Class Size Effects in Asia: A Multi-Cutoff Regression Discontinuity Design

Authors and Affiliations:

Wei Li
College of Education
University of Alabama
307C Carmichael Hall, Tuscaloosa, AL 35487
Phone: (205) 348-8087
Email: wli59@ua.edu

Spyros Konstantopoulos
College of Education
Michigan State University
461 Erickson Hall, East Lansing, MI 48824
Email: spyros@msu.edu

Class Size Effects in Asia: A Multi-Cutoff Regression Discontinuity Design

Background

In the past decades, many Asian countries have enacted class size reduction (CSR) policies. For instance, Japan, Hong Kong, Singapore, and Taiwan have set maximum class size rules for kindergarten, primary school or secondary school grades aiming to increase student achievement. Although the literature has documented evidence about class size effects in the U.S., there is quite limited evidence about class size effects in Asian countries. There are only a few exceptions that focused on eighth-grade achievement with mixed findings (e.g., Pong & Pallas, 2001; Wossmann & West, 2006). Also, prior studies have used older data and thus could not inform the effects of current CSR policies. In a word, our review of the literature points to lack of evidence of class size effects in early grades using latest national probability samples in Asian countries.

Purpose

Our study is designed to address these limitations. Specifically, we examine the effects of class size on student mathematics achievement in Japan, Hong Kong, Singapore, and Taiwan using data from the 2011 and 2015 fourth grade sample of the Trends in International Mathematics and Science Study (TIMSS). We utilize a fuzzy multi-cutoff regression discontinuity design (RDD) coupled with the instrumental variable (IV) approach to facilitate causal inference.

Data

TIMSS is the largest international database that measures trends in mathematics and science achievement at fourth and eighth grades every four years. We selected Hong Kong, Japan, Singapore and Taiwan as our analysis sample because they had known rules about maximum class size limits for fourth graders. Table 1 provides details about the selected countries and maximum class size rules in 2011. Appendix provides detailed explanations of variables that we used in our analysis.

Research Design

The assignment of students and teachers to classrooms is not random typically, and thus class size could be correlated with unobserved factors related to students, parents, and teachers' characteristics. In that case, the estimated class size effect produced from ordinary least square regression would be biased. To overcome this potential shortcoming and facilitate causal inferences, the current study evaluates class size effects through RDD that utilizes the maximum class size rules (see Angrist & Lavy, 1999). Specifically, we compute the average class size for each school through the following equation

$$f_i = \frac{E_i}{int(\frac{E_i - 1}{rule}) + 1},$$

where E_i denotes the enrollment in grade four in a school; f_i denotes the computed average class size based on the maximum class size rule; *rule* denotes the maximum class size rule in each country; for any positive number *n*, the function *int(n)* is the largest integer less than or equal to *n*. Figure 1 shows plots of the computed average class size and school and grade specific average actual class size against enrollment in fourth grade in 2011. When schools follow the maximum class size requirement well (e.g., Japan), this rule generates a discontinuity in terms of the actual and the computed average class size as school enrollment increases. In other words, the class size displays an up-and-down form as the school enrollment increases, and there were multiple cutoffs when school enrollment reached to multiples of maximum rules.

As shown in Figure 1, the maximum class size rules were not strictly implemented in practice, and thus it is a fuzzy RDD. Researchers usually use IV to estimate the treatment effects in a fuzzy RDD (Lee & Lemieux, 2010). The IV methods have been widely used in previous work to evaluate class size effects through parametric linear regressions (i.e., two-stage least squares or 2SLS) with an arbitrary bandwidth (e.g., +/-5 around the cut offs) and conventional inference methods (e.g., Angrist & Lavy, 1999; Li & Konstantopoulos, 2016). However, the exact function forms around the cutoffs are unknown. Therefore, local polynomial estimators and robust nonparametric inference are now preferred and routinely employed in applied research (see Calonico, Cattaneo, & Titiunik, 2014). What is more, one unique feature of our RDD is that there were multiple cutoffs and thus current methods of calculating the optimal bandwidth (e.g., Imbens & Kalyanaraman, 2012) cannot be directly adopted. RDD with multiple cutoffs is very common in applied work, and thus researchers usually normalize the running variable (e.g., school enrollment) and estimate a pooled RDD treatment effect (Cattaneo, Keele, Titiunik, & Vazquez-Bare, 2016). Cattaneo et al. (2016) showed that the pooled estimator could be interpreted as a weighted average treatment effect across cutoffs. One advantage of this normalizing-and-pooling method is that it allows researchers to compute an optimal bandwidth easily since there is only one cutoff after the normalizing. However, this normalizing method has not been used in class size effects research. To address these limitations, our study applies the normalizing-and-pooling method and estimate the causal effects of class size on student achievement. Specifically, we use the nonparametric local polynomial regression with the MSE-optimal bandwidth and bias-corrected robust inference approach using the STATA routines *rdbwselect* and *rdrobust* (Calonico, Cattaneo, Farrell, & Titiunik, 2017).

