NOTES

IDENTIFYING CLASS SIZE EFFECTS IN DEVELOPING COUNTRIES: EVIDENCE FROM RURAL BOLIVIA

Miguel Urquiola*

Abstract—This note implements two research designs that attempt to isolate the effect of class size on achievement. A first strategy focuses on variation in class size in rural schools with fewer than 30 students, and hence only one classroom, per grade. Second, an approach similar to Angrist and Lavy's exploits regulations that allow schools with more than 30 students in a given grade to obtain an additional teacher. Both designs suggest class size negatively affects test scores.

I. Introduction

A large body of research suggests that class size has only small effects on educational outcomes.[1] Nonetheless, economists have generally suspected that it may matter in certain contexts, and with nonlinear effects that are hard to isolate empirically.[2] For example, Lazear (2001) presents a model in which a causal negative effect of class size on achievement is obscured as schools assign more disruptive students to smaller classes. More generally, if class size is endogenously determined, cross-sectional variation may yield biased estimates of its effects.

Angrist and Lavy (1999) highlight a particular way in which such bias can arise. Namely, class size is often positively related to enrollment, which may in turn be correlated with socioeconomic status (SES) and achievement. These channels seem to be empirically relevant: Lavy (1995) finds them in Israel, and points out that the Coleman (1966) report described a similar finding in the United States.

This paper illustrates that such correlations are also relevant in Bolivia, and presents two empirical strategies to identify the effects of class size. A first approach focuses on variation in class size in rural schools with fewer than 30 students, and hence only one classroom, per grade. This eliminates these schools' ability to allocate weaker students to smaller classes. Additionally, they are likely to be in remote areas and enjoy a monopoly in educational supply, reducing the scope for parental choice. This may account for why—at least on observables—enrollment SES correlations are weak among these schools. In short, this strategy focuses on a situation where it is impossible to separate class size from enrollment, because they are in fact equal. Instead, it relies on the assumption that in this sample, enrollment-SES correlations may be less important.

Second, an approach similar to Angrist and Lavy's (1999) exploits regulations that allow schools with more than 30 students in a given grade to obtain an additional teacher. This generates a discontinuity in the enrollment class-size relation, making it feasible to identify the latter's effect using the resulting class size function as an instrumental variable. In short, this strategy accepts the presence of an enrollment-SES link, but identifies a situation where the effects of class size can potentially be separated from those of enrollment.

Both designs suggest larger classes result in lower test scores, with results that are consistent with specific nonlinearities in this effect.

II. Data

The data come from Bolivia's Educational Quality Measurement System (SIMECAL), a testing program that samples private and public institutions. The schools and grades covered vary from year to year, so that the system does not provide panel data.

Table 1 describes the cross section used—that collected for the third grade in 1996. In addition to language and math scores, it includes data on students' SES, their classrooms, and their teachers. Although there is information for higher grades, we focus on the third grade because enrollment rates are higher at earlier ages, and the endogeneity of attendance is therefore less relevant. Specifically, more than 93% and 100% of rural and urban children, respectively, eventually complete the third grade.[3]

III. Empirical Challenges

Angrist and Lavy (1999) emphasize that if enrollment and class size are positively related, and enrollment is in turn positively correlated with SES, cross-sectional data might even suggest that raising class size improves achievement, as illustrated by simple comparisons of means in table 1. To see how this arises, suppose that the j^{th} class in school k has size

$$C_{jk} = \frac{E_k}{n_k}, \tag{1}$$

where E_k is enrollment, n_k is the number of classes, and all variables are grade-specific.[4]

The clearest association between enrollment and class size arises in sparsely populated rural areas, where schools often have only one class per grade. Figure 1 illustrates this by plotting class size against enrollment. For very small institutions, the two variables are identical, and a positive association remains over most ranges of enrollment.

The simplest way to see how enrollment SES correlations originate once again involves comparing small and large schools. Because the former are more often in rural and the latter in urban areas, welfare differences between these two (frequently observed in developing countries) will produce such correlations. For instance, table 2 explores the correlation between enrollment and SES in different samples, and shows, that children in larger schools

Received for publication November 14, 2000. Revision accepted for publication May 24, 2005.

* SIPA and Economics Department, Columbia University.

For valuable feedback, I thank David Card, Kenneth Chay, Darren Lubotsky, David Stern, and two anonymous referees. I am also grateful to the Universidad Católica Boliviana, where most of this research was carried out. All remaining errors are my own.

