable proxies–teachers' education–and is completed by indicators of school infrastructure and a set of controls.

Identification proceeds in two steps. First, fixed-effects panel regressions with municipality, state, and year effects absorb time-invariant local heterogeneity, broader state-level differences, and common temporal shocks. Second, a Difference-in-Differences design leverages institutional and policy variation in education governance to sharpen causal interpretation. The project tests whether improvements in teacher quality raise median achievement. It also examining whether observed score gains align with changes in failure rates. Analyzing these margins, the study aims to provide policy-relevant evidence on which educational investments most plausibly translate into genuine learning gains rather than shifts in selection or promotion dynamics.

Introduction

This paper proposes a research project on educational data. The aim is establish causality between variables of interest using state of the art causal inferene techniques. The Brazilian context offers interesting sources of variation among data as well as granular observations. These are key factors in providing rubust empirical analysis.

Motivation: It is well established (Mokyr, 2005)[1] that density in the upper tail is crucial for innovation and diffusion of modern technology. Nonetheless, under many aspects there is little knowledge of what causes good educational outcomes (**<empty citation>**). Many factors concur in formig students, thus scholars have always had hard times in distinguishing chains of causes and effects. The proposed paper could provide some more evidence.

Data: such a high aim requires the best tools. The author put a lot of effort in the data collection step, many institutional websites have been surfed until the best source of data come out. The Brazilian Education Panel Databases (Huberts et al., 2025), appears the best source of observation for the proposed task. Not only it provides researchers with granular data, it also comes from a reliable and well know institution, namely IDB (Interamerican Development Bank).

Methodology: If we are to establish causality in a credible way, we need to use causal inference methodologies. The author will provide the reader with extensive explantion of employed identification strategy in the methodology section (Methodology). The following lines briefly anticipate that section: the identification strategy selected for our aim is splet in two: first, due controls come along with two regressions. Secondly, a Difference in Difference application of these regressions, should isolate with more efficacy the causal effec of independent vari-

[1]Prof. Joel Mokyr won the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 2025.

ables on dependent one. An appropriate review of the literature provides for suitable control variables.

Policy value: The following pages are intended to create value for policy makers. Putting aside cumbersome vocabulary and contex related dictions, they should inform policy making in the field of education in a simple and transparent manner. The results of the analysis may open new ways of investing in education and shed light on wiser policy aims.

Hypothesis

My goal is to distinguish two channels: standardized test performance (*extensive margin*) and failure rates (*intensive margin*). Comparing extensive and intensive margin helps determine which mechanisms affect the test score.

The analysis in the next sections will test the following hyposthesis:

$$\textit{Hp n}^{o}1\textit{: Teacher quality} \Uparrow \Rightarrow \textit{median student performance} \Uparrow$$

$$\textit{Hp n}^{o}2\textit{: Teacher quality} \Uparrow \Rightarrow \textit{worst students performance} \Uparrow$$

Research Question

To what extent do changes in teacher quality–proxied by teachers' educational attainment–causally affect students' educational outcomes in Brazil (measured as standardized test achievement and failure rates)?

1 Literature Review and Historical Setting

Knowledge and insights on the historical context and institutional settings are drawn from Encyclopedia Britannica (Ball, James, et al., 2026; Ball, Schneider, et al., 2026), Glossario of Atlas Geográfico Escolar (CEON, n.d.) and Southey (2012).

A preliminary analysis of the literature on education in developing countries, highlighted a study from Turmena and Bitencourt (2022). The journal article constitutes the milestone of this article and serves as main reference for the literature on education in Brazil.

2 Data Collection

All projects begin with data collection, which is a crucial step. However it takes a lot of time and effort.

In order to select the best source, many datasets have been explored and many institutional websites have been visited. Potential data sources included IPUMS ("IPUMS Online Data Analysis System", n.d.) and the Instituto Brasileiro de Geografia e Estatística ("Portal Do IBGE", 1967). Additionally, aggregated data may be retrieved from other sources[2]. Eventually, an article from Rubiane Daniele Cardoso de Almeida et al. (2023) offers panel data on some demographic aspects.

