

Charter School Competition: Evidence From Illinois Public Schools

NaLette Brodnax

DRAFT: DO NOT CITE WITHOUT PERMISSION

STAT 626

December 11, 2014

1 Executive Summary

Over the past twenty years, charter schools have become an increasingly popular alternative to failing neighborhood schools. Though they are publicly funded, charter schools are exempt from many of the state- and district-level policies governing structure, curriculum, personnel, and other aspects of learning. Since the first charter school opened in 1992, over 4,000 charter schools have opened in forty states and the District of Columbia (Zimmer and Buddin, 2009).

Charter schools face a tall order: in exchange for public funding and increased autonomy, they are expected to improve achievement and create competition for neighborhood schools, the administrations of which must adapt or face sanctions. This analysis explores the extent to which traditional public schools (TPSs) respond to competition from charter schools. Using data from Illinois public schools, I test whether proximity to a charter school is associated with an increase in aggregate achievement at the school level. Using Bayesian estimation techniques, I find that charter school competition has a small, positive effect on the share of fourth grade students who meet or exceed the Illinois state standard for math achievement.

The remainder of this report is structured as follows. Section two provides a brief review of recent studies on charter school competition. Section three discusses the sources and construction of the data. Section four comprises analyses of the data using exploratory graphics, linear regression, and hierarchical modeling. Sections five and six discuss the findings and limitations of the analysis, respectively.

2 Literature Review

Table 1 summarizes earlier studies' findings on the effects of charter school competition. There are a number of challenges in comparing the results of the studies. First, researchers measure competition in different ways. Some consider competition to be the percentage of students transferring from TPSs to charter schools, referred to as the charter's market *penetration*, while others measure a public school's *proximity* to one or more charter schools. Second, researchers employ different units of analysis: six of the studies use school or district data, while the remaining use student data. Finally, conclusions are mixed. Early studies of charter school competition show positive effects on the achievement of traditional public schools, while more recent studies find no effects, negative effects, or small positive effects. The mixed results may be attributed to several factors, such as state-level variation in charter school laws, local demography, or the political environment.

In order to avoid bias due to differences in state laws, I analyze charter school competition for just the state of Illinois. Though student-level data is ideal, it is also the most difficult to acquire. I use schools as the unit of analysis due to the ease of obtaining comprehensive data on school location, achievement, and characteristics. I use proximity to the nearest charter school as the key

measure of competition. Proximity measures best capture the competitive effects that arise from charter schools' unique ability to enroll students from outside district boundaries. Eight of the ten studies summarized in Table 1 feature proximity measures.

Table 1: Summary of Competitive Effects on Traditional Public Schools

Author(s)	State(s)	Year	Units	Measure	Effects
Holmes et al.	NC	2003	Schools	Proximity	+
Hoxby	MI,AZ	2004	Schools	Penetration	+
Bohte	TX	2004	Districts	Both	+
Bifulco & Ladd	NC	2004	Students	Both	None
Bettinger	MI	2005	Schools	Proximity	None
Sass	FL	2006	Students	Both	+
Booker	TX	2008	Students	Both	+
Ni	MI	2009	Schools	Penetration	-
Zimmer & Buddin	CA	2009	Students	Both	None
Kamienski	IL	2011	Schools	Proximity	None

3 Data

The state of Illinois enacted charter school legislation in April 1996 and has since authorized over 100 charter schools, two-thirds of which are part of the Chicago Public Schools (Illinois Network of Charter Schools, 2012). Most charters are authorized by local school districts and governed by an independent board of directors. Charter schools enroll 11% of public school students in Chicago and 2.4% of public school students statewide. However, the Illinois Network of Charter Schools estimates that 15,000 students are on waiting lists for existing charters.

School Achievement Data The Illinois State Board of Education publishes annual report card data¹, including school-level assessment results for the Illinois State Achievement Test (ISAT). The ISAT is administered to all students in grades third through eighth in math, reading, and science (Illinois State Board of Education, 2012). Since most students learn to read by the end of third grade, I obtained the 2011 ISAT results in reading and math for all fourth grade students. However, because the findings are similar for both subjects, this report provides only results related to math achievement. ISAT performance is measured by the share of students who fall into one of four categories relative to the state standard: warning status, below standard, meets standard, or exceeds standard. Throughout this report, I refer to the achievement rate as the share of students who meet or exceed the state standard.