Results

We are still working on the analysis of TIMSS 2015 data. In this section, we only provide our preliminary results from TIMSS 2011. Table 2 presents descriptive statistics and samples sizes.

The first stage regression results are summarized in Table 3. Only Japan and Taiwan had IVs whose F-statistics were larger than 10. For Hong Kong and Singapore, the IVs were weak. When the IV is weak, the estimate and its standard errors are not reliable (Stock, Wright, & Yogo, 2002). Therefore, we only discussed the IV results from Japan and Taiwan, which are summarized in Table 4. Results from the parametric 2SLS indicate no class size effects. However, the results from nonparametric local polynomial regression show that, with a reduction in class size of one student, the average mathematics achievement increases by 1.52 points in Japan; also, the average difference in mathematics achievement between students in smaller classes and those larger classes was 23.07 points, which is equivalent to 0.23 standard deviation (SD). This effect is considerable because the average small class size effect in Project STAR was about 0.20 SD. These results indicate that CSR is a promising policy option in Japan.

Appendix

Variables:	Description (TIMSS Variable Name)
Student Variables	
Mathematics Achievement	Set of five overall mathematics score plausible value variables (ASSMAT1–ASSMAT5)
Female	Binary indicator for the student whose gender is female (ASBG01)
Age	Student age at the time of testing (ASDAGE)
Speaking the Tested Language at Home	Binary indicator for the student who spoke the tested language at home “always or almost always” (ASBG03)
SES: Books in the Home	Number of books in the home (ASBG04)
SES: Items in the Home	Sum of eleven wealth-related household possessions variables (ASBG05A–ASBG05K)
Positive Affect to Mathematics	Average of five self-reported student's affect to mathematics variables, with negatively-worded items reverse-coded (min = 1, max = 4; ASBM01A–ASBM01E)
Parents Asked What the Student was Learning in School	Binary indicator for the parents asking the student what he/she is learning in school every day or almost every day (ASBG07A)
Student Talked about the Schoolwork with Parents	Binary indicator for the student talking about the schoolwork with parents every day or almost every day (ASBG07B)
Parents Made Sure the Student Set Aside Time for the Homework	Binary indicator for the parents making sure that the student sets aside time for the homework every day or almost every day (A
Parents Checked if the Student Did the Homework	Binary indicator for the parents checking if the student does the homework every day or almost every day (ASBG07D)
Teacher/Classroom Variables	
Class Size	Number of students in the classroom (ATBG12A)
Classroom SES: Books	Average number of books in the home
Classroom SES: Items	Average number of items in the home
Proportion Female	Proportion of female students in a class
Average Students' Positive Affect to Mathematics	Average self-reported student's affect to mathematics in a class
Teacher Experience in Years	Teacher's year of teaching (ATBG01)
Teacher Completing Post-Secondary Education	Binary indicator for the teacher who completed post-secondary education (ATGB04)
Female	Binary indicator for the teacher who is female (ATBG04)
Instruction Time	Time spending teaching mathematics to the students in the class per week (ATDMHW_M)
School Variables	
Percent Disadvantaged Students	Set of four indicators for categorical percentage of economically-disadvantaged students (ACBG03A)
Percent of Students Having Tested Language as Native Language	Binary indicator for categorical percentage of the students having tested language as their native language more than 90% (ACB
Students Having Early Numeracy Skills	Set of four indicators for categorical percentage of the students entering the primary grades with early numeracy skills (ACDG
City Size	Set of six indicators for categorical city population (labels = 0–3,000, 3,001–15,000, 15,001-50,000, 50,001-100,000, 100,001–500,000, greater than 500,000; ACBG05A)
Income Level of the School's Immediate Area	Set of three indicators for the income level of the school's immediate area (ACBG05C)
Grade 4 Enrollment	Total enrollment of fourth graders in the school (ACBG02)

Reference:

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 114*(2), 533-575.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295-2326.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). rdrobust: Software for regression discontinuity designs. *The Stata Journal, 17(2)*, 372-404.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G., & Keele, L. (2016). Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics*, *78*(4), 1229-1248.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), 933-959.

Lee, D. S., & Lemieuxa, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), 281-355.

Li, W., & Konstantopoulos, S. (2016). Class size effects on fourth grade mathematics achievement: Evidence from TIMSS 2011. *Journal of Research on Educational Effectiveness, 9(4),* 503-530.

Pong, S. L., & Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad. *Educational Evaluation and Policy Analysis, 23*(3), 251-273.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics, 20*(4), 518-529.

Wossmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review, 50*(3), 695-736.

