Methodological Appendix for Housing Costs, Zoning, and Access to High-Scoring Schools

By Jonathan Rothwell

This document provides methodological detail for the Brookings Metropolitan Policy Program publication "Housing Costs, Zoning, and Access to High-Scoring Schools." See the Brookings website for the full report.

School Quality

The measure of school quality used in this study is the percentage of students scoring at or above proficiency, according to state-specific tests. The test score data were purchased from GreatSchools, which collects test score data for public schools from various state sources.[1] The final database had 84,077 schools, out of a total of 99,806 listed by the National Center for Education Statistics.[2] 273 schools were missing demographic data for all racial groups, but thousands were missing data on native Hawaiian enrollment. For that reason, schools with missing Hawaiian enrollment were assigned a value of zero for that group, but no other missing values were imputed.

School proficiency does not isolate the advantage or disadvantage of being at a specific school for a child, but does provide an important general indicator of how well students perform at that school.[3] The analysis assumes that test scores are determined through a combination of factors that operate at the school level, the family and neighborhood level, and the individual level. With that said, school test scores are the most salient measure available on how well a given student is likely to perform who is enrolled in that school.

This study analyzes test scores for public schools only, as the data were not available for private schools. This clearly leaves out a large number of schools, but the percentage of students that attend private schools is very small. In the 2009-2010 school year, Brookings analysis of data from the NCES shows that 91 percent of students in the United States (enrolled from Kindergarten through high school) attend public schools. The share attending public school is slightly less for whites, at 90 percent, and very high for blacks and Hispanics, at 96 percent each. Whatever the advantages might be of private education, it is rare for most children, especially those from disadvantaged backgrounds. In fact, the number and share of students enrolled in private schools has fallen since the 1997-1998 school year, when it was 11 percent.

For each school, multiple test scores are reported across grades, subjects, and for various years. For this analysis, a single score was calculated for each school. The first step was to keep only observations from the latest year of data. For two-thirds of schools, this was 2011, and for 90 percent, it was 2011 or 2010.[4] The next step adjusted for specific test characteristics by calculating the average difference between the state proficiency rate and the school proficiency rate for the given grade and subject. As the formula below shows, this yields an average state-adjusted-test score for each school. Here P is the percentage of students scoring at or above proficiency on the state exam, and $\bar{P}$ is the state average proficiency rate. The subscript *i* denotes the school, *s* the state, *t* the subject, *g* the grade level, and *y* the year the exam was administered.

$$A1. Average\ State\ Adjusted\ School\ Test\ Score_s = \frac{1}{n}\sum_{S_n}^{n}(P_{i,s,t,g,y} - \bar{P}_{s,t,g,y})$$

This state adjusted score includes a grade effect, a subject effect, year effect, and a state effect, which should mitigate much of the bias from comparing schools. One caveat is that many states were missing grade level observations, meaning that their schools were compared only to the state and subjects averages for that year.

The analysis also calculates national and metropolitan level indicators of the level and distribution of school performance by aggregating school level data. School specific data on enrollment by race and eligibility to receive free or reduced price lunches (an indicator of poverty or low income), student-teacher ratios, and school district spending were downloaded from the NCES's 2009-2010 Common Core of Data (the latest available) and matched to the test scores using the identification numbers provided by GreatSchools and the NCES. At the national and metropolitan levels, sample weights were constructed based on the share of students at each school, and those weights were applied to the data to calculate attendance by quality for the average student and by group (using group specific weights in that case).

The school test-score gap is defined as the difference in percentile ranking (on a scale of 1-100) for the average school attended by two different groups of students. For the metro test score gap, each state adjusted score is ranked against all other schools in the metro. In online data, schools are also ranked against all other schools in the nation to facilitate cross-metro comparisons.

For national and metropolitan summary data, the school test score gap is reported using data on all public schools. For parts of the analysis that compare the test score gap to housing costs, the test score gap is calculated only for the 48,008 schools in which a majority of students are enrolled in elementary grades (i.e., kindergarten to fifth grade), since that is the universe for the housing calculation—as described below.

Supplementing NCES Data

The NCES, which is part of the U.S. Department of Education, receives data from state government agency officials. Yet, the NCES data does not always match publicly available data from state education departments. When analyzing the NCES data, Brookings found a number of errors that ranged from important to trivial. They were of two kinds: missing lunch program data, when that data is reported for the school by state agencies; and schools reported as having zero lunch program eligible students, when state data says otherwise.