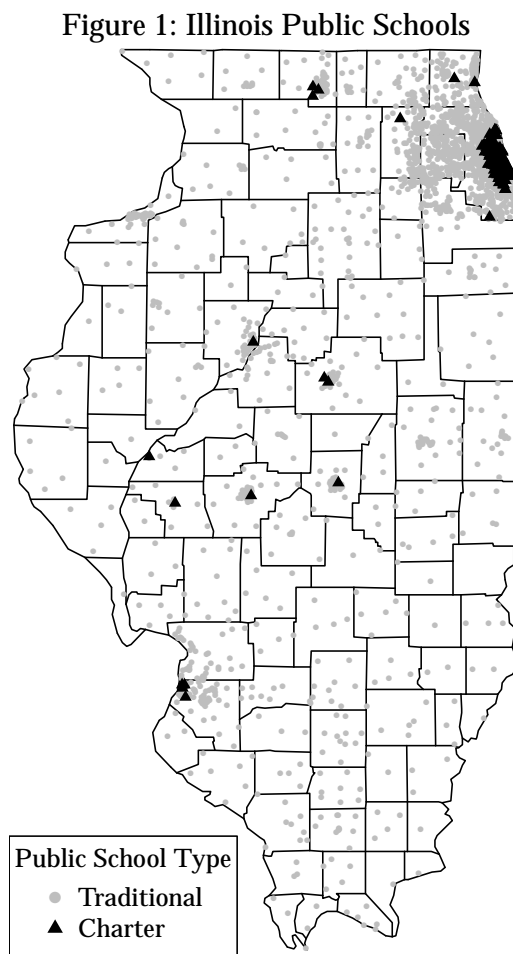
The report card dataset also includes school characteristics, such as total enrollment, enrollment by race, student proportion receiving free or reduced lunch, student proportion with limited english proficiency, and student proportion receiving special education services. The final dataset includes 1,685 elementary schools.

Proximity Data As a measure of proximity, I use the shortest distance in miles between each school and the nearest charter school. Note that distance is inversely related to proximity: as distance increases, proximity decreases. To calculate the distance to the nearest charter school, I

¹See http://www.isbe.net/assessment/report_card.htm

obtained latitude and longitude coordinates for each school from the National Center for Education Statistics². The NCES data includes only the primary campus location for charter schools with multiple locations. In order to incorporate all charter school campuses, I obtained the coordinates manually using the Bing Maps API and a web geocoding application³.

Figure 1 provides a map of traditional public schools and charter schools. Most schools are concentrated in the upper right corner of the map, which is the Chicago metropolitan area. However, there are a number of other clusters with charter schools around the state.



