January 8, 2025

Dear Dr. Bell,

Thank you very much to you and the reviewers for your consideration of our manuscript, titled, “Disruption to Test Scores after Hurricanes in the United States.” We appreciate the reviewers’ insightful comments and feedback. You will find each of their comments included, in turn, below, followed by our bolded author responses.

Reviewer 1

This research is an important step to identify educational attainment after storms. The limitations you provide impact the robustness of the article and explanation of impact of storms on test scores. The results are not unexpected: Educational scores are highly variable by state, racial/ethnic, and sociodemographics. Further research into how return-to-school policies in place allow for better access to education and their long-term impact on test scores.

**Thank you very much for acknowledging the importance of this research. We agree with you that the results are not unexpected given geographic and sociodemographic variability in the United States and that further research is needed to enhance school system and student recovery in the aftermath of hurricanes and other disaster events.**

Reviewer 2

It is not clear how confounders were selected and they seem to only be factors

that would be associated with the outcome and not the exposure of interest.

Adjusting for factors that are only associated with the outcome may attenuate

observed associations so I recommend the authors reconsider their adjustment

variables. This paper may be helpful with thinking about that.

Schisterman, Enrique F.a; Cole, Stephen R.b; Platt, Robert W.c. Overadjustment

Bias and Unnecessary Adjustment in Epidemiologic Studies. Epidemiology

20(4):p 488-495, July 2009. | DOI: 10.1097/EDE.0b013e3181a819a1

It would be very useful to incorporate any available information on school

closures into this analysis. Grouping together all counties that have high wind

speeds as being exposed may smooth over associations you wish to observe

because the impacts may be much worse for counties that experienced actual

school closures during the time period as a result of the hurricanes. With this, I

think the authors should also explain more why wind speed is the measure they

decided to use as a proxy for hurricane exposure. Wind speed is only one small

part of a hurricane’s impacts and it is no surprise that by using that as the

exposure, the majority of the counties included in the analysis are mostly directly

along the coast because hurricanes lose speed as they are on land longer.

However, it is important to note that in examples like Hurricane Florence, larger

impacts were observed further inland in North Carolina due to extreme flooding

that occurred as a result of the hurricane. These counties that were impacted

with things like school and hospital closures further inland due to hurricane-

related flooding would be included in your control group in this analysis when you

use wind speed as the exposure determinant.

Similarly, I think the authors should consider doing additional analyses examining

the impacts of repeated exposures as some of these may be areas that

experience more than one hurricane in the time period studied.

A lot of descriptive information that is useful for interpreting these results is

missing and making it difficult to interpret these results. The authors should

report baseline standardized test scores. The authors state that scores go up in

florida and down in North Carolina, but the baseline scores are not reported so it

is hard to understand what this means in comparison to each other. Additionally,

if I understand the methods correctly, the authors assume that the baseline

scores are the same across all states, but the baseline scores are not reported

so this cannot be confirmed.

In the figure 2A the authors report associations that seem to be associations with

demographic groups and test scores which is not the objective of this research

study. I do not understand why these are reported as these should be covariates,

not the exposure of interest. I recommend refencing this paper on the table 2

fallacy.

Daniel Westreich, Sander Greenland, The Table 2 Fallacy: Presenting and

Interpreting Confounder and Modifier Coefficients, *American Journal of*

*Epidemiology*, Volume 177, Issue 4, 15 February 2013, Pages 292–

298, https://doi.org/10.1093/aje/kws412

The authors should provide more background literature on why this outcome

matters for health. Especially given the comments made in the limitation section,

it is known that standardized test scores do not do the best job giving a fair

assessment of student’s performance and it is unclear what the impacts are on

health of children. More background literature on what this means for health

would be useful if this paper is to be published in a health journal.