To the first problem, 32 percent of public schools in New York state included in this sample had missing lunch status data—meaning that schools did not report how many children were eligible for free or reduced price lunches, even though they reported total enrollment. No other state was missing lunch program observations for more than 5 percent of students. For five large metropolitan areas—New York City, Richmond, and Chicago, more than five percent of students could not be identified by lunch status, with New York City being by far the highest, with 38 percent).

To the second problem, Alaska, Illinois, Ohio, and Texas reported having a very high percentage of schools in which zero students were reported as being eligible for free or reduced lunch (above 5 percent) relative to other states (in 35 states it is less than 0.5 percent).

Since Alaska does not have any of the 100 largest metropolitan areas, data quality issues were not investigated for those schools, but for New York, Virginia, Illinois, Ohio, and Texas, state resources were consulted to correct potential errors. Schools in the state downloads were matched to the larger national database using phone numbers, principal names, and school names combined with either counties or other unique indentifying information.

School level information for New York on national lunch program status was downloaded from the New York State Education Department.[5] The author was able to replace 998 missing school-level observations regarding students eligible for free or reduced lunch. This identified roughly 579,000 extra low-income students, which were added to the database. Virginia was also missing a small number of observations, especially in the Richmond metro area, but only one replacement could be found with non-missing data. In that case, lunch eligibility enrollment data was obtained from the Virginia Department of Education.[6]

NCES data on eligibility for free or reduced price lunch status is reported with a large error for the state of Texas. The reason seems to be that Texas schools use an alternative indicator they call “economically disadvantaged” and so many schools do not report free or reduced price lunch status separately (or rather report it as zero). For example, 369 schools report having no free or reduced lunch eligible students (with roughly half from Hidalgo County, the McAllen metropolitan area, one of the poorest in the country). By comparison, only four schools in California report zero reduced price lunch students. In Texas, all students eligible for free or reduced price lunches are deemed economically disadvantaged; also included are students eligible for other poverty programs.[7] Student data on this low-income indicator is available for each school from the Texas Education Agency’s (TEA) Academic Excellence Indicator System. These TEA data were used to replace low-income enrollment data for 2,491 Texas schools. Free and reduced lunch specific enrollment was replaced as missing for schools reporting zero students in the NCES, since it was not available from the TEA. Finally, total enrollment from the TEA replaced total enrollment from the NCES for 12 schools.

Similar changes were made to school data in Ohio and Illinois. For Ohio, 3039 schools were reclassified with the appropriate state value from the Ohio Department of Education (30 of which had reported no eligible students).[8] For Illinois, NCES data for 397 schools on lunch eligibility was replaced with from the state of Illinois (replacing 247 schools reported as having no eligible students).[9]

Location and Housing Costs

In the absence of data on actual attendance zones boundaries for every school in the United States, hypothetical attendance zones are created using census tracts. The first step uses GIS to assign every census tract to all schools within a ten mile radius. Geographic coordinates and enrollment by grade level for schools are obtained from the NCES. Only schools that have at least 50 percent of their students in enrolled in elementary school (kindergarten through fifth grade) are included. This is done to avoid cases where students attend schools further away from residence, as is likely to be the case for schools that are strictly for middle or secondary students.

The next step is to calculate the distance between the school and (centroids of) the census tracts located within the ten mile radius. This is done based on longitude and latitude of both schools and tract using the *Vincenty* program written for STATA. For each census tract, school enrollment data was used to assign a weight for the tract between 0 and 100 percent that would be equal to the tract's contribution to total enrollment in the school. Starting with the nearest census tract and radiating outwards, housing costs for a tract were "assigned" to a school until cumulative tract enrollment equaled school enrollment. When tract enrollment shares pushed cumulative enrollment shares from below one to above one, only the fraction of students that would bring the total to 100 percent were included, and the tract weight was adjusted accordingly. The last step calculates a weighted average of housing prices using tract enrollment shares as the weight.

This method has the advantage of creating a natural attendance boundary around each school based on the student population surrounding it. It is limited to some extent by census tract boundaries, which are needed in the absence of actual housing cost data for each family unit. The analysis allows schools nearby one another to have overlapping attendance boundaries and, therefore, identical or nearly identical housing costs, depending on distance to surrounding tracts. This gives it a natural relationship to the school district's housing stock, even as actual attendance boundaries can be gerrymandered from year to year for administrative or political purposes.