4 Analysis

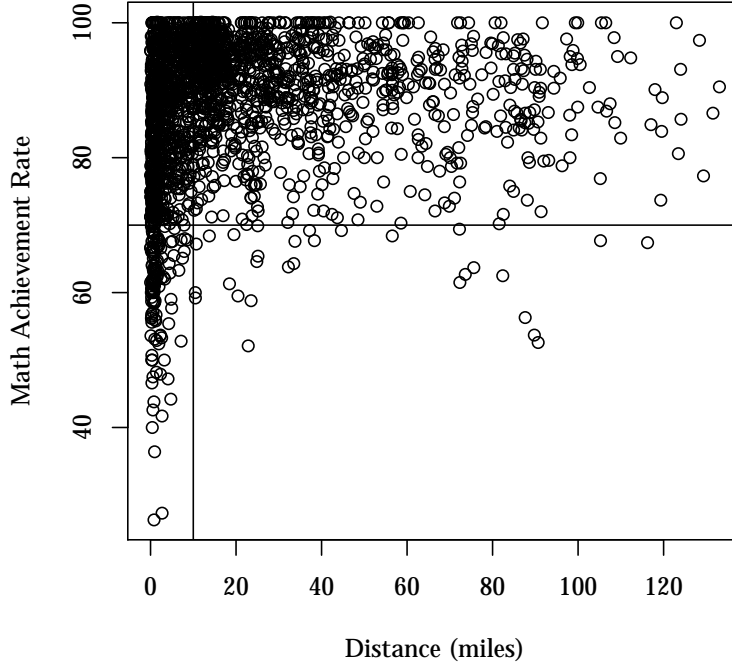
4.1 Exploratory Analysis

An exploratory graphical analysis of the data suggests that traditional public schools in close proximity to a charter school may have different achievement rates than schools located farther away. Figure 2 shows the proportion of students meeting or exceeding the state standards for math in each school, plotted against the distance to the nearest charter school.

²See <http://nces.ed.gov/ccd/elsi/tableGenerator.aspx>

³See <http://www.gpsvisualizer.com/geocoder/>

Figure 2: Math Achievement by Distance to the Nearest Charter School



Schools within a ten-mile radius of a charter school exhibit greater variation in their achievement rates. For those schools with a math achievement rate higher than 80%, the achievement rate appears to be randomly distributed without regard to distance. However, schools with math achievement rates below 70% appear to be clustered within a ten-mile radius of the nearest charter school. The plots confirms what we already know: since charter schools are intended to be alternatives to failing schools, they tend to be located closest to the worst performing schools. However, the relationship between distance and achievement is likely moderated by factors that are not apparent from a graphical review of the data. Accounting for these factors is the topic of the next section.

4.2 Empirical Method: Linear Regression

Linear regression is a powerful technique used to estimate the relationship between two variables. Assuming that the relationship between achievement and distance to a charter school is linear, the achievement rate can be represented by the following model:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon$$

where Y_i is the *achievement rate*, defined as the proportion of students in school i meeting or exceeding the state math standard on the ISAT, and X_i is the shortest distance to the nearest charter school. If proximity to a charter school has a positive effect on TPS achievement, then I expect to see a negative coefficient for $\beta_1(\text{distance})$.

Achievement is influenced by a variety of factors related to school practices and student characteristics. To control for these observed characteristics, I include a number of covariates that are known to explain achievement outcomes. First, achievement is strongly correlated with the racial

and socioeconomic composition of a school, as low-performing schools tend to have high concentrations of poor students and racial minorities (Orfield et al., 2005; Gamoran and Long, 2006; Jencks and Phillips, 1998). Second, students with limited English skills score lower than their peers on achievement tests and are concentrated in poor schools (Reardon and Galindo, 2009; Orfield et al., 2005). Finally, achievement may be influenced by the surrounding neighborhood, which may be in a densely populated urban area or in a rural area. To account for these factors, I include school-level measures for the proportion of black students in the school, the proportion of students with limited English proficiency, the proportion of students receiving free or reduced lunch, and total enrollment. Descriptive statistics for all variables are given in Table 2.

Table 2: Descriptive Statistics

	Mean	S.D.	Min	Max
Achievement Rate	86.807	11.162	26.3	100.0
Proximity	22.901	26.503	0.003	133.062
Pct Black	21.357	33.059	0.0	100.0
Enrollment	395.161	200.666	29.0	1071.0
Pct English Learners	9.677	15.126	0.000	70.0
Pct Low Income	54.809	31.114	0.20	100.0
N = 1685				

Regression coefficients, β_j are typically estimated by minimizing the sum of the squared residuals, a technique known as Ordinary Least Squares. These coefficients can also be estimated with a Bayesian approach: given a sampling model for the data and a set of prior beliefs, posterior beliefs can be approximated via Bayes Theorem. Applied to the regression case, the posterior distribution of β is

$$p(\beta | y, X, \sigma^2) \propto p(y | X, \beta, \sigma^2) \times p(\beta).$$

I assume that the sampling model for y has a multivariate normal form, the prior distribution of β has a conjugate multivariate normal form, and the prior distribution of $\phi = 1/\sigma^2$ has a semi-conjugate gamma form, which give the following posterior distributions:

$$p(\beta | y, X, \sigma^2) \propto \exp \left[\beta^T \left(\Sigma_0^{-1} \beta_0 + \frac{X^T y}{\sigma^2} \right) - \frac{1}{2} \beta^T \left(\frac{\Sigma_0^{-1} + X^T X}{\sigma^2} \right) \beta \right]$$

$$p(\phi | y, X, \beta) \propto \phi^{(\nu_0 + n)/2 - 1} \exp(-\phi [\nu_0 \sigma_0^2 + \text{SSR}(\beta)]/2)$$