[1] See Hanushek (1999) for a review.

[2] For instance, see Pritchett and Filmer (1997) and Krueger (2000).

[3] Urquiola (2000).

[4] The data provide both the measure (1) and the actual number of pupils in each class. These are identical for 51% of schools, and for 80% they are within 5 percentage points of each other.

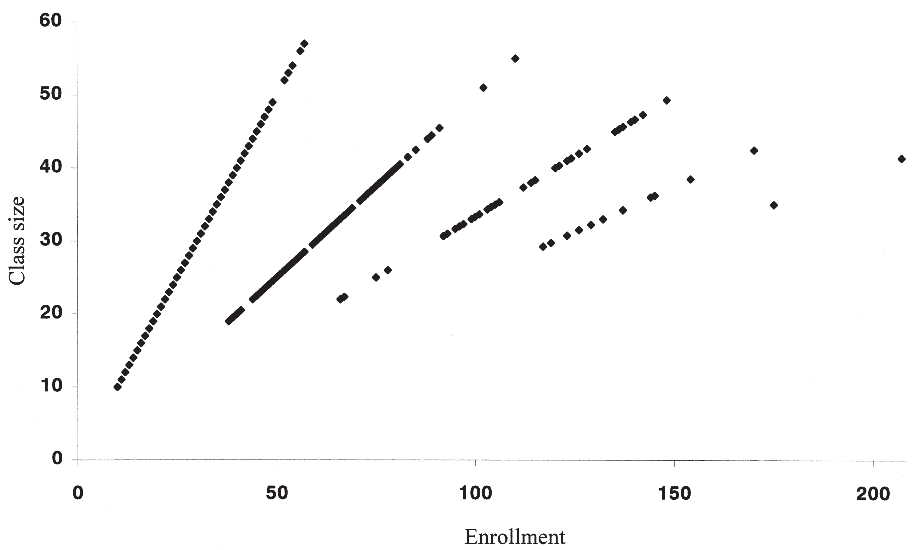
The Review of Economics and Statistics, February 2006, 88(1): 171–177

TABLE 1.—DESCRIPTIVE STATISTICS BY TYPE OF SCHOOL

	Total (1)	Rural Public (2)	Urban Public (3)	Urban Private (4)
Test Scores				
Mean language score	49.8 (9.3)	47.2 (9.3)	49.8 (8.4)	58.2 (8.2)
Mean math score	53.2 (9.2)	50.9 (9.7)	53.2 (8.4)	60.8 (6.9)
Sample size*	10,018	3,594	5,285	1,139
Class and Teacher Characteristics				
Mean third-grade class size	29.9 (9.9)	23.3 (8.4)	35.5 (6.4)	37.4 (9.8)
Mean third-grade enrollment	63.3 (44.1)	38.2 (27.4)	91.7 (39.0)	77.6 (59.6)
Mean number of third-grade classes	1.6	1.3	2.2	1.4
% of classrooms with a blackboard†	95.2	94.9	94.6	100.0
% of classrooms with textbooks‡	36.2	23.4	40.3	99.1
% of classrooms with libraries†	18.5	22.4	9.9	39.5
Mean age	40.0	38.1	41.7	45.1
% with Spanish as first language	80.0	74.5	85.5	89.2
% certified	79.3	74.1	82.5	97.6
% in top 2 (of 9) pay scale categories	21.1	14.0	25.5	46.4
% in bottom 2 (of 9) pay scale categories	34.9	51.0	19.3	4.4
Sample size*	608	310	252	46
Student Characteristics				
Mean age	9.2	9.5	9.1	8.7
% with Spanish as first language	75.0	60.2	81.2	93.3
% who work at least occasionally	67.9	84.1	62.9	40.2
% whose mother finished at least grade school	87.3	78.6	90.6	99.5
% whose mother finished at least middle school	47.1	27.5	50.9	91.2
% whose mother finished at least high school	30.6	12.9	31.4	82.5
% whose mother has at least some higher education	14.0	4.8	11.7	53.9
% whose mother has a college degree	3.8	0.8	2.4	20.1
% whose house has a sewerage connection	30.7	9.6	35.8	74.1
% whose house has a phone line	17.5	2.7	15.9	71.6
Sample size*	10,018	3,594	5,285	1,139

Notes: Standard deviations are in parenthesis. *The sample size refers to the total number of observations in each category. Not all variables have valid data for that many observations.
† The teacher indicates the classroom has these inputs.
‡ Computed using an indicator equal to 1 if the teacher responds that more than half of students in the class have their own textbooks.