Enrollment Segregation

Enrollment segregation captures how evenly students from disadvantaged backgrounds are enrolled across zip codes. The purpose of this index is to examine the correlation between school test-score gap and neighborhood sorting into schools. The calculation uses the dissimilarity index, which is described on the Census Bureau website.[10] In this case, the index refers to the percentage of low-income students that would have to switch the zip code of their school with middle/high-income students to attain equal percentages of enrollment across all zip codes in a metropolitan area. The index is also calculated for black-white segregation and Hispanic-white segregation. To the extent that ZIP codes do not determine school enrollment, this index measures school segregation based on parental or administrative sorting; to the extent that ZIP codes dictate enrollment, it measures residential segregation. Thus it makes no assumptions about how students are assigned administratively.

Data on Zoning

The only recent survey known to the author at this time that takes a representative survey of jurisdictions within metropolitan areas was conducted by Rolf Pendall in 2003. The results were analyzed and reported in a Brookings publication with Robert Puentes in 2006.[11] Using this database, recent academic publications have found that zoning restrictions on density cause higher levels of racial segregation and higher housing prices at the metropolitan level.[12] The database is described in detail in those publications.

The analysis above uses the Pendall database on zoning to measure density restrictions at the metropolitan scale. The main survey question used here is: “What is the maximum number of units allowed in the jurisdiction per acre of land?” This was available for 1,677 local governments in 50 of the largest metropolitan areas in the United States, as of 2000. The local measures were aggregated to metropolitan areas, since the survey sample was designed to be representative of local governments in those areas. Although the survey was taken in 2003, academic work finds that zoning restrictions tend to stay relatively constant over 10 year periods and respond historically to changes in metropolitan population density.[13]

To get a larger sample of zoning for states and metropolitan areas, an alternative index of zoning was calculated based on the share of law firms in the state that specialize in zoning laws. This was done, as described above, using lawyers.com. For each state, the number of law firms with zoning specialists was divided by the total number of law firms. Each school level observation was assigned a news index zoning score based on its state. Metropolitan measures took the average score for each of its schools, such that metros that cross state lines were given a blended index. This index was highly correlated with the Pendall measure for the 49 metropolitan areas in which they overlapped (correlation coefficient of 0.57).

This index shows that Connecticut, Rhode Island, New Hampshire, and New Jersey are the most restrictive states, each being roughly two standard deviations above the mean. For New Jersey, especially, this seems reasonably accurate given the contentious state Supreme Court history on zoning and a historical analysis from academics.[14] Likewise, Massachusetts and Pennsylvania score highly (more than one standard deviation above the mean), which conforms to research on zoning in those states.[15] Yet, high values for

Idaho, Utah, and Hawaii may indicate that the index is picking up non-exclusionary zoning (see below).

The zoning law firm index is highly correlated with year of statehood and the number of rural housing units as a share of land area (rural housing density). This is consistent with the idea that urbanization during the late 19th and early 20th century led to the settlement of rural jurisdictions, as affluent residents fled cities to gain more autonomy over local government affairs.[16]

Finally, the zoning law firm index is interpreted here to proxy exclusionary zoning. There are two pieces of empirical evidence to support that interpretation. One, it is highly correlated with the Pendall survey measure of anti-density zoning, as noted above. Second, it is highly correlated with a newspaper index of exclusionary zoning created for each state and assigned to metros using Proquest. From 1975 to January of 2012, Proquest news archives were searched separately for each state using the key words “zoning” and “exclusionary zoning.” The results were deflated by instances of the word January and July (averaged together). For the 100 largest metropolitan areas, the zoning law firm index is more highly correlated with the exclusionary zoning index than the generic zoning index. Moreover, the generic index is only weakly correlated with the Pendall measure (.17), while the “exclusionary” news index is highly correlated with it (.40).

Aggregated measures lose some of the local detail and introduce error. To provide that detail, two other surveys were used. One was a nationally representative survey of local governments conducted by scholars at the Wharton School at the University of Pennsylvania. It is described in the methods section above. To link this database to the variables used in the report, schools—with accompanying data for housing costs, enrollment, test scores, and metropolitan areas—were assigned to the places (e.g. town, cities, etc) with zoning data in the Wharton database, using place names and states.

