Empirical Project 2

# Do Smaller Classes Improve Test Scores? Evidence from a Regression Discontinuity Design

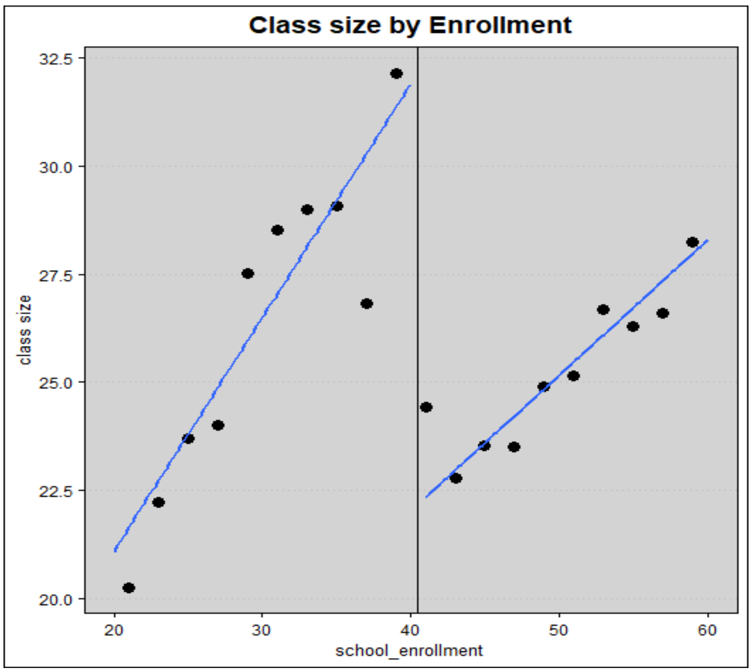
## Importance of causal estimation

Isolating the causal effect of class size is not as simple as comparing outcomes across different class sizes. Running a regression of class size on test scores would result in a correlational estimate, i.e., how they move in tandem, but would not be causal. To obtain causal estimates we would need to hold every other possible effect on test scores constant which would be nigh impossible with just a simple regression as we likely have not captured all these effects in the data available. To overcome this obstacle, we can use quasi-experimental data analysis which is the next best way to isolate specific effects outside of controlled experiments which in this setting would be a long and expensive process to implement.

## Evidence from Israel

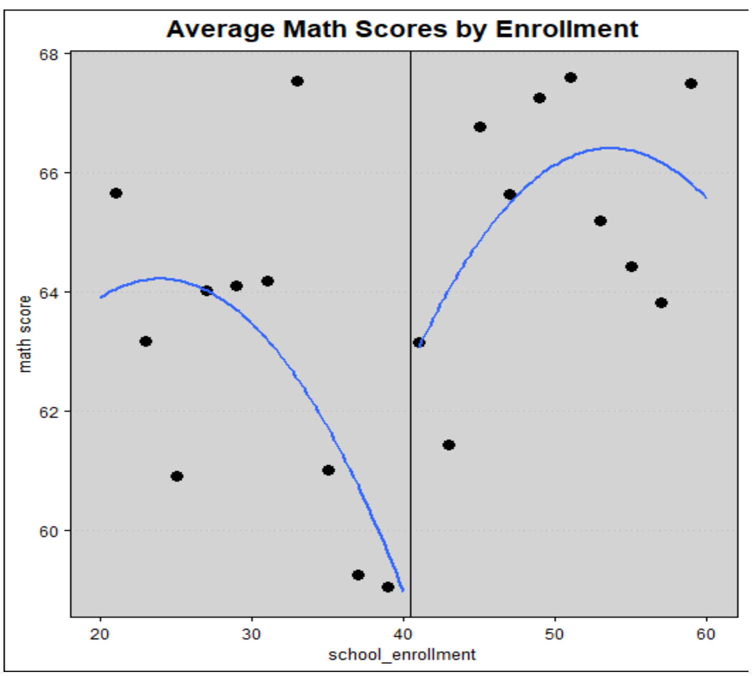
Luckily, there is data readily available from Angrist and Lavy (1999) which examines the relationship between class size and test scores in math and language using a quasi-experimental technique known as Regression Discontinuity. Regression Discontinuity examines differences in trends across some exogenous treatment. In this setting, class enrollment in multiples of 40 is used as a cutoff; In Israel, there is a rule that no class can have a student-teacher ratio of greater than 40-1. So, a school that moves from 40 to 41 students must split their grade into 2 classes. This rule is used as an independent (exogenous) change to class size to view the differences in test scores for classes that are nearing the threshold (large classes) and those that have been split (small classes). The significant effects in the data come from 4$^{th}$ and 5$^{th}$-grade students. 3$^{rd}$-grade students were also studied but no significant effects were found which may or may not be attributed to changes in teaching not related to class size (Angrist). Before discussing the evidence, I will briefly mention the underlying assumptions that make regression discontinuity work. It must be the case that other covariates that affect class size must be “smooth” across the area of interest. That means that there should only be a discontinuous jump as pictured below caused by the exogenous mechanism of enrollment affecting class size. If there are other mechanisms besides enrollment enacting similar effects on class size (and are not controlled for in the regression equation) then the discontinuity model breaks down.