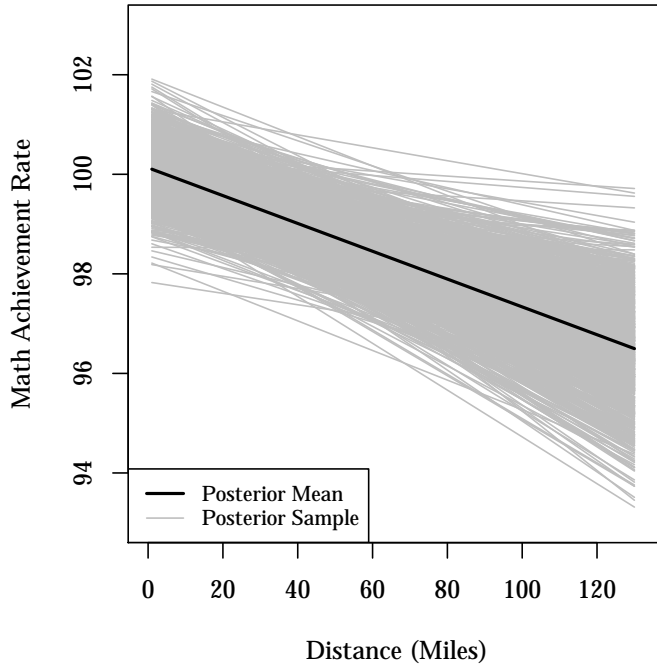
My specification of prior parameters is simplified by using Zellner’s g-prior (Hoff, 2009). Zellner’s g-prior stabilizes parameter estimates, prevents model overfitting, and allows posterior approximation via Monte Carlo sampling. Using achievement data from 2009, I set $\sigma_0^2 = \sigma_{ols}^2$ and $g = n = 1685$, with a weakly-informative prior sample size of $\nu_0 = 1$. Estimation results are given in Table 3.

Accounting for other factors that influence achievement, the coefficient for distance is small and negative. In other words, the smaller the distance between a traditional public school and the nearest charter school (and the greater the proximity), the higher the achievement. The 95% confidence region for $\beta_1(\text{distance})$ does not contain zero, so I conclude that proximity has a small, positive effect on achievement. Figure 3 provides a plot of the posterior β samples generated using Monte Carlo sampling. The samples in gray show that there is considerable uncertainty in the estimates, particularly for distances greater than 60 miles, where the sample of schools is much smaller. Despite this uncertainty, all samples show a negative relationship between distance and achievement.

Table 3: Bayesian Estimation Results for β_j

	β_j	95% Confidence Region
Intercept	99.569	$98.174 < \beta_0 < 101.003$
Proximity	-0.031	$-0.050 < \beta_1 < -0.013$
Pct Black	-0.087	$-0.109 < \beta_2 < -0.065$
Enrollment	0.000	$-0.002 < \beta_3 < 0.003$
Pct English Learners	-0.097	$-0.138 < \beta_4 < -0.058$
Pct Low Income	-0.172	$-0.193 < \beta_5 < -0.149$

Figure 3: Posterior Estimation with Uncertainty (1000 Samples)



