Appendix E: Methodologies for comparing student learning growth

We analysed student Progressive Achievement Test (PAT) scores and demographic data to estimate the impact of receiving tutoring on student achievement. We performed a type of matched analysis, using propensity score weighted regression of PAT gain on students' participation in the initiative.

Matched analysis

Overview	Propensity score analysis is a technique that attempts to estimate the effect of a treatment (in this case, the initiative) by accounting for potentially confounding covariates that predict receiving the treatment. **Achievement analysis** We performed propensity weighted regression analysis of PAT gain (the difference of 2023 and 2022 scores) on tutoring participation and 2022 PAT score. We split by year level (years 3 to 10) and domain (reading and maths). **Absence analysis** Similarly, we performed propensity weighted regression analysis of the increase in student absences (the difference between 2023 and 2022 absences) on tutoring participation and 2022 absences. We broke this down by year level.
Data	We used 2022 and 2023 term 4 PAT scores and yearly absences for students between years 3 and 10 in 2023. For each student, we took their semester 2, 2022 teacher judgement data. We sourced demographic data for each student from the 2023 August census.
Outline of steps	**Step 1: propensities** Using the above datasets, we chose a variety of covariates that may influence participation in tutoring. This included individual demographic factors as well as school-level variables. For each year level and domain, we performed a logistic regression of tutoring participation on: • disadvantage status • Aboriginal status • English as an Additional Language status • disability status • 2022 PAT scale score • indicator if student was below the expected level in teacher judgement in that domain.

Step 2: weighted least squares regression
We estimated the average treatment effect of tutored students. To do this, we assigned each student a weight based on their propensity of being selected for tutoring.

We assigned tutored students a weight of 1, and non-tutored students a weight of p_i/(1-p_i). We clipped weights to the 5th and 95th percent quantiles within each year level and domain to remove the effect of extreme weights.

We performed the following regressions with errors clustered by school:

- PAT_GAIN_i = beta_0 + beta_1 TLI_i + beta_2 SCORE_2022_i;
- DIFFERENCE_ABSENCES_i = beta_0 + beta_1 TLI_i + beta_2 2022_ABSENCES_i.

We took beta_1 to be the treatment effect of tutoring.

Sensitivity analysis

We tested a variety of combinations of covariates in the propensity regression.

We also tested various functional forms of the regression equation, such as difference score and residual change.

We performed our analysis with an overall tutoring flag and separate literacy and numeracy tutoring flags.

In addition to the average treatment effect on the tutored group, we also estimated the average treatment effect for the whole population.

Throughout all sensitivity analysis, our analysis arrived at similar results and conclusions.