FIGURE 1.—ENROLLMENT AND CLASS SIZE



Note: The figure uses class-level observations for the third grade. Class size is computed using equation (1).

TABLE 2.—CORRELATIONS AND DIFFERENCES IN MEANS FOR SELECTED SAMPLES

	Correlations between Enrollment and			Difference in Means for Schools above and below the 1st Enrollment Cutoff†	
	Full Sample (1)	Rural Schools (2)	Rural Schools with ≤30 Students (3)	±5 Students Inclusive (4)	±3 Students Inclusive (5)
Teacher Characteristics					
Certified	0.10**	0.03	0.08	−0.09	−0.15
In one of two highest pay scales	0.12***	−0.07	0.04	0.00	0.00
In one of two lowest pay scales	−0.20***	−0.02	−0.09	−0.17	−0.17
Age	0.17***	0.02	−0.04	−0.14	1.93
Spanish is first language	0.12***	0.09	−0.06	−0.17	−0.04
Classroom Characteristics					
Has a blackboard	−0.00	0.04	0.08	0.00	0.00
More than 50% of students have textbooks	0.24***	0.13**	0.18**	0.44**	0.35**
Student Characteristics					
Spanish is first language	0.11***	0.11***	−0.03	−0.18***	−0.08
Mother did not complete primary	−0.17***	−0.15***	−0.00	−0.07	−0.08
Works at least part time	−0.14***	−0.03**	−0.09***	0.12**	0.15***
Lives in a house with sewerage connection	0.26***	0.17***	0.04	0.01	−0.00
Lives in a house with phone connection	0.13***	0.04**	−0.03*	−0.00	0.00

Notes: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively. In columns 1–3 and for indicator variables, simple linear probability or logit specifications produced similar results on significance.
† These columns focus on the first enrollment cutoff because it is the one with the greatest sample sizes. We omit results for the second cutoff, but the qualitative findings are similar.

are significantly less likely to work or have mothers with low levels of schooling.[5]

IV. Empirical Strategies

To address these issues, this paper uses two identification strategies. As background, suppose that the test score of individual i in class j and school k is given by

$$T_{ijk} = \alpha + \beta_C C_{jk} + X_{ijk}\beta_X + \mu_j + \eta_k + \varepsilon_{ijk}, \tag{2}$$

where C_{jk} is class size, X_{ijk} is a vector of controls, and μ_j and η_k allow for within-class and within-school correlations in scores.

A. *Focusing on Small Towns*

A first empirical strategy begins from the observation that the endogeneity of C_{jk} may be less of a concern among rural schools so small as to have one class per grade. Due to their size, these schools cannot seek additional instructors, or allocate students in a compensatory manner. They are more likely to be in small communities where the class size a child experiences is largely determined by her cohort's size, and where parents can exercise little school choice. Among these institutions, therefore, enrollment-SES correlations might be less of a concern.

This is a significant assumption, because even in this sample, larger communities will tend to have larger cohorts, enrollments, and class sizes. One can study its validity as far as observables are concerned by returning to table 2. Column 1 shows that in the full sample there are significant correlations between enrollment and student, school, and teacher characteristics—in every case but one, these are significant at the 5% level. Columns 2 and 3 restrict the sample to rural schools and to rural schools with fewer than 30 students, respectively. All correlations become weaker—by column 3, only two out of twelve are still significant.[6]

Unfortunately, there is no way to verify that these schools are truly isolated, and one therefore cannot rule out the existence of children who live at the margin of two schools' areas of influence, whose parents could send them to the one they perceived to be better.[7] Additionally, though parents could move in reaction to observed class sizes, that would typically be rather costly in this rural setting.

B. *Teacher Allocation Patterns As an Instrumental Variable*

A second design exploits the fact that schools with more than 30 students in a given grade can apply for another instructor. For now, assume one is always requested and granted, so that the number of classes in a school becomes a discontinuous function of its enrollment:

$$n_k^* = \frac{E_k}{\text{int}\left(\dfrac{E_k - 1}{30}\right) + 1}. \tag{3}$$

With this, class size should fall discretely at enrollment intervals of 30 students.