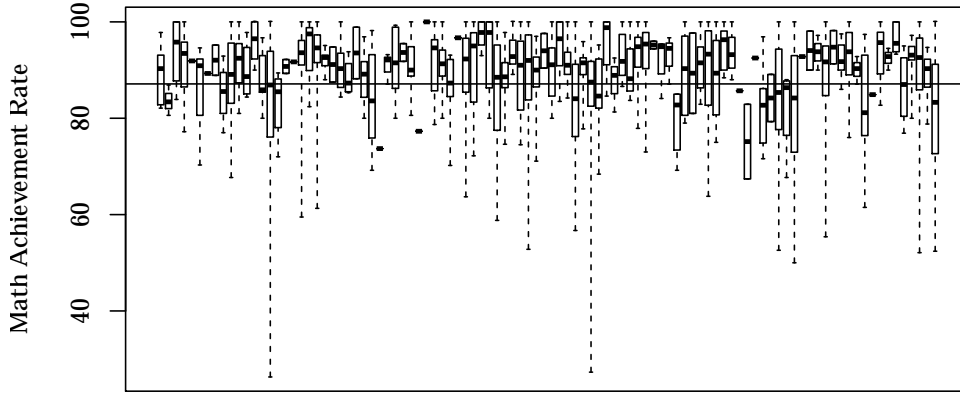
The implication that proximity to a charter school results in higher achievement contradicts the exploratory analysis, which seemed to show higher achievement with larger distances. The regression results demonstrate the importance of controlling for confounding factors, such as racial composition, the proportion of students with limited English proficiency, and the proportion of students in low-income households. However, the linear regression model does not address the spatial structure of the data. The next section considers ways to account for spatial clustering of schools.

4.3 Cluster Analysis: Hierarchical Modeling

A review of the map of Illinois in Figure 1 shows that neither public schools nor charter schools are randomly distributed; rather, they are clustered within counties and districts. Since charter schools are not required to adhere to district boundaries, a district-level analysis may be too noisy. However, most of the geographic clusters appear to be located in distinct counties.

Figure 4 provides box plots of the math achievement rate for all 101 counties containing elementary schools. The mean achievement rate of 89% for all schools is represented by the horizontal bar. Two issues are apparent. First, there is considerable variation between counties, as shown by the locations of the means in each box plot in comparison to the mean for all schools. Second, there is variation within districts, as the individual county boxes are different sizes.

Figure 4: Math Achievement Rates for All Illinois Counties



In order to determine the extent to which achievement is influenced by school factors relative to district factors, I estimate the achievement rate using a hierarchical normal model. Given $n = 1685$ schools clustered in $m = 101$ counties⁴ and a normally distributed sampling model, Gibbs sampling can be used to approximate the posterior distributions from the full conditional distributions of four key parameters:

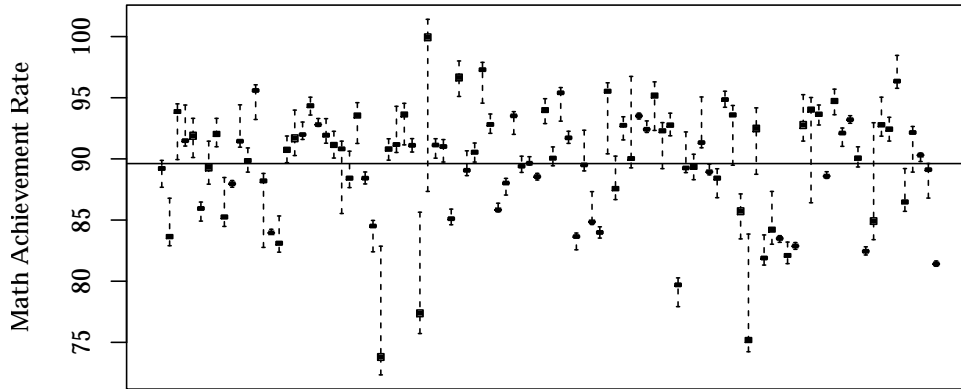
1. The mean achievement rate for each county, $\theta_1, \dots, \theta_m \sim \text{normal}(\mu, \tau^2)$;
2. The mean achievement rate of the population of counties, $\mu \sim \text{normal}(\mu_0, \gamma_0^2)$;
3. The variance for the population of counties (between-group variation),
 $\tau^2 \sim \text{inverse-gamma}(\eta_0/2, \eta_0\tau_0^2/2)$;
4. The variance for each sample of schools (within-group variation),
 $\sigma^2 \sim \text{inverse-gamma}(\nu_0/2, \nu_0\sigma_0^2/2)$

