Zika virus and declining birth rates in Brazil: Replicating a causal difference-in-differences study with linear and nonlinear models

Ella Bartt[1]

UC San Diego, Masters in Computational Social Science.

UC San Diego
SCHOOL OF SOCIAL SCIENCES

Introduction

- Zika virus (ZIKV) is a mosquito-borne disease found in mosquitoes (*Aeges aegypti)* known to cause a range of birth defects and autoimmune illnesses. [6]
- In March 2015, Brazil reported a number of clustered Zika cases, and WHO soon declared a Public Health Emergency in 2016. [7]
- Reported birth rates declined, with many choosing to postpone or abort pregnancy due to the disease's birth defects. [5]
- The temporally localized nature of the outbreak can be treated as an observational epidemiological experiment, in which we may use a Difference-in-Differences (DiD) approach to measure the disease's impact on birth rates in Brazil. [5]
- This work replicates the results found in the first application in Taddeo et al. [5], in which the authors use linear and nonlinear models for a DiD approach to various public health risks in Brazil.
- I test the effectiveness of adding a socioeconomic covariate to the original paper's estimation model.

Data and Assumptions

The Brazilian states Pernambuco and Rio Grande do Sul are treated as a treatment and control group, respectively. Pernambuco reported 65% of total Zika cases in Brazil, while Rio Grande do Sul reported 0 cases. The states are geographically far away, preventing spillover in treatment. [5]

The dataset includes:

- Municipality and state
- Rate: Birth rate per 1000 people
- BirthSchool: live births to mothers with 8+ yrs. Schooling
- Population of Municipality
- GDP per capita by municipality, 2014 *(added in replication)* [1]
- Rate_wSchool, Birth rate per 1000 people to mothers with 8+ yrs. schooling *(added in replication)* [1]
- Prop_wSchool: Proportion of overall births in Municipality to mothers with 8+ yrs. schooling (0-1) *(added in replication)*

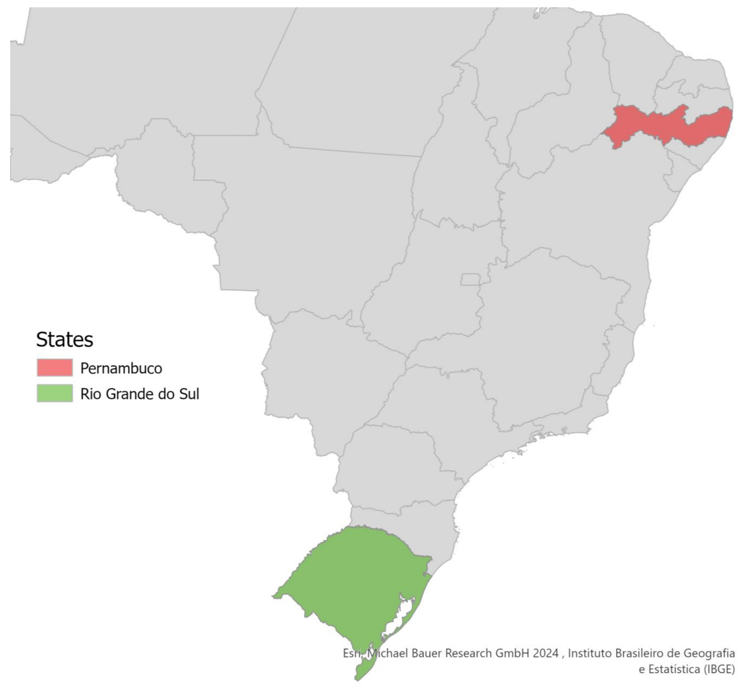

Fig 1. Brazilian states Pernambuco (PE) and Rio Grande do Sul (RS). Red = Zika outbreak group (treated), green = control group. (Map credit: Kiana Vaghefi).

Materials and Methods

The states satisfy all necessary assumptions for a causal DiD framework [paper]. The two states must first show similar (parallel) birth rate trends with no anticipatory treatment difference to accurately measure the treatment effect (fig 2). They satisfy these assumptions, as seen below, so we can proceed to measure the causal parameter τ.