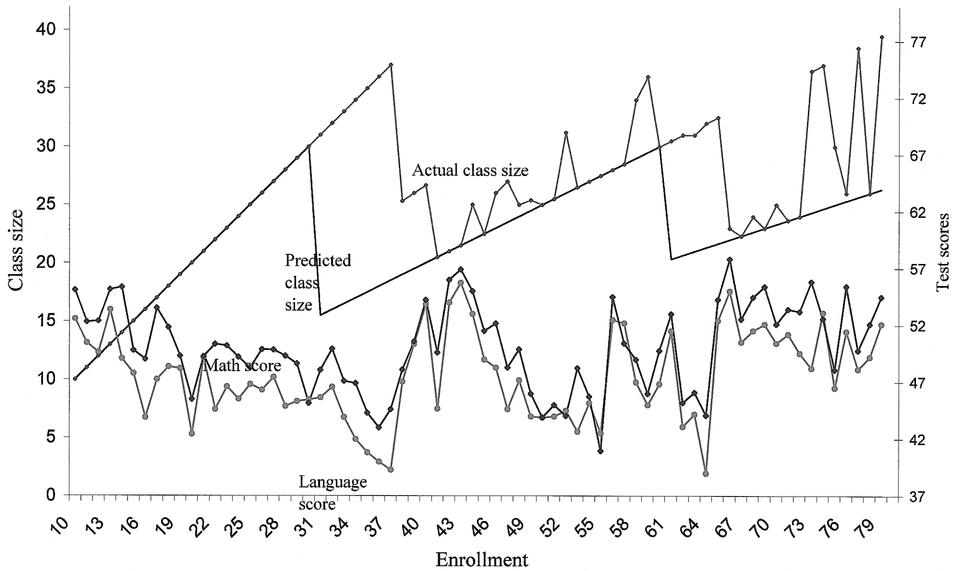
The top segments of figure 2 show that in practice the actual does not always track the predicted class size (for reasons discussed later, figure 2 considers only rural schools). Rather, authorities wait until

[5] Enrollment-SES correlations may also arise within urban areas, for instance, if wealthier or more motivated parents prefer certain schools.

[6] One can also examine the fitted values of regressions of test scores on all the available observables. The relationship between this measure and enrollment is generally smooth and flat among rural schools. As expected, in the full sample there is a clear upward trend.

[7] In general, this would tend to bias the estimated class size effect upward.

FIGURE 2.—ENROLLMENT AND PREDICTED AND ACTUAL CLASS SIZE FOR MATH AND LANGUAGE SCORES



Note: The figure plots averages (by enrollment level) of actual and predicted class sizes, and of math and language scores. It uses class-level observations for the third grade in rural schools only. The data are not smoothed.

enrollments reach approximately 38 to assign new teachers.[8] This may reflect resource constraints, but also illustrates that the teacher allocation mechanism reflects practice rather than a legally binding rule.[9] At approximately 38 students, class sizes drop for most schools, creating a clear discontinuity. The fact that they do not immediately fall for all schools, however, yields a fuzzy discontinuity, making it impossible to rule out all biases.[10] There is another drop at approximately 67 students (one generated by fewer observations), but beyond that the rule ceases to predict class size effectively.

Although equation (3) does not eliminate the relation between class size and enrollment, it does provide useful variation as long as SES is a smooth function of enrollment—an application of the regression discontinuity design introduced by Campbell (1969). Specifically, *if* the schools right above each peak in the class size function are similar to those right below, it is plausible that differences in their outcomes are caused by differences in class size.

To check whether this holds for observables, table 2 presents differences in means for several characteristics for schools within enrollment bands that include the first enrollment cutoff.[11] These do not go in a single direction, and few are statistically significant, particularly within narrower bands.

In contrast, the bottom segments of figure 2 show that average test scores display an up-and-down behavior inverse to the class size function. In the first segment, covering enrollments up to approximately 38, there is a negative relation between test scores and class size. At almost the precise level where the rule first causes class sizes to fall discretely, test scores jump abruptly. Although their behavior between the two discontinuities is more erratic, scores jump again at enrollments (66–68 students) at which there is another discrete reduction in class size.