I implemented the Gibbs sampler with 5,000 iterations, using the mean and variance of the 2009 achievement rates as prior starting points: $\mu_0 = 84.7$ and $\sigma_0^2 = \tau_0^2 = 169.2$. I set the population variance prior $\gamma_0^2 = 25$ in order to cap the upper bound of the achievement rate at 100. Prior sample sizes reflect a high degree of uncertainty at $\eta_0 = \nu_0 = 2$. The posterior achievement rates for all counties are shown in Figure 5 with a line for the posterior mean of 89.6%. The distribution

⁴There are 102 counties in Illinois. For future analyses, I intend to investigate which county is missing and assess the impact of the missing data on analysis outcomes.

of the posterior samples is consistent with the degree of variation shown across counties in Figure 4. The posterior mean is ≈ 3 percentage points higher than the mean achievement rate in the 2011 data, despite a prior (but weakly informative) rate of 84.6%.

Figure 5: Posterior Achievement Rates by County



Gibbs Sampler Convergence In order to check whether there were any problems with the Gibbs sampler, I compute effective sample sizes and review stationarity plots to check for convergence. The stationarity plots for the posterior samples of μ , τ^2 , and σ^2 are shown below. All samples have converged well before the 5,000th iteration. Effective sample sizes for μ , τ^2 , and σ^2 are sufficiently large, at 5,000, 4,433, and 5,000, respectively. The effective sample sizes for $\theta_1, \dots, \theta_m$ are also sufficiently large, ranging from 4,469 to 5,000.