The Brazilian Education Panel Databases (Huberts et al., 2025), which covers the period from 1996 to 2015, was selected as main source of data.

[2](Instituto Brasileiro de Geografia e Estatística (IBGE), n.d.-a, n.d.-b, n.d.-c, n.d.-d, n.d.-e, n.d.-f, n.d.-g, n.d.-h, n.d.-i, n.d.-j, n.d.-k, n.d.-l, n.d.-m, n.d.-n, n.d.-o, n.d.-p, n.d.-q, n.d.-r, n.d.-s, n.d.-t, n.d.-u, n.d.-v, n.d.-w, n.d.-x, n.d.-y, n.d.-z, n.d.-aa, n.d.-ab, n.d.-ac, n.d.-ad, n.d.-ae; Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), n.d.-a, n.d.-b).

3 Methodology

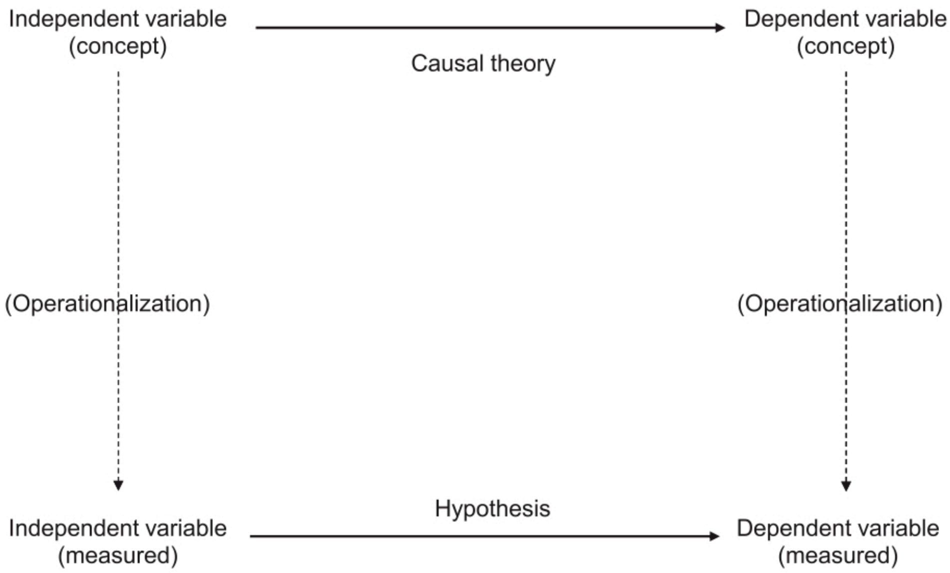

Figura 1: *Figure 1.2 From theory to hypothesis (Kellstedt & Whitten, 2018, chap. 1)*

Figure one illustrates the path from causal theory to an empirically testable hypothesis and it clearly shows the difference between what we mean and what we can measure. At the conceptual level, a causal model theorises that an independent concept influences a dependent concept–an abstract claim about how the world works. Yet concepts are not directly observable, this means that empirical analysis requires operationalization. This process can be thought as the translation of each concept into measurable indicators. The dashed vertical arrows capture this translation, emphasizing that measurement is not automatic but a set of choices that must be theoretically justified. Once concepts are operationalized, the researcher can formulate a hypothesis linking the measured independent variable to the measured dependent variable, represented by the lower horizontal arrow. The figure supports measurement validity: weak indicators can lead to precise estimates of the wrong relationship. In this proposal, we operationalise good education through test score (or failure rates) and teaching quality through techers education.

3.1 Data Generating Process

Causal inference is fundamentally about the data-generating process (DGP): the (usually unobserved) mechanism that maps underlying conditions, choices, and shocks into the data we observe. In causal questions, we are not merely interested in how variables move together in the realized dataset; we want to know how outcomes would change if we were to intervene-if the DGP were run under a different input, such as a different policy, treatment, or institutional setting.