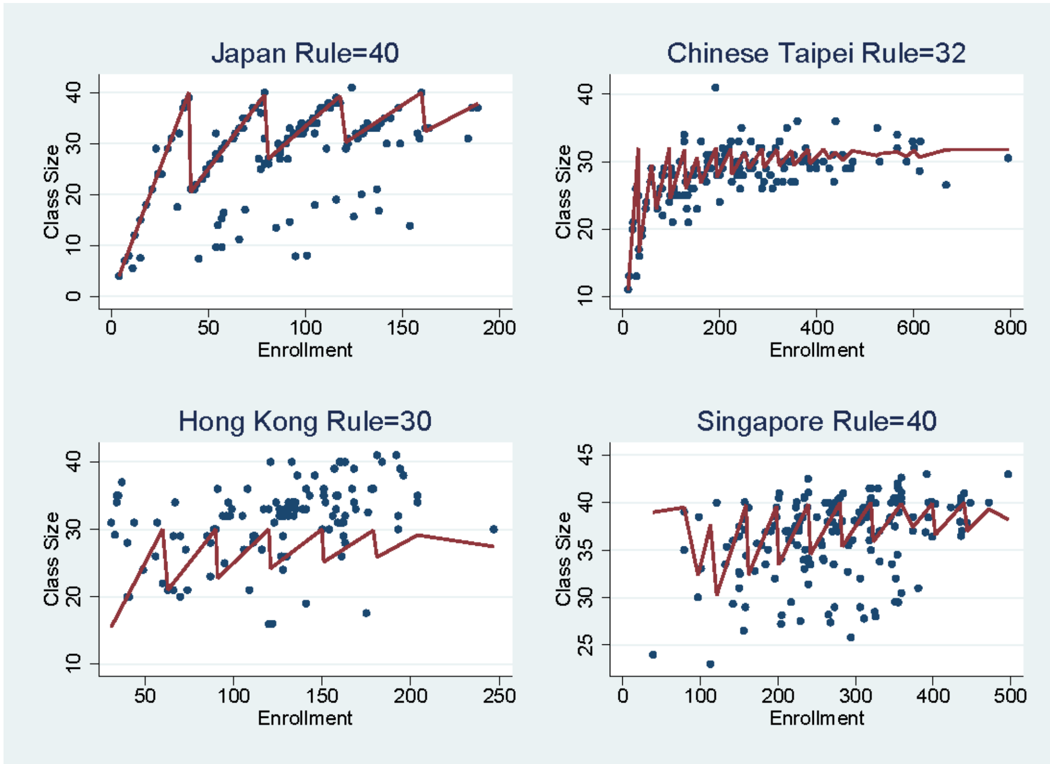

Figure 1. Computed and reported average class size by grade enrollment
Note: The straight line represents the computed average class size using rules about maximum class size. The dots represent reported average class sizes for each school.

Table 1. Maximum Class Size Rules

Hong Kong	30
Japan	40
Singapore	40
Taiwan	32

Table 2. Descriptive Statistics

	Hong Kong	Japan	Singapore	Taiwan
Student Variables				
Mathematics Achievement	601.61	585.37	605.79	591.21
	(66.42)	(72.31)	(78.18)	(73.22)
Class Size	32.13	28.90	37.00	28.03
	(5.38)	(8.54)	(5.57)	(4.60)
Grade 4 Enrollment	105.78	58.61	273.05	108.96
	(48.65)	(41.50)	(88.83)	(120.63)
Sample Size				
Schools	136	149	176	150
Classes	137	149	351	155
Students	3957	4411	6368	4284

Note: Weighted means are reported. Standard deviations are in parentheses.

Table 3. First Stage Estimates

	Mathematics			
	Japan	Hong Kong	Singapore	Taiwan
Parametric (2SLS)				
IV-Computed class size	0.82*	-0.50*	0.28	0.73*
	(0.11)	(0.16)	(0.24)	(0.11)
F-Statistic for IV	54.17	9.80	1.37	44.09
Nonparametric: Actual class size				
IV-Computed class size fell below the cutoffs	-29.37*		0.98	-0.87
	(5.71)	N.A.	(5.51)	(1.42)
F-Statistic for IV	26.47		0.03	0.37
Nonparametric: Binary smaller classes indicator				
IV-Computed class size fell below the cutoffs	2.00*			0.14
	(0.26)	N.A.	N.A.	(0.10)
F-Statistic for IV	59.96			2.02

Note. N.A. indicates the STATA routine "*rdbwselect*" could not compute the local polynomial bandwidth.
Standard errors in parentheses; * $p<0.05$

Table 4. IV Estimates of Class Size Effects

	Mathematics	
	Japan	Taiwan
Parametric (2SLS)		
Class size	-0.81	-0.83
	(0.46)	(1.23)
Nonparametric: Actual class size		
Class size	-1.52*	N.A.
	(0.67)	
Nonparametric: Binary smaller classes indicator		
Class size	23.07*	N.A.
	(7.73)	

Note. N.A. indicates IV is too weak to compute the valid IV estimate of class size effects.
Standard errors in parentheses; * $p<0.05$