Figure 6: Convergence of Population Mean

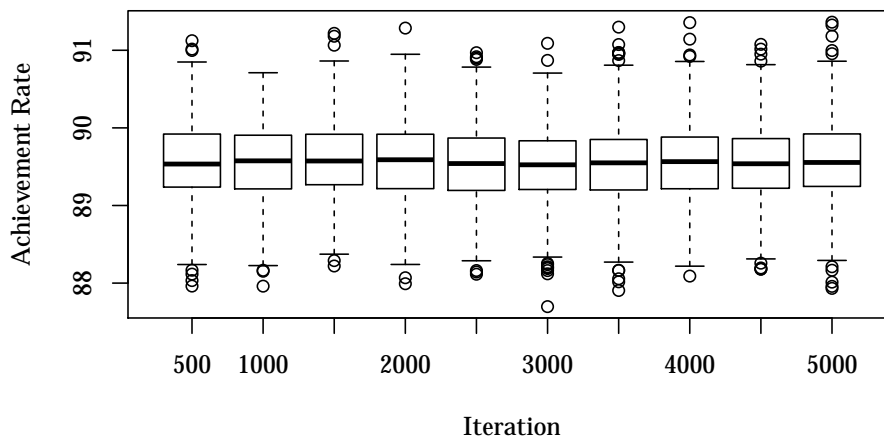


Figure 7: Convergence of Between–County Variation

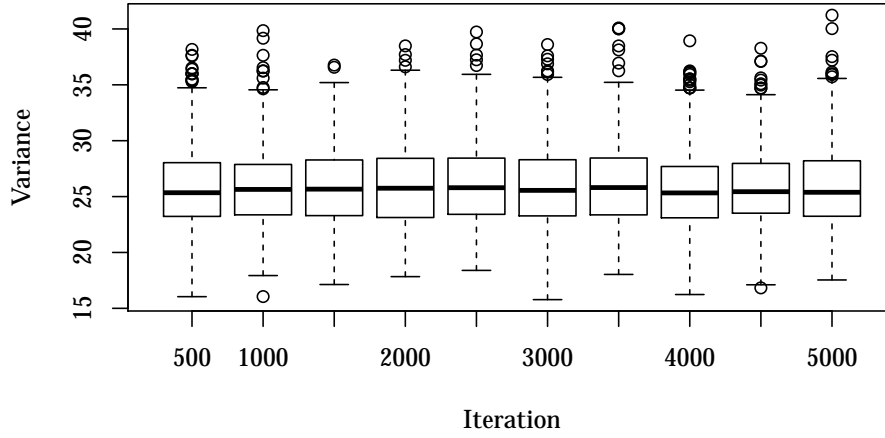
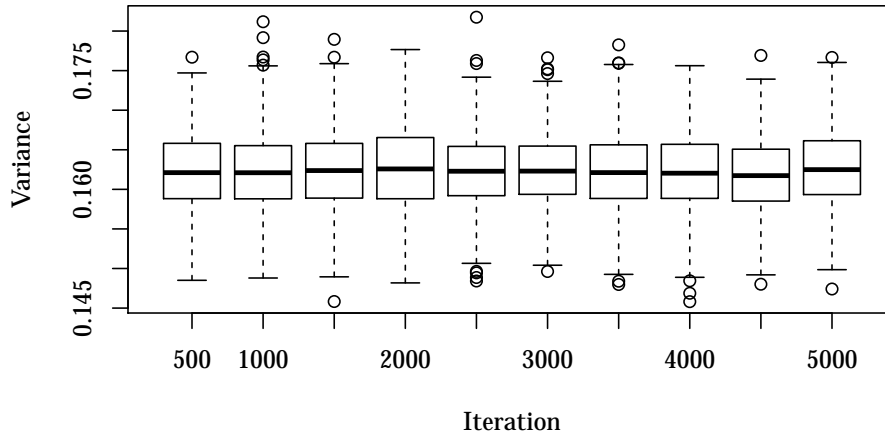
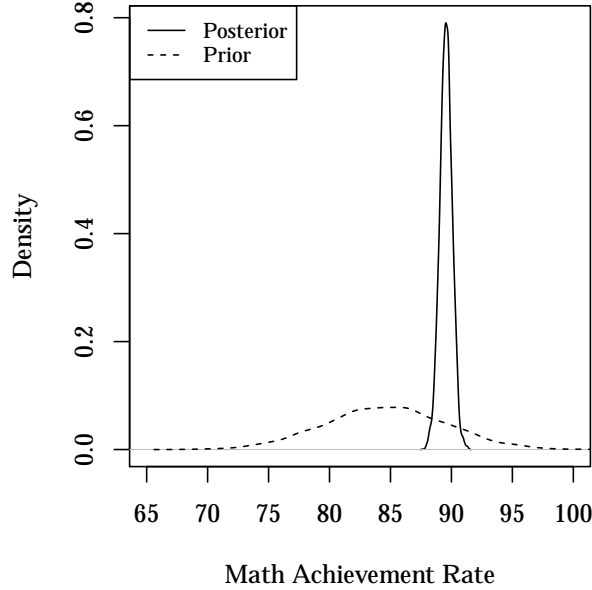


Figure 8: Convergence of Within–County Variation



Posterior Inference The parameter μ is a “hyper-parameter” for the mean of the population of counties. Its posterior mean $E[\mu|\theta_1, \dots, \theta_m, \tau^2]$ is 89.55, with a 95% confidence region between 88.53 and 90.52. The posterior and prior densities of the mean achievement rate are shown in Figure 9 below. In comparison to the prior, the posterior density has a higher mean and much smaller variance.

Figure 9: Population Mean



The parameter τ^2 is also a hyper-parameter for the variance of the population of counties. It is commonly referred to as *between-group* variation. The posterior mean for $\tau^2|\theta_1, \dots, \theta_m, \mu$ is 25.91, with a 95% confidence region between 19.64 and 34.22. This means that 95% of the county means will fall within ≈ 5 percentage points of the population mean—a substantively large amount of variation. A comparison of the posterior and prior densities of τ^2 is given in Figure 10.

The parameter σ^2 is the variation between individual schools within a given county, commonly referred to as *within-group* variation⁵. The posterior mean for $\sigma^2|\theta_1, \dots, \theta_m, y_1, \dots, y_m$ is 0.162 and the 95% confidence region is between 0.152 and 0.172). In comparison to the between-group variation, within-group variation is much smaller. The posterior and prior densities of σ^2 are given in Figure 11.