Finally, regulatory data was downloaded from a 2004 study of housing regulations in Massachusetts by the Pioneer Institute and Rappaport Institute.[17] For researchers wishing to replicate the analysis in Massachusetts, here are details of how the index was constructed. To measure zoning, the percentile rank of four variables related to restrictions on density were averaged to get one index. These variables included a measure of the minimum lot size required for any housing developed under the most flexible cluster rules. The average was 8,370 square feet. A dummy variable was created if the town either does not permit multi-family housing or allows it only by special permit, or special permit within a more flexible cluster zone. 62 percent of towns met these criteria. Another variable was the longest frontage requirements in the town for single family housing—a measure of orientation towards expensive homes. 161 feet was the average, with 300 the maximum. Finally, combining two variables, the author calculated the percentage of zoning districts in the town that require at least 150 feet in front of the housing unit. A high number on this measure indicates zoning that blocks affordable housing. The average was 43 percent. For details, see the codebook provided by the Pioneer Institute website.[18]

To assign schools—along with enrollment data, test scores, and housing costs—to the towns listed in the Massachusetts database, the first step to match town names to a crosswalk between towns and zip codes, provided by Moody's Economy.com. Then schools were assigned to towns based on their zip codes. This was done instead of directly linking town names to schools because of the idiosyncratic spellings of Massachusetts municipalities, townships, and villages, which differ across databases.

Regression Results

In the section on metropolitan areas measures of zoning and housing prices, the tables and figures imply that more restrictive exclusionary zoning is associated with high housing cost to live near good schools compared to bad schools (measured by test score performance). That analysis is described below.

In equation (2), the dependent variable, R, is the ratio of housing costs in the average neighborhood of top-quintile schools to housing costs in the average neighborhood of bottom-quintile schools. Z stands for zoning, and M is a vector of these metropolitan level variables: median household income, the Bachelor's degree or higher educational attainment rate, household income inequality (measured with the Gini coefficient), the share of the population that is black or Hispanic, the log of the population density and a dummy variable if the MSA is one of the 100 largest (i.e. it has a population above 500,000, roughly). ϵ is an error term that is correlated within states areas, since state constitutions and histories affect zoning differently. The regression to be estimated is

$$(2)\ R_m = \beta_1(Z_m) + \beta_2(M_m) + \epsilon_s$$

This looks at the correlation between zoning and housing premium, holding other factors constant. The data is summarized in Appendix Table 1 and the results are shown in Appendix Table 2, both below.

Appendix Table 1. Summary Statistics of Main Variables

Variable	Obs	Mean	Std. Dev.	Min	Max	Source of data
Housing costs in tract of top-quintile school/bottom quintile	347	1.79	0.54	0.59	3.54	U.S. Census, ACS 2005–2009
Standardized Index of Anti-density zoning	49	0.00	1.00	-1.89	1.79	Pendall survey, 2003
Standardized Index of zoning law firms per law firm	347	-0.09	0.94	-1.63	3.26	Lawyers.com
MSA Median household income (in thous)	347	47.02	8.12	31.74	84.52	2010 Census
MSA household income inequality (Gini coefficient)	347	0.45	0.02	0.39	0.54	Census, two-year 2008–2010 ACS
MSA black and Hispanic share of population	347	0.22	0.16	0.01	0.94	Census, 3-year 2007–2009 ACS
MSA population density (per sq miles of land)	347	298	332	7	2826	2010 Census
MSA Bachelor's Degree attainment rate, 2010	347	0.26	0.08	0.11	0.58	2010 Census
Binary variably for one of the 100 largest MSAs	347	0.29	0.45	0.00	1.00	2010 Census
Elementary school test score gap*	347	20.0	7.5	-4.6	37.8	GreatSchools and NCES
Year of Statehood (primary city)	347	1829	36	1787	1959	U.S. Mint

*Note: This variable compares schools attended by low-income students to middle/high-income students. “Low-income” here refers to students eligible for either free or reduced price lunch; middle/high-income refers to students who are not eligible for either program.

Appendix Table 2. Regression of metropolitan zoning on housing cost gap

	Housing cost gap	
	1	2
Pendall measure of anti-density zoning	0.170**	
	(0.0661)	
Zoning law firm index		0.0811**
		(0.0360)
Median household income in thousands, 2010	0.00523	0.00950
	(0.0203)	(0.00667)
MSA household income Gini coefficient, 2008–2010	4.596	5.500***
	(5.784)	(1.280)
MSA black and Hispanic share of population, 2007–2009	0.568	0.579**
	(0.538)	(0.234)
MSA population density in 2010	0.000166	0.000180*
	(0.000108)	(0.000105)
Bachelor's degree or higher attainment rate, 2010	-2.107	-0.459
	(2.983)	(0.466)
100 largest metropolitan area		0.378***
		(0.0803)
Constant	0.377	-1.282*
	(2.650)	(0.678)
Observations	49	347
Adjusted R-squared	0.264	0.347

Robust standard errors in parentheses, clustered on states. *** p<0.01, ** p<0.05, * p<0.1

To calculate the effect on the housing cost gap of changing zoning from one extreme to another, the range of zoning is multiplied by the regression coefficient from Appendix Table 2. The range equals 3.68 using the Pendall data and 4.92 using the zoning law firm index. The product yields .63 and .40, respectively.