With this suggestive graphical evidence as background, below we use the predicted class size C_k^* as an IV:

$$C_k = \alpha_1 + \gamma C_k^* + f(E_k) + e_k, \tag{4}$$

$$T_{ijk} = \alpha + \beta_C \hat{C}_k + X_{ijk}\beta_X + g(E_k) + \mu_j + \eta_k + \varepsilon_{ijk}, \tag{5}$$

where $\hat{C}_k$ stands for the predicted values from equation (4), and f and g are flexible functions of enrollment.[12]

Finally, figure 2 considers only rural schools because the teacher allocation rule does not bind in the urban area, where class sizes are substantially higher. Though no reasons for this asymmetry are immediately apparent, note that the existence of sparsely populated rural areas implies that authorities must be willing to tolerate lower class sizes there in order to achieve target enrollment ratios. Also, additional rural teachers are less costly because they are frequently less well trained.

V. Results

Columns 1–4 in table 3 present the baseline OLS specifications for the full sample. The simplest bivariate regressions (1 and 3) suggest larger classes produce higher language and math scores, a result that is not surprising, given the previous discussion. When specifications 2 and 4 add a series of controls, the coefficient on class size becomes insignificant and approaches 0. Something similar happens in the rural sample (columns 5–8). In short, these baseline results are broadly consistent with the idea that class size does not matter.

[8] The IV results below are for a discontinuity at 38 students. The key conclusions are not sensitive to setting this break at 37 or 39.

[9] Nonetheless, several factors suggest a commitment to limiting class sizes. For instance, the agency in charge of rural school building designs its classrooms for 30 students.

[10] See van der Klaauw (2002).

[11] The table considers the first cutoff because it is the one with the most observations. Comparisons around the second cutoff yield similar conclusions.

[12] If f and g have the same form, this reduces to two-stage least squares. See van der Klaauw, Hahn, and Todd (1999).

TABLE 3.—OLS SPECIFICATIONS: CLASS SIZE

	Full Sample				Rural Schools				Rural Schools with Enrollments of 30 or Less			
	Language		Mathematics		Language		Mathematics		Language		Mathematics	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	0.09**	−0.01	0.07**	0.01	−0.10**	−0.09	−0.12**	−0.06	−0.22**	−0.23**	−0.19**	−0.19**
	(0.04)	(0.03)	(0.03)	(0.03)	(0.06)	(0.08)	(0.05)	(0.07)	(0.10)	(0.10)	(0.09)	(0.10)
	[0.09]	[−0.00]	[0.07]	[0.01]	[−0.09]	[−0.08]	[−0.10]	[0.05]	[−0.14]	[−0.14]	[−0.11]	[−0.11]
Student, teacher, and school characteristics†	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.008	0.158	0.005	0.126	0.009	0.037	0.011	0.043	0.018	0.041	0.012	0.032
Number of students	9,993	9,993	9,993	9,983	3,594	3,594	3,594	3,594	1,555	1,555	1,555	1,555
Number of classes	589	589	589	589	298	298	298	298	159	159	159	159

Notes: *, **, *** — indicate significance at the 10%, 5%, and 1% levels, respectively. Huber-White standard errors are in parentheses, and are adjusted for clustering at the class level. Brackets contain the fraction of a standard deviation change in the dependent variable brought about by increasing class size by 1 standard deviation.

† Indicates whether regressions include all the variables in table 1, as well as enrollment, as controls.

TABLE 4.—REDUCED-FORM REGRESSIONS: PREDICTED CLASS SIZE

	Class size		Language score		Math score	
	(1)	(2)	(3)	(4)	(5)	(6)
	0.89***	0.65***	−0.16***	−0.19**	−0.18***	−0.15*
	(0.04)	(0.07)	(0.06)	(0.08)	(0.06)	(0.08)
	[0.69]	[0.50]	[−0.11]	[−0.13]	[−0.12]	[−0.10]
Controls: student, teacher, and school characteristics†	No	Yes	No	Yes	No	Yes
R^2	0.520	0.552	0.014	0.044	0.017	0.047
Number of students	3,594	3,594	3,594	3,594	3,594	3,594
Number of classes	298	298	298	298	298	298

Notes: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively. Huber-White standard errors are in parentheses, and are adjusted for clustering at the class level. Brackets contain the fraction of a standard deviation change in the dependent variable brought about by increasing predicted class size by 1 standard deviation.

† Indicates whether regressions include all the variables in table 1, as well as enrollment, as controls.

Columns 9–12 then implement the first empirical strategy, focusing on schools with enrollments less than or equal to 30, and hence only one third-grade class.[13] The coefficients on class size are uniformly significant in this sample, and they are robust to the introduction of controls. It is relevant to note that if in this set of schools class size is indeed influenced by cohort size, these estimates may represent the cumulative effect of up to three years of small or large classes. This is also relevant, although perhaps less so, for the next research design.