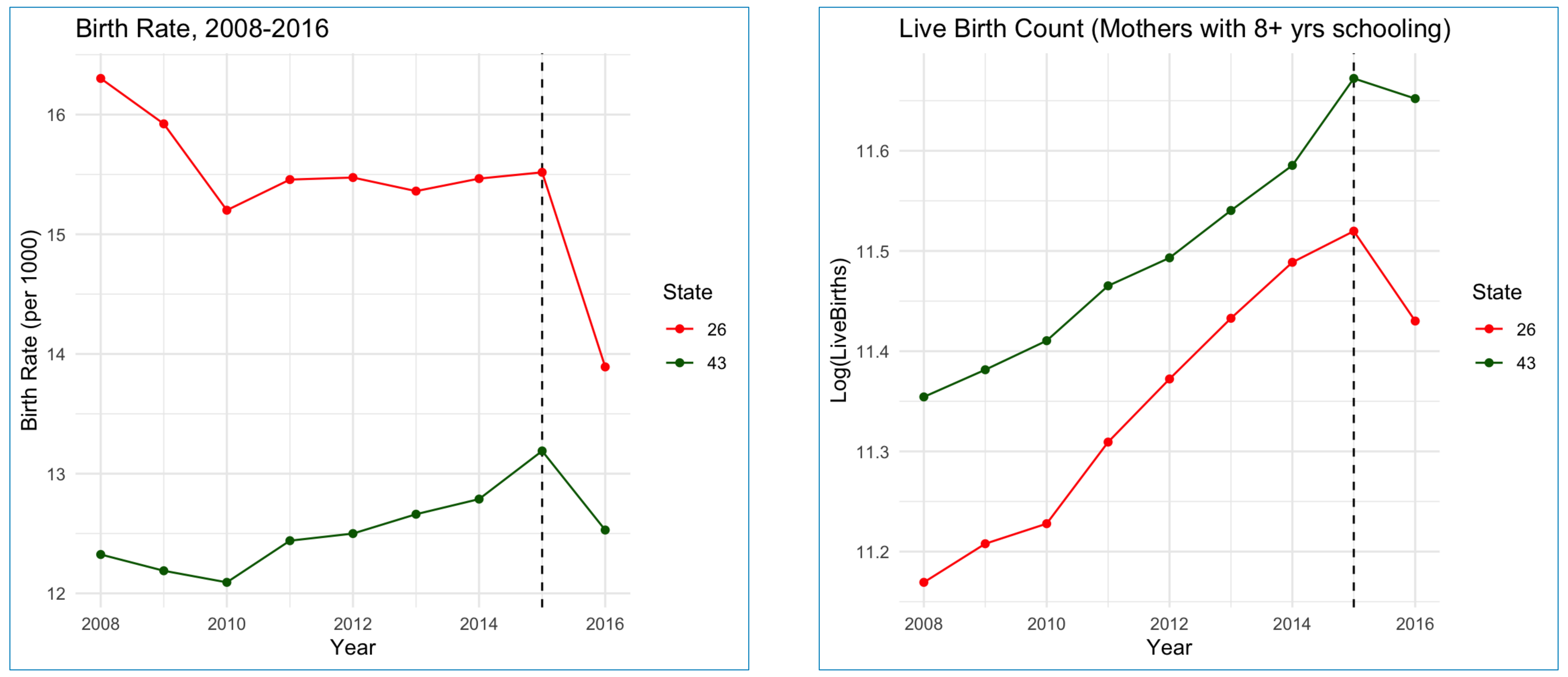

Fig 2. Birth rate (left) and live birth count (log) from mothers with 8+ yrs education (right). State 26 = Pernambuco (PE) and 43 = Rio Grande do Sul (RS).

Fixed effects models are deployed to control for time-invariant differences between municipalities and to extract clustered standard errors.

- t = time period
- T = treatment (0 for RS, 1 for PE)
- Y = outcome: BirthRate for linear and Poisson, BirthSchool for Negative Binomial
- X = prop_wschool, controlled for t=0 to avoid post-treatment bias.
- U = Municipality

Balance in X: The proposed covariate, prop_wschool, is imbalanced in the dataset between the treatment and control groups. Adding IPW weighting brings the balance test statistic down from 5.83 to 0.27 (<0.25 is best).

Other potential covariates: GDP per capita in each municipality was also tested as a potential covariate, since Lobkowicz et al. [3] cites a strong correlation between relative poverty and Zika outbreak areas. The results showed high correlation with other variables and weak instrumental variable measures, so this has been omitted from the estimate.