Given the theory outlined above, one needs to consider not only how zoning affects housing costs, but whether or not it affects inequality in schooling. The argument is that exclusionary zoning siphons poor children into poor neighborhoods with low-performing schools, and affluent children into affluent neighborhoods with high-performing schools. One may test this directly is to use zoning as an instrument for the housing premium in order to test the causal effect of the housing premium on school inequality.

Zoning is a valid instrument if it meets the two conditions that it strongly predicts the housing premium and that there is no way for it to affect school inequality, except through its effect on the housing premium, which is itself a measure of economic segregation. There is no way to definitely prove the second criteria, but it is difficult to imagine how the disproportionate presence of zoning lawyers could cause differences in access to schooling if not by indicating more exclusionary zoning. Fortunately, one can add a second instrumental variable and use the Hansen-J statistic to test if the equation is over-identified (meaning the instruments are invalid). This second instrument indicates the degree of rural settlements and is measured by dividing the number of rural housing units in the metropolitan area by total metropolitan land area. One can think of it as an alternative measure of zoning, and the two are highly correlated. If rural settlements are disproportionately common in a large metropolitan area, then it is likely because zoning has shaped the landscape accordingly by isolating rural suburbs from urban development.

Assuming the assumptions are valid, the equation would look like this:

$$(3)\ S_m = \beta_1(\bar{R}_m) + \beta_2(M_m) + \epsilon_m$$

Here, S stands for the difference in school performance for the average low-income and middle/high-income student. R hat is the housing premium, instrumented with zoning as in equation (2) and M is the same vector of control variables from equation (2).

These results are reported in Appendix Table 3. As predicted, the housing premium is strongly associated with greater school inequality, and if zoning is a valid instrument, the results can be interpreted causally.

Appendix Table 3. 2SLS Regression of housing cost gap from zoning on school test-score gap

	Test-score gap, elementary	Test-score gap, all schools
	1	2
Instrumented: Housing costs top quintile schools/bottom quintile	9.143**	18.79***
	(4.152)	(5.843)
Median household income in thousands	0.196*	0.0223
	(0.101)	(0.123)
MSA household income Gini coefficient	20.28	-43.25
	(32.18)	(39.77)
MSA black and Hispanic share of population	2.793	-3.079
	(3.179)	(3.743)
MSA population in millions, 2010	-8.71e-05	-0.00222
	(0.00145)	(0.00205)
Bachelor's degree or higher attainment rate, 2010	3.424	11.81
	(7.945)	(8.821)
100 largest metropolitan area	1.358	-1.201
	(1.630)	(2.202)
Constant	-16.49	-1.321
	(12.13)	(14.50)
Observations	347	347
Anderson-cannon coefficient	10.4	10.4
Hansen-J	0.28	0.58
R-squared	0.55	0.10

Robust standard errors in parentheses, clustered on states. *** p<0.01, ** p<0.05, * p<0.1. Instrument is the zoning law firm index and the number of rural housing units per acre of total MSA land in 2010.

In order to predict how changing zoning could change the school test-score gap, the same thought experiment from above can be conducted above. From the previous set of regressions, the effect of moving from an extremely restrictive state's zoning laws to a state with minimal zoning laws yields a reduction in the housing cost gap of approximately 0.40 points. This effect size can be multiplied by 9.1 (the coefficient on zoning in column 1 of Appendix Table 4) and 18.8 (the coefficient on zoning in column 2) to obtain an estimated effect that ranges from -4 to 7 on the test score gap, depending on whether one looks at only elementary schools or all schools.