Class size and enrollment



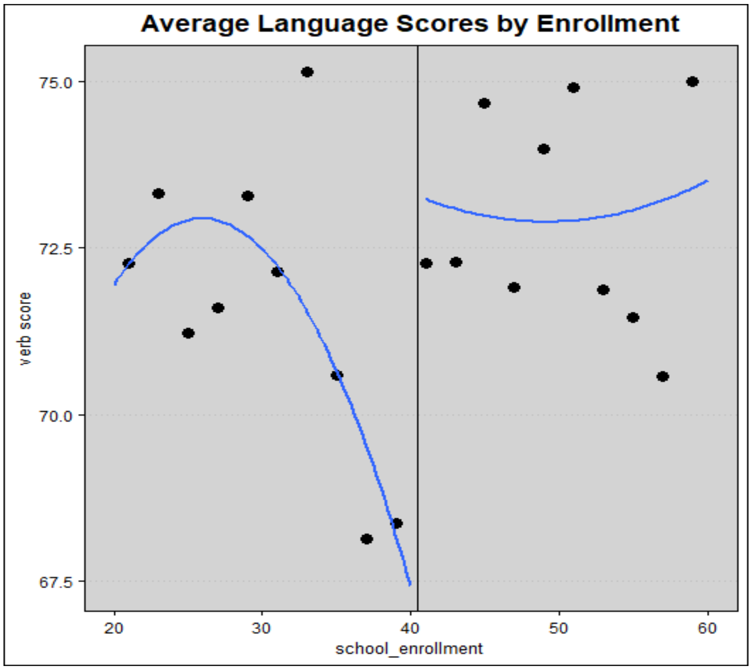

Above is a binned scatter plot showing class size across school enrollment. A binned scatter plot simply aggregates the data into averaged points rather than each having its own point, all graphs in this memo have 20 entries per point. There is quite a significant change in the trend line on either side of the 40-student enrollment cutoff with class sizes returning. All this is showing is evidence of the rule in action. To formalize the relationship, I also reran the discontinuity regression to get the official estimate of crossing the threshold. Going from 40 to 41 students in a grade decreases class size by around 11 students on average. You notice that class size never actually reaches the 40-student limit and caps out at an average of 32, this is because many schools have the resources to reduce class size before the 40-student limit is reached. Additionally, class size does not always reduce by half as might be expected if a new class is being created. There are two primary reasons for this, the first being enrollment in this setting is expected enrollment before the start of the school year, actual enrollment might be low enough to not require splitting. The other possible reason is that not many classes actually reach the 40-student threshold and are mandated to be split so reduction effects are rarely ever at their maximum (Angrist).

Math scores and enrollment




Unlike the relationship between class size and enrollment, the relationship between class size (through enrollment) and math scores does not appear to be linear or positive. This is to be expected as the assumption is that less individual attention from the teacher should lead to worse performance. On average, scores increase by 3.4 percentage points from the average of 61.

Language scores and enrollment




The relationship between language scores and class size show a similar pattern to that between math and class size. Language scores appear to have a much more consistent increase in scores after passing the cutoff. That would imply that language acquisition is more sensitive to class size. That being said, the quantitative results show the opposite, crossing the enrollment threshold increases scores by 2.6 percentage points on average, which is almost a whole percentage point smaller than the effect on math test scores. The average score for language tests is already almost 10 percentage points higher than those for math so perhaps there are simply decreasing marginal returns to smaller class size the better baseline performance gets; in other words, students are already decently informed in language arts and do not need/benefit from more teacher attention as much as in math.

## Proposed changes

A reduction in class size from 40 to 35 would likely have a positive effect on math and language test scores. The proposed change in class size is half that found in the Israeli data, that being said, the relationship between class size and scores does not appear to be linear. The majority of the gain might be concentrated around the cutoff as seen in the graphs. This could lead to an increase in scores larger than half of that found in the data. I would estimate a safe range to be 50-75% of the observed effect when moving from 40 to 35 students. If instead, class size was reduced from 20 to 15 students, I do not think there would be as large gains as from 40 to 35. Again, looking at the graphs there do not seem to be large positive changes in scores moving between smaller class sizes, only from a large class to a small one as well as large negative effects from a medium to a large. I would urge caution in expecting the same results as there might be reduced marginal returns the smaller class size gets. One might also be worried about teacher quality as classes get substantially smaller, requiring a much larger teaching base. Luckily, Jepsen and Rivkin (2009) find little to no evidence that there are negative effects of having new and uncertified teachers fill the increased demand. There are short-run costs that will be borne by the students who get first-year teachers with no experience or certification, but these effects disappear over time; that is not to say that those short-run effects are not to be considered. There is also heterogeneity in schools' ability to accommodate the increased demand for teachers; low-income schools with larger shares of minority students tend to get larger proportions of new teachers and will face steeper short-run costs.

I would recommend that policy be enacted to reduce class size, especially if improved long-term outcomes are desired. In the long run, teacher quality effects fall away leaving only the benefits of smaller classes and short-run costs may be alleviated if reductions are implemented in a way so that no one class would face inexperienced teachers in consecutive grades. That could be to do all grades at once which would be extremely taxing, or to work from the top down one grade at a time. Improved test scores would imply more skill retention and higher GPAs which could lead to higher college acceptance and attendance which would lead to higher earnings which in a way pays for itself via more taxes overall making class reduction policy worthwhile.

# References


Angrist, Joshua D., and Victor Lavy. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." The Quarterly Journal of Economics, vol. 114, no. 2, 1999, pp. 533-575.

Jepsen, Christopher, and Steven Rivkin. "Class size reduction and student achievement the potential tradeoff between teacher quality and class size." Journal of Human Resources 44, 1 (2009): 223-250.