The difficulty is that we observe the DGP only once, under the conditions that actually occurred. The outcomes generated under alternative conditions–the counterfactual realizations of the same DGP–are not in the data. A causal model is therefore essential because it provides an explicit representation of how the DGP operates and, crucially, what is assumed to remain invariant when we imagine an intervention. Under those assumptions, causal inference uses observed data to learn about the parameters or structural relationships of the DGP, and then leverages them to predict what the outcome would have been in the counterfactual scenario. In this sense, the DGP is central to causal inference because it is the bridge between what we observe and the unobserved "what if" quantities we seek to estimate.

3.2 DAG

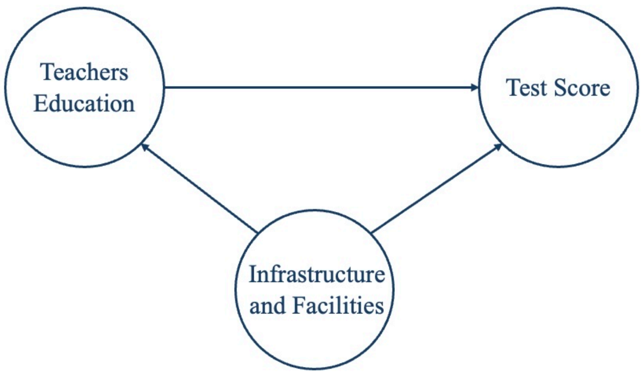

Figura 2: *Causal DAG.*

Figure 2 presents a directed acyclic graph (DAG) that summarizes the study's theoretical framework and makes the intended causal interpretation of the empirical model explicit. In a DAG, each node represents a variable of interest and each arrow denotes a hypothesized direct causal relationship, with the direction indicating the assumed flow of influence. Here, the graph encodes the idea that teachers' education may affect students' test scores both directly and indirectly through school infrastructure and facilities.

A key purpose of drawing a DAG is to clarify where confounding variables may arise. A confounder is a variable that influences both the explanatory variable and the outcome, thereby creating a non-causal association that can bias naïve regression estimates. In this setting, infrastructure and facilities can act as a confounder for the relationship between teachers' education and test scores if, for instance, better-resourced schools both attract more educated teachers and produce higher achievement. To isolate the causal effect of teacher education on test outcomes, the empirical specification therefore includes relevant confounders as controls, consistent with the adjustment set implied by the DAG.

3.3 Regressions

To establish causality between the dependent and independent variables, the analysis will employ causal inference techniques. As a possible solution to the identification problem the study will provide results from a Difference-in-Difference. This technique is able to isolate the effect produced by the introduction of the laws in the variables that approximate education quality. Nonetheless, without appropriate control variables no identification strategy is reliable. Following best practices of the political science field, only a deep analysis of the literature will provide for suitable control variables.

3.3.1 Extensive Margin

The first of the two regression presented in the paragraph looks at the extensive margin. Therefore, it investigates the relationship between teachers quality and test scores. The equation used is the following:

$$Y_{smt}^{score} = \beta_0 + \beta_1 TQ_{smt} + \beta_2 INFRA_{smt} + \gamma X_{smt} + \mu_m + \lambda_t + \varepsilon_{smt}$$

Teachers quality is measured as teachers education.

Identification also relies on:

-*Municipal FE*: this is a way to control for possible unobserved variables that might bias the analysis. To the eyes of statistitians this is a mere intercept that captures the mean value for a town.

-*State FE*: the same concern we had for the municipal level, motivates the use of FE at the state level. However, concern comes along with a great fortune[3]. What Cunningham (2021, chap. 2, p. 46) said about USA perfectly suits the Federation of Brazil.

-*Year FE*: since our dataset (Huberts et al., 2025) offers several years, we will exploit time variation too. The methodological solution to make use of panel data is again FE. In fact, thanks to this instrument, we are able to isolate the variation among years and discard the magnitude of variation in a signle year.

This strategy isolates the effect of teachers quality on students test scores, which is commonly refered to as *outcome variable.*

Controls: The vector of controls is X_{smt}, while γ is the vector of coefficients. It represents the effects of control variables on our outcome variable.