To introduce this strategy, table 4 presents reduced-form regressions using the predicted class size as the key independent variable. In the first-stage regressions (columns 1 and 2), the coefficient on predicted class size is highly significant, and the R^2 always exceeds 0.5.[14] Columns 3–6 present reduced forms for language and math. The coefficients are significant at least at the 10% level, and their magnitude is always greater than that observed in the OLS regressions. They are essentially unaffected by the addition of controls.

Table 5 presents the IV results. In the simplest specifications, the coefficients on class size are negative and significant. With controls, these become somewhat more negative. The addition of a cubic in enrollment renders them less significant, but leaves the point estimates essentially unchanged.

These effects are substantial in magnitude: a 1-standard-deviation reduction in class size (approximately eight students) raises scores by up to 0.3 standard deviations—sufficient to move a student from the 25th percentile to the median of the language score distribution.[15] These effects might also be substantial in terms of the actual skills, like basic literacy, that they entail. Unfortunately, there is no outside evidence on what they might mean for outcomes such as dropping-out probabilities or earnings.

A further note is that combined with figure 2, these estimates suggest that the teacher allocation rule may be operating at a margin at which there are large returns to class size reductions.[16] Note that

[13] The class size rule suggests that to focus on small schools, one should select only institutions with 30 or fewer students. Figure 2, however, could be interpreted as arguing for a higher cutoff. These results use the 30-student cutoff because it is more likely to select schools in isolated communities.

[14] With class-level observations, the R^2 is approximately 0.44—comparable to that in Angrist and Lavy (1999).

[15] Using a 10-student reduction as a benchmark for the STAR experiment, Finn and Achilles (1990) find an effect size, relative to the individual distribution, of 0.13 to 0.25. Krueger (1999) presents similar results. For Israel, Angrist, and Lavy (1999) find effects of 0.10–0.20; for NELS, Akerhielm (1995) and Boozer and Rouse (2001) find effect sizes of 0.15 and 0.29, respectively. Using a 10-student reduction for comparability, the effect sizes found here range between 0.19 and 0.35.

[16] I am indebted to an anonymous referee for this observation.

TABLE 5.—INSTRUMENTAL VARIABLES REGRESSIONS: CLASS SIZE

	Language Score			Mathematics Score		
	(1)	(2)	(3)	(4)	(5)	(6)
	−0.18***	−0.29**	−0.33*	−0.21***	−0.22*	−0.25
	(0.07)	(0.12)	(0.17)	(0.07)	(0.13)	(0.20)
	[−0.16]	[−0.26]	[−0.30]	[−0.18]	[−0.19]	[−0.20]
Controls: student, teacher and school characteristics†	No	Yes	Yes	No	Yes	Yes
Cubic in enrollment	No	No	Yes	No	No	Yes
Number of students	3,594	3,594	3,594	3,594	3,594	3,594
Number of classes	298	298	298	298	298	298

Notes: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively. Huber-White standard errors are in parentheses, and are adjusted for clustering at the class level. Brackets contain the fraction of a standard deviation change in the dependent variable brought about by increasing class size by 1 standard deviation.
† Indicates whether regressions include all the variables in table 1 as controls.

whereas there is a decline in scores between enrollments of 10 and 31, the gradient becomes significantly steeper after 32 (and again before the second threshold).[17] This sudden worsening in performance—particularly among schools that do not get a new teacher until their enrollment approaches or even surpasses 40—combined with the positive impact of a significant reduction for those that do, may account for the magnitude of the results.[18]

Put differently, the IV procedure yields a local average treatment effect, or the average effect for the schools that are granted an additional teacher when their enrollment reaches the threshold; this effect could be larger than that observed for randomly selected schools.[19] Note also that because not all schools receive a new teacher at the same enrollment level, one cannot rule out that schools likely to get the biggest return from an additional teacher are the ones actually granted one. Again, in this case the IV results would overestimate the returns to reducing class size at a random school (even one with an enrollment close to 38).

