

Causal Inference and COVID: Contrasting Methods for Evaluating Pandemic Impacts Using State Assessments

CITATION: Shear, B. R. (2023). Causal inference and COVID: Contrasting methods for evaluating pandemic impacts using state assessments. *Educational Measurement: Issues and Practice*, 42(1), 99–109. <https://doi.org/10.1111/emip.12540>

BACKGROUND: State assessment results from Spring 2021 have been widely used to quantify the impact of the COVID-19 pandemic on student learning. This paper describes three threats to making valid causal inferences about pandemic impacts based on cross-year comparisons of average test scores: measurement artifacts affecting the comparability of scores, secular trends, and changes to tested populations.

DATA: The paper uses statewide administrative test score data for grades 3–8 math and English Language Arts (ELA) tests in Colorado from 2017, 2019, and 2021 to illustrate and contrast different methods being used to study the impacts of the pandemic.

RESEARCH DESIGN: The paper compares three different statistical adjustments that have been used to study changes in state test scores during the pandemic to make inferences about pandemic impacts: the Fair Trend, baseline student growth percentiles, and multiple regression with demographic covariates.

KEY FINDINGS:

- Changes to the tested population of students in 2021 relative to prior years complicates the comparison of average scores across years among Colorado students. Direct comparisons of average scores before and after the pandemic that do not account for population changes may either under or overestimate the impact of the pandemic on student test scores.
- The three statistical adjustment methods used lead to similar inferences about the average impacts statewide, and suggest larger score declines in math than in ELA. The estimated effect sizes ranged from -0.07 to -0.21 standard deviations across grades, subjects, and estimation method. These findings are consistent with analyses in other states.
- The three methods can lead to substantially different inferences about which specific demographic groups of students were most impacted.

CONCLUSION: State assessment results provide a useful data point for understanding pandemic impacts on student learning. However, statistical adjustments should be used to ensure that cross-year comparisons do not provide misleading results. Selection of statistical adjustment methods should align with the intended inferences and policy goals of using state assessment data. The considerations in this paper also apply to analyses based on more recent years of state assessment data. Finally, state assessments provide only one source of information about pandemic impacts. Researchers should continue to use additional data sources to provide a comprehensive understanding of pandemic impacts and to implement and evaluate recovery efforts.