[3]Cunningham (2021, chap. 2, p. 462) says: "I have a bumper sticker on my car that says "I love Federalism (for the natural experiments)". [. . .] United States is a never-ending laboratory. Because of state federalism, each US state has been given considerable discretion to govern itself with policies and reforms. Yet, because it is a union of states, US researchers have access to many data sets that have been harmonized across states, making it even more useful for causal inference."

The selected controls are:

1. School Characteristics & Size
   - Functional status: TP_SITUACAO_FUNCIONAMENTO, STATUS
   - Management type: TP_DEPENDENCIA, with dummies for federal, state, municipal, private.
   - Location: TP_LOCALIZACAO, URBANA (urban/rural).

2. Infrastructure and Facilities

Availability of resources: - Libraries (IN_BIBLIOTECA), labs (IN_LABORATORIO_INFORMATICA, IN_LABORATORIO_CIENCIAS), sports field (IN_QUADRA_ESPORTES), computers/internet.

- Management type: TP_DEPENDENCIA, with dummies for federal, state, municipal, private.

Utilities: electricity, water, sewage (both public and general availability).

Quantities: NU_SALAS_EXISTENTES, NU_SALAS_UTILIZADAS, NU_COMPUTADOR.

3. Teachers & Staff
   - Staff totals: NU_FUNCIONARIOS, PROFESS, PROFFUNDTOT.
   - Teacher education indicators: PCPROFMED, PCPROFSUP, etc.
   - Derived metric: EDUCTEACH (average years of teacher education).

3.3.2 Intensive Margin

The second regression presented in the paragraph looks at the intensive margin. Therefore, it investigates the relationship between teachers quality and rates of failure. The equation used is the following:

$$Y_{smt}^{failure} = \beta_0 + \beta_1 TQ_{smt} + \beta_2 INFRA_{smt} + \gamma' X_{smt} + \mu_m + \lambda_t + \varepsilon_{smt}$$

Teachers Quality is measured teachers education. Again, identification relies on municipal FE, state FE and year FE. This regression uses the same controls as above.

4 Limitations

From the very first page, in this paper I adopted a transparet approach. Therefore, this section highlights all limitation of the analysis conducted so far. In this way, constructive criticism become active part of the scientific process.

I noticed three weakeness of the identification strategy:

- Omitted Variable Bias.
- Provide tests for observable implications in as many as possible narrow, focused, controlled circumstances, as suggested by Clarke (2005).
- It maybe that I included "control variables that are a consequence of the key IV" (Sieberer, 2007, chap. 8, pp. 163–182).

Omitted Variable Bias is a common burden of all empirical scientists. There is no way to include all possible sources of variation in the identification strategy and hampers the possibility of identifiyng the "perfect" causal mechanism. However, there are painkiller to this issue. They do not come from causal inference or econometrics. Instead, they come from institutional and qualitative knowledge of the topic under study.

Thus, we can reassure the reader of the reliability of the empirical result deepening the qualitative knowledge of the problem at stake.

A second important limit of this study comes from the fact that the more we generilise the results, the less we can be sure of what we state. Empirical tests have increadible internal validity. Nonetheless, they all suffer of a noticeable restriction, namely external validity. The identification strategy is solid only when we test the causal mechanism in data it was though to work for. When we apply the identification strategy to external data, coming from a similar data generating process, the results tend to become fuzzy, and the ground of the analysis becomes slippery.

Eventually, we have another source of doubts about the research proposal. Let's imagine a state investing more in school with better tests scores or with lower failure rates. In this way, school managers or regional governments are forced to act in accordance with an incentives scheme.

However, if this were the case for Brazil, there would be a huge bais in the identification strategy used so far. This bias is due to the fact that the outcome variable (test score for example) shapes investments in the school infrastracture, that at the same time influences students performances.

Further exploration of the institutioinal setting can shed light on this issue.

Conclusion

This research proposal sets out to turn a familiar policy question–what actually improves educational outcomes?–into an empirical design for the Brazilian case. Building on a panel framework, the project is motivated by a core inferential ambition: identify credible causal effects.