VI. Conclusion

The two research designs introduced in this paper suggest that larger class sizes have a negative effect on achievement. Further, they are consistent with a nonlinear effect: they are larger for the IV strategy, which emphasizes substantial class size reductions from high initial levels; and smaller for the first design, which emphasizes schools with smaller classes. However, it is also possible that part of the IV results originate in tracking that becomes feasible when schools get an extra teacher. If so, the returns to further increases in the number of instructors might be smaller, giving rise to a different type of nonlinearity. With the data used here, it is impossible to isolate to what extent tracking is responsible for the estimated effects, in part because few rural schools receive a third teacher.[20]

Finally, in part because of such potential nonlinearities, these results must be interpreted with care. They cannot settle, for instance, whether across-the-board class size reductions would be worthwhile in Bolivia, let alone elsewhere. They also illustrate how difficult it is to rule out or confirm class size effects in a policy-relevant manner.

[17] This dip could also reflect nonrandom sorting—for example, more motivated principals getting additional instructors earlier. This would be problematic, for even the IV results would be biased by such selection.

[18] In other words, the above results may reflect a combination of the impact of of going from a class size of approximately 35 to 40 or more students (among institutions not immediately assigned new instructors) with that of going from approximately 38 to approximately 20 students (among those who do immediately receive them). Unfortunately, the data are not rich enough to parse out such nonlinearities. For instance, in narrow ranges around the cutoff, the IV specifications result in generally more negative point estimates, but also significantly less precision.

[19] See Angrist (2003).

[20] Nonetheless, there is not much prima facie evidence of tracking—among schools with 45 or fewer students and two third-grade classes, there are few significant differences in scores or SES across classes.

REFERENCES

Akerhielm, K. "Does Class Size Matter?" *Economics of Education Review* XIV (1995), 299–341.
Angrist, J., "Treatment Effect Heterogeneity in Theory and Practice," NBER working paper no. 9708 (2003).
Angrist, J., and V. Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114:2 (1999), 533–575.
Boozer, M., and C. Rouse, "Intraschool Variation in Class Size: Patterns and Implications," *Journal of Urban Economics* 50:1 (2001), 163–189.
Campbell, D., "Reforms as Experiments," *American Psychologist* 24 (1969), 409–429.
Coleman, James S., Ernest Q. Campbell, Carol J, Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York, *Equality of Educational Opportunity* (Washington DC: Government Printing Office, 1966).
Finn, J., and C. Achilles, "Answers and Questions about Class Size: A Statewide Experiment," *American Educational Research Journal* (Fall 1990), 557–577.
Hanushek, E., "The Evidence on Class Size" (pp. 131–168), in Susan E. Mayer and Paul Peterson (Eds.), *Earning and Learning: How Schools Matter* (Washington, DC: Brookings Institution, 1999).
Krueger, A., "Experimental Estimates of Educational Production Functions," *Quarterly Journal of Economics* 114:2 (1999), 497–531.
——— "Economic Considerations and Class Size," Princeton University Industrial Relations Section working paper no. 447 (2000).
Lavy, V., "Endogenous School Resources and Cognitive Achievement in Primary Schools in Israel," Hebrew University, Falk Institute discussion paper no. 95.03 (1995).
Lazear, E., "Educational Production," *Quarterly Journal of Economics* 116:3 (2001), 777–803.
Mizala, A., P. Romaguera, and T. Reinaga, "Factores que inciden en el rendimiento escolar en Bolivia," Universidad de Chile mimeograph (1999).

Pritchett, L., and D. Filmer, "What Educational Production Functions Really Show: A Positive Theory of Education Spending," World Bank policy working paper no. 1795 (1997).

Urquiola, M., "Universal Education," in *Defeating Poverty: Eight Goals at a Time* (La Paz: EDOBOL, 2000).

van der Klaauw, W. "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach,' *International Economic Review* 43:4 (2002), 1249–1287.

van der Klaauw, W., J. Hahn, and P. Todd, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," UNC–Chapel Hill mimeograph (1999).

FORECLOSING ON OPPORTUNITY: STATE LAWS AND MORTGAGE CREDIT

Karen M. Pence*

Abstract—Foreclosure laws govern the rights of borrowers and lenders when borrowers default on mortgages. In states with laws favoring the borrower, the supply of mortgage credit may decrease because lenders face higher costs. To examine the laws' effects, I compare approved mortgage applications in census tracts that border each other but are located in different states. Using a regression-discontinuity design and semiparametric estimation methods, I find that loan sizes are 3% to 7% smaller in defaulter-friendly states; this result suggests that defaulter-friendly laws impose material costs on borrowers at the time of loan origination.

I. Introduction

State foreclosure laws balance the rights of lenders and borrowers when borrowers default on mortgages. Some states attempt to protect borrowers who have fallen upon hard times by providing extensive protections to defaulters. These protections, however, can impose substantial costs upon lenders. If lenders attempt to recoup these costs by increasing interest rates or requiring larger down payments, laws intended to protect homeowners in distress may impose costs on all borrowers.