Results Table

		Linear				Poisson				Negative Binomial		
		Coefficient	95% CI	SE (clustered)	R^2	Coefficient	95% CI	SE (clustered)	Adj. Pseudo R^2	Coefficient	95% CI	SE (clustered)
Original Paper	$\hat{\tau}$	-1.21	[-1.56, -0.85]	0.161	0.804	-0.087	[-0.10, -0.07]	0.008	0.992	-0.131	[-0.160, -0.107]	0.013
	GATT	**-1.21**	[-1.56, -0.85]	0.161	0.804	**0.916**	[0.905, 0.932]			**0.877**	[0.852, 0.899]	
With Covariate (prop_ wschool)	$\hat{\tau}$	-0.846	[-1.19, -0.50]	0.174	0.843	-0.072	[-0.11, -0.04]	0.143	0.998	-0.14	[-0.164, -0.116]	0.012
	GATT	**-0.846**	[-1.19, -0.50]	0.174	0.843	**0.931**	[0.896, 0.961]			**0.869**	[0.849, 0.890]	

Table 1. Model results. Standard errors shown are clustered at the municipal level. $\hat{\tau}$ is the coefficient for each model that estimates the GATT. The link functions between the coefficient value and the GATT are the identity function (linear) and log (Poisson and Negative Binomial).

Results

Original study results:

- **Linear*:*** -1.21 births per 1000 people in PE compared to RS → -7.8% of birth rate from 2015 *due to outbreak.*
- **Poisson:** 0.916 Incidence Rate Ratio (IRR) → 2016 birth rate in PE decreases to 91.6% of the 2015 birth rate *due to outbreak.*
- **Negative Binomial:** 0.877 IRR → 2016 PE live birth count for educated mothers reduced to 87.7% of the 2015 count *due to outbreak.*

→ Models agree: ~8% decrease in overall birth rate is due to the outbreak (even if total decrease is more than 8%).

Modified results (with covariate and IPW):

- **Poisson** and **Negative binomial**: GATT estimates are very similar to original measures.
- **Linear**: Smaller estimated decrease that **does not agree** with the Poisson and Negative binomial models using the same covariate and IPW components.

Conclusion

→ Adding a socioeconomic covariate and IPW weighting *do not* improve the GATT estimate in the case of this study. In fact, this worsens the accuracy of our estimates, with greater standard errors and wider confidence intervals.

→ The results suggests that adding a socioeconomic variable is already adequately controlled for by including the treatment status and municipality fixed effect in our estimates.

→ We can conclude that the state of Pernambuco in Brazil experienced an additional 8% drop in birth rate due to the effects of Zika virus.

References

[1] *Base dos Dados—Conjuntos de dados abertos*. (n.d.-a). Base dos Dados. Retrieved June 2, 2026, from https://basedosdados.org
[2] *Humanity's deadliest predator is evolving to be even more dangerous | CBC The Nature of Things*. (2026, April 9). CBC.
[3] Lobkowicz, L., Power, G. M., De Souza, W. V., Montarroyos, U. R., Martelli, C. M. T., de Araùjo, T. V. B., Bezerra, L. C. A., Dhalia, R., Marques, E. T. A., Miranda-Filho, D. de B., Brickley, E. B., & Ximenes, R. A. de A. (2021). Neighbourhood-level income and Zika virus infection during pregnancy in Recife, Pernambuco, Brazil: An ecological perspective, 2015–2017. *BMJ Global Health*, *6*(12), e006811.
[4] *Renda média domiciliar per capita—Notas Técnicas*. (n.d.). Retrieved May 25, 2026.
[5] Taddeao, M., Amorim, L., & Aquino, R. (2022). Causal measures using generalized difference-in-difference approach with nonlinear models. *Statistics and Its Interface*, *15*, 399–413. https://doi.org/10.4310/21-SII704
[6] *Zika virus*. (n.d.). Retrieved May 29, 2026, from https://www.who.int/news-room/fact-sheets/detail/zika-virus
[7] *Zika virus disease outbreak 2015-2016*. (n.d.). Retrieved May 30, 2026, from https://www.who.int/emergencies/situations/zika-virus-outbreak

Acknowledgements: Thank you to Prof. Guoer Liu and Annie Lin for their support in POLI 273. Thanks to the CSS classmates and program.