In results not shown, the year of statehood and the number of rural housing units in the metropolitan area per acre of land were used as instruments for the zoning law firm index in a regression of zoning on the housing cost gap, controlling for the same variables used in Appendix Table 1. The relationship is statistically significant and the coefficient on zoning increased to 0.43—implying that a change in zoning could cause a reduction in the housing cost gap of 1.6 percentage points, which is almost enough to eliminate the test score gap. Moreover, the instruments pass the necessary validity tests: they strongly predict zoning (Anderson coefficient has a p-value of 0.0) and the Hansen J-statistic over-identification test suggests that they are exogenous (p-value of .20). This implies that the results shown in Appendix Table 2 are biased towards finding no relationship between zoning and the housing gap. This could be the case if state or local government policies counteract some of zoning's exclusionary effect with inclusionary policies, perhaps as a result of political pressure or state court rulings.

The correlation between state of birth and zoning is -0.44 for 49 states with metropolitan areas in the sample. Older states had more time to settle and urbanize, and hence suburbanize and set up exclusionary zoning in the 1920s. Previous research by the author finds that 1920s population density is a very strong predictor of zoning today.[19] This does not prove that zoning causes large housing premiums near good schools or that zoning causes school segregation, but it is difficult to think of other plausible explanations for these associations.

Metropolitan Test Scores and Individual Labor Market Outcomes

In the main report, at the end of the first finding in the report, the analysis addresses the following question: Does the performance of public schools attended by minority groups affect their probability of employment, college education, or wages?

The literature cited in the background section suggests that the answer is yes, but it is possible that schools make such a small difference that it does not show up in aggregated data. To look for evidence either way, the analysis compares the outcomes mentioned above for young individual blacks and Latinos living in different metropolitan areas. The idea is that recent graduates from the school system are the most likely to have been affected by it, and that young people are less likely to have switched metropolitan areas since high school than middle-aged adults.

The individual data come from the 2010 American Community Survey as provided by the Integrated Public Use Microdata Series (IPUMS).[20] To make individuals comparable, the analysis adjusts outcomes for observable characteristics that might otherwise bias the results: family income, age, sex, age of immigration to the United States (if foreign-born), whether or not he or she moved states since birth, marital status, disability status, and number of children. Since metros differ by more than just school performance, metropolitan area variables were included to control for school performance for the average white student, the share of the population aged 25 and older with a Bachelor's degree or higher, the unemployment rate in 2010, the population size, the years of education required by the average job, and the share of residents from the ethnic group in question. Those data were calculated from the U.S. Census Bureau by Brookings. Family income was calculated by subtracting individual income from total family income. For both Latinos and blacks, the sample was limited to data on roughly 22,000 to 32,000 individuals who have lived in the United States since the age of 13 or younger and are currently between the ages of 18 and 25. All results apply sample weights from the Census Bureau.

Appendix Table 4 shows the results from six regressions of metropolitan level test scores on individual outcomes for black or Latino young adults aged 18 to 25. Specifically, the regression looks at how average state-adjusted test scores (ranked against all schools) for the average school attended by blacks (or Latinos) are correlated with economic outcomes for young-adult blacks (or Latinos). To isolate the effects of test scores from the effects of living in a highly-educated metro, the results also adjust for school test scores for the average white student. These results were used to predict the summary statistics displayed in Table 4 in the main body of the report by calculating the predicted effect of living in the metro with the lowest test scores to the effect of living in the metro with the highest test scores.

The results are very strong for blacks. On all indicators—income, post-secondary attendance, and employment—blacks do better if they live in metros with better test scores. Since the analysis adjusts for family income, it is difficult to dismiss these findings as stemming from any obvious selection bias (i.e. more economically successful black families living in metros with better test scores for blacks). The results for Latinos are very similar, except the probability of post-secondary school attendance is not significantly correlated with metro test scores.

Still, caveats are needed. As stated in the endnotes of the manuscript, this analysis cannot definitely conclude that schools are the cause of the better economic outcomes, since individuals who live in higher-scoring metropolitan areas may have unmeasured advantages.