Two conceptual contributions structure the empirical strategy. First, the proposal distinguishes between two outcome margins: first, standardized test performance (the "extensive margin"), and secondly, failure rates (the "intensive margin"). This distinction is meant to reveal whether improvements in observed performance reflect genuine learning gains, changes in selection into testing, shifts in promotion standards, or other institutional mechanisms. Second, the project operationalizes "teacher quality" through measurable proxies–teachers' education and class size–then embeds these in a causal framework that explicitly anticipates confounding, as illustrated by the proposed causal DAG.

From the standpoint of methods emplyed, the design combines fixed-effects panel regressions with a Difference-in-Differences (DinD). The fixed-effects architecture (municipality, state, and year) is intended to absorb time-invariant heterogeneity across places and common national shocks across years.

The DiD is proposed as a second line of support: by focusing on changes induced by reforms, it aims to isolate the marginal effect of shifts in educational inputs from other background trends. In policy terms, this is the kind of design that can speak to decision-makers: its targets are realistically adjustable–teacher qualifications, class size, infrastructure. The design asks how changes in these targets map onto outcomes that matter for human capital formation.

At the same time, the proposal is transparent about what could go wrong. Omitted variable bias remains the classic specter: even with extensive fixed effects and controls, some determinants of educational performance (local political capacity, school leadership quality, household shocks, informal labor dynamics) may vary over time and correlate with both inputs and outcomes.

External validity is the second constraint: estimates that are internally persuasive in one institutional setting, may not apply cleanly to other contexts or periods.

A third risk is reverse causality: performance itself may shape investment, producing a feedback loop in which outcomes partially determine the "treatments" meant to explain them. This is especially salient in contexts where funding formulas, administrative incentives, or reputational competition allocate resources based on observable performance.

These limitations do not invalidate the design; they specify the conditions under which it can succeed. In practical terms, they point toward concrete refinements that would elevate the proposal from a promising design to a highly persuasive empirical study.

The most important is institutional context: mapping the precise timing and scope of the targeted reforms and documenting whether–and how–funding and administrative responses are conditioned on performance is crucial for the credibility of the study.

Eventually, the "two-margins" structure proposes a richer interpretation of mechanisms: if teacher quality improves test scores but not failure rates (or vice versa), that divergence is analytically meaningful and can adjudicate between competing narratives about learning, retention policies, and classroom sorting.

In sum, the proposal's value lies in its attempt to convert a broad and politically consequential question into a design that can generate defensible causal claims. Brazil is not merely an interesting case; its federal structure, policy variation, and the availability of granular panel data make it a powerful empirical laboratory for education policy research. If the project successfully integrates its econometric strategy with careful institutional analysis it can offer findings that are both academically informative and directly relevant to policy choices.

Acknowledgements

Artificial intelligence-based tools were employed solely to improve linguistic clarity and grammar. No AI system contributed to the development of the research questions, theo-

retical framework or conclusions presented in this paper.

Bibliography

Books

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press.

Kellstedt, P. M., & Whitten, G. D. (2018). *The Fundamentals of Political Science Research* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108131704

Southey, R. (2012). *History of Brazil*. Cambridge University Press.

Articles

Clarke, K. A. (2005). The Phantom Menace: Omitted Variable Bias in Econometric Research. *Conflict Management and Peace Science*, *22*(4), 341–352. https://doi.org/10.1080/07388940500339183

Mokyr, J. (2005). The intellectual origins of modern economic growth. *The Journal of Economic History*, *65*(2), 285–351. https://doi.org/10.1017/S0022050705000112

Rubiane Daniele Cardoso de Almeida, Benjamin M. Tabak, & Tito Belchior Silva Moreira. (2023). Demographic aspects and regional income convergence in Brazil: A panel data approach. *CEPAL Review*, *2023*(139), 197–210. https://doi.org/10.18356/16840348-2023-139-10

Turmena, L., & Bitencourt, J. C. (2022). A reforma de 1ř e 2ř graus de 1971 e a reforma do ensino médio de 2017: Algumas aproximações. *Educ. Escr.*, *13*(1), e43895. https://doi.org/10.15448/2179-8435.2022.1.43895