In this paper, I examine the effect of foreclosure laws on the size of approved mortgage loans. This effect is, a priori, ambiguous because foreclosure laws may have offsetting effects on mortgage supply and demand. As mentioned above, the supply of mortgage credit may be lower in defaulter-friendly states because lenders experience higher costs. However, defaulter-friendly laws provide borrowers with wealth insurance against falling house prices. If borrowers value this insurance, mortgage demand may be higher.

Identifying the effect of foreclosure laws on the mortgage market is difficult because both the laws and the factors that affect the mortgage market vary by region. The northeastern and midwestern states, for example, are more likely to require the judicial foreclosure process examined in this paper. Factors that affect housing and mortgage supply and demand, such as weather, fertile soil, and proximity to amenities, also vary across the country, and real estate downturns generally affect some regions more than others. In a simple cross-section regression, then, a regional shock to a given housing market could be misinterpreted as an effect of foreclosure law. Although previous authors have established a connection between foreclosure laws and the supply of mortgage credit, no study has taken regional patterns into account.[1]

I address this identification problem by comparing approved mortgage applications in census tracts that border each other but are located in different states. Mortgages in these census tracts are subject to different laws, but because of their proximity may take on similar values for important unobserved variables. Analogous *borders* identification strategies were used by Holmes (1998) to estimate the effect of state right-to-work laws on the location of manufacturing, and by Black (1999) to explore the effect of school quality on house prices.

I implement the borders strategy with a fixed-effects model and a partial linear model. The fixed-effects model assumes that unobserved variables are constant across a given urban area. The partial linear model lets unobserved variables take on a different value for each census tract and requires only that these variables change smoothly over space. This semiparametric estimator fits within the regression discontinuity framework of Hahn, Todd, and Van der Klaauw (2001) and Porter (2002) and is a useful tool for estimating the effects of state policies while controlling flexibly for spatially varying unobserved factors. Taking account of these unobserved variables proves to be important for my results.

II. Judicial Foreclosure Requirement

In 21 states, lenders must go through the courts to foreclose on a property. This judicial foreclosure process takes 5 months longer on average than the nonjudicial alternative (Wood, 1997) and imposes more transaction costs. Several studies have verified that the judicial requirement can significantly raise lender foreclosure costs,[2] perhaps by as much as 10% of the loan balance.[3] From the borrower's perspective, this requirement provides several months of free rent and protection against lender excesses.

Received for publication January 27, 2004. Revision accepted for publication April 26, 2005.

*Board of Governors of the Federal Reserve System.

I am grateful to John Karl Scholz, Yuichi Kitamura, and Bob Haveman for their guidance and to David Brown, Mary DiCarlantonio, Chris Greene, Greg Grothe, Michele Rambo, Keiko Snell, and James Tedrick for outstanding assistance. For helpful comments, I thank the editor, two anonymous referees, the faculty and students of the University of Wisconsin, my Federal Reserve Board colleagues, Chuck Capone, and numerous seminar participants. I received generous financial support from a U.S. Department of Housing and Urban Development doctoral dissertation grant, the Christensen Award in Empirical Economics, and the Social Science Research Council Program in Applied Economics, funded by the MacArthur Foundation. This research does not necessarily reflect the views of the Board of Governors of the Federal Reserve System, the U.S. Department of Housing and Urban Development, the Social Science Research Council, or the MacArthur Foundation.

This paper is dedicated to the memory of Bernie Saffran.

[1] See Jones (1993), Alston (1984), Wood (1997), Schill (1991), and Meador (1982).

[2] See Touche, Ross and Co. (1980), Clauretie (1989), Clauretie and Herzog (1990), Ciochetti (1997), Wood (1997), Pennington-Cross (2003), DeFranco (2002), and Bridewell (1938).

[3] Foreclosure costs come from three sources: time-dependent expenses such as maintenance, taxes, and forgone interest; transaction costs; and the loss on the sale of the property. The 5-month delay in judicial states implies that time-dependent expenses are higher by 5% of the loan balance (Wilson, 1995). I assume transaction costs are higher by 1% of the loan balance. At the foreclosure sale, houses in judicial states sell for 4% less (Pennington-Cross, 2003), presumably because of greater depreciation during the longer foreclosure period.