Appendix Table 4. Regression of individual labor market outcomes on MSA measures of school performance

	Earnings		Attained some college or higher education		Employed or in school	
	Blacks	Latino	Blacks	Latino	Blacks	Latino
	1	2	3	4	5	6
Percentile rank of average black student's elementary school	60.98**		0.00175**		0.00172**	
	(23.47)		(0.000774)		(0.000707)	
MSA Black Share of population, 2007–2009	1,441		-0.129*		-0.0607	
	(1,780)		(0.0749)		(0.0650)	
Percentile rank of average Hispanic student's elementary school		71.31**		0.00117		0.00239***
		(31.55)		(0.00169)		(0.000543)
MSA Hispanic share of population, 2007–2009		-1,382*		0.148***		0.0556***
		(813.5)		(0.0470)		(0.0175)
Percentile rank of average white student's elementary school	-88.26**	-67.24*	-0.00260*	-0.00259	-0.00341***	-0.00215***
	(33.72)	(38.95)	(0.00140)	(0.00213)	(0.000917)	(0.000814)
Family income	-0.0120***	-0.00837***	1.28e-06***	8.69e-07***	2.25e-07***	1.96e-07***
	(0.00271)	(0.00240)	(1.01e-07)	(7.57e-08)	(8.30e-08)	(4.00e-08)
Sex	351.3*	-1,641***	0.146***	0.122***	0.0715***	0.0543***
	(195.2)	(179.7)	(0.00861)	(0.00832)	(0.00752)	(0.00362)
Age 18	-14,959***	-17,182***	-0.448***	-0.393***	-0.0265**	-0.0662***
	(717.5)	(629.1)	(0.00835)	(0.0150)	(0.0122)	(0.0105)
Age 19	-13,451***	-14,759***	-0.241***	-0.197***	-0.0842***	-0.0750***
	(714.0)	(653.7)	(0.0134)	(0.0167)	(0.0160)	(0.0110)
Age 20	-11,193***	-12,323***	-0.142***	-0.131***	-0.0770***	-0.0598***
	(798.2)	(686.2)	(0.0122)	(0.0131)	(0.0177)	(0.0102)
Age 21	-9,959***	-10,176***	-0.117***	-0.0988***	-0.0725***	-0.0607***
	(715.0)	(693.6)	(0.0183)	(0.0155)	(0.0154)	(0.00903)
Age 22	-7,794***	-8,143***	-0.0886***	-0.0758***	-0.0514***	-0.0444***
	(584.2)	(681.1)	(0.0146)	(0.0120)	(0.0155)	(0.00864)
Age 23	-5,767***	-6,525***	-0.0431***	-0.0539***	-0.0196	-0.0379***
	(606.6)	(621.6)	(0.0140)	(0.0125)	(0.0131)	(0.0100)
Age 24	-4,138***	-4,300***	-0.0347**	-0.0163*	-0.0324**	-0.0186*
	(523.4)	(558.2)	(0.0148)	(0.00908)	(0.0139)	(0.00994)
Year immigrated to the USA X Foreign-Born status	-0.729*	-0.219**	1.78e-05	-6.75e-05***	1.42e-05*	8.59e-06***
	(0.376)	(0.0948)	(1.42e-05)	(6.35e-06)	(8.49e-06)	(3.29e-06)
Born in state of residence	-1,988***		-0.0952***	-0.0455***	-0.0268***	-0.00884
	(406.9)		(0.0100)	(0.00780)	(0.00888)	(0.00638)
MSA share of population aged 25 and older with Bachelor's or higher	-5,863	14,264***	-0.343	0.00667	-0.309	0.212***
	(7,018)	(4,870)	(0.261)	(0.257)	(0.212)	(0.0810)
MSA 2010 unemployment rate	-260.5**	-107.7	0.000843	-0.00558	-0.0203***	-0.00766***
	(115.3)	(90.89)	(0.00321)	(0.00425)	(0.00336)	(0.00207)
MSA Population	1.07e-05	1.07e-05	3.16e-09***	2.00e-09	3.82e-09***	1.28e-09**
	(2.41e-05)	(2.42e-05)	(8.84e-10)	(1.73e-09)	(7.97e-10)	(5.30e-10)
Years of Education Demanded by Average Occupation in MSA	6,574**	-4,348	0.233**	0.178*	0.0355	-0.0632**
	(2,967)	(2,821)	(0.107)	(0.105)	(0.0736)	(0.0318)
Number of own children in the household	-1,150***	-1,629***	-0.0820***	-0.120***	-0.0381***	-0.0320***
	(176.5)	(147.2)	(0.00597)	(0.00680)	(0.00534)	(0.00352)
Wife	-3,567***	-6,889***	0.0270	-0.0321*	2.88e-05	-0.105***
	(1,286)	(609.4)	(0.0314)	(0.0195)	(0.0276)	(0.0184)

Married	5,480***	7,349***	0.0343*	0.00687	0.0552***	0.0784***
	(1,039)	(509.2)	(0.0192)	(0.0195)	(0.0193)	(0.00737)
Has disability (cognitive or ambulatory)	-5,406***	-6,714***	-0.271***	-0.233***	-0.0858***	-0.0680***
	(219.5)	(319.3)	(0.0175)	(0.0134)	(0.0190)	(0.0153)
Constant	-63,921*	79,595**				
	(37,807)	(36,753)				
observed probability			0.48	0.46	0.75	0.83
predicted probability			0.47	0.45	0.75	0.84
Observations	21,662	32,406	21,662	32,406	18,928	28,679
Adjusted R-squared	0.157	0.193				

Robust standard errors in parentheses, clustered on metropolitan areas. *** p<0.01, ** p<0.05, * p<0.1. Columns 1-2 use OLS and columns 3-6 use probit

[1] GreatSchools, available at www.GreatSchools.org. (January 2012).

[2] It was unclear why some schools did not have data, but they were smaller and more likely to be irregular. Of those missing test score data, only 54 percent were "regular" schools, meaning not alternative, not vocational, or not special-ed. This compared to 96 percent of all schools in the database. The regular schools without test score data were also smaller, with average enrollment of 250 students, compared to 556 students for schools with test data.

[3] Hastings and Weinstein, "Information, School Choice, and Academic Achievement."

[4] For a very small percentage of schools, the latest year of data was from 2006 or 2007—0.18 percent total; for all others it was 2008 or later as described in the text above.

[5] New York State Education Department, available at https://reportcards.nysed.gov/databasedownload.php (April 2012).

[6]Virginia Department of Education, available at http://www.doe.virginia.gov/support/nutrition/statistics/index.shtml (April 2012).

[7] Texas Education Agency definition of economically disadvantaged students, available at http://ritter.tea.state.tx.us/cgi/sas/broker (March 2012).

[8] Ohio Department of Education, available at http://ilrc.ode.state.oh.us/Downloads.asp (April 2012).

[9] Illinois State Board of Education, available at http://www.isbe.state.il.us/research/htmls/fall_housing.htm (April 2012).

[10] U.S. Census Bureau; Housing Patterns, definitions of segregation, available at http://www.census.gov/hhes/www/housing/housing_patterns/app_b.html (January 2012).

[11] Rolf Pendall, Robert Puentes, and Jonathan Martin, "From Traditional to Reformed: A Review of the Land Use Regulations in the Nation's 50 Largest Metropolitan Areas," (Washington: The Brookings Institution, 2006).

[12] Rolf Pendall, "Local Land Use Regulation and the Chain of Exclusion" *Journal of the American Planning Association* 66 (2) (2000):125-142; Douglas S. Massey, Jonathan Rothwell, and Thurston Domina, "The Changing Bases of Segregation in the United States" *The Annals of the American Academy of Political and Social Science* 629 (1) (2009): 74-90; Jonathan Rothwell and Douglass S. Massey, "The Effect of Density Zoning on Racial Segregation in U.S. Urban Areas," *Urban Affairs Review* 44 (6) (2009): 779-806; Jonathan Rothwell "Racial Enclaves and Density Zoning: The Institutionalized Segregation of Racial Minorities in the United States," *American Law and Economics Review* 13 (1) (2011): 290-358; Jonathan Rothwell, "Density Regulation and Metropolitan Housing Markets," Working Paper 1154146 (Social Science Research Network, 2009).

[13] Rothwell and Massey, "The Effect of Density Zoning on Racial Segregation in U.S. Urban Areas."

[14] Southern Burlington County NAACP v. Township of Mount Laurel, 336, A. 2d 713 [N.J. 1975]; Norman Williams Jr. and Thomas Norman, "Exclusionary Land Use Controls: The Case of North-Eastern New Jersey," *Syracuse Law Review* (1970-1971): 475-507.

[15] Glaeser and Ward, "The Causes and Consequences of Land Use Regulation"; Mitchell, "Will empowering developers to challenge exclusionary zoning increase suburban housing choice?"

[16] Rothwell "Racial Enclaves and Density Zoning"

[17] Pioneer Institute for Public Policy Research and Rappaport Institute for Greater Boston. 2005. *Massachusetts Housing Regulation Database*. Prepared by Amy Dain and Jenny Schuetz.

[18]Housing Regulation Database, available at http://www.masshousingregulations.com/dataandreports.asp (January 2012).
[19] Rothwell "Racial Enclaves and Density Zoning: The Institutionalized Segregation of Racial Minorities in the United States."
[20] Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.