⁵For the purpose of this analysis, I assume that σ^2 is constant. However, this is a strong assumption given the within-county variation that is present in Figure 4. It is possible to specify a hierarchical model that allows σ^2 to vary by school.

Figure 10: Between–County Variation

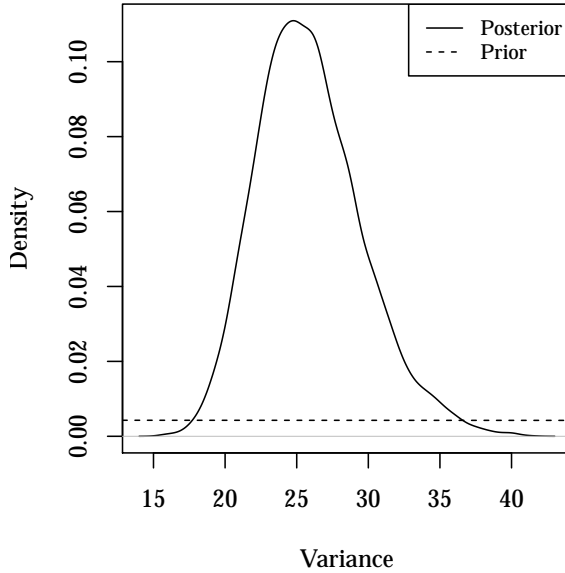
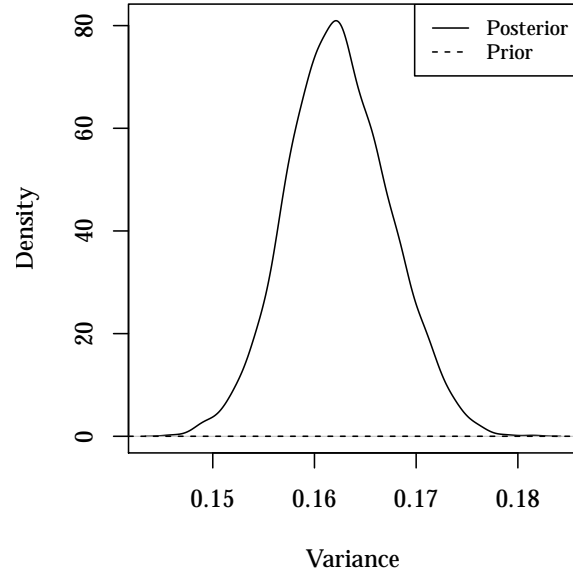


Figure 11: Within–County Variation



5 Findings and Implications

This study aims to test the relationship between school-level achievement and proximity to a charter school, resulting in three key takeaways. First, the linear regression results provide evidence that proximity to a charter school has a small, positive effect on a traditional public school’s math achievement rate. The coefficient on the proxy for proximity, *distance*, is -0.031: the farther away the nearest charter school, the lower the achievement rate.

Second, math achievement is strongly correlated with student characteristics. The coefficients in the regression model for the proportion of black students, the proportion of students with limited English proficiency, and the proportion of students in low-income households are all negative and significant, at -0.087, -0.097, and -0.172, respectively. The negative effect of having a larger proportion of low-income students is more than five times the positive effect of being located near a charter school.

Third, the cluster analysis via hierarchical modeling demonstrates a strong relationship between location and achievement. Variation in achievement rate is much larger between counties than between schools within counties.

The findings of this analysis are relevant for policymakers who continue to champion charter schools as beneficial for students. The results of this study suggest that when charter schools open in a neighborhood, even students who do not attend the charter school may benefit from its presence. However, additional work is needed to better estimate the effect of charter school proximity on traditional public schools.

6 Limitations

Statistician George E. P. Box remarked that, “Essentially, all models are wrong, but some are useful” (Box and Draper, 1987). While I find the results of this analysis useful for policymakers

and education practitioners, the comments below highlight aspects of the models that are known to be ‘wrong’ and suggest ways to improve them.

1. **Lack of Causal Inferences** This study relies on observational data and therefore does not feature a precise treatment and counterfactual group on which to base causal inferences. Though covariates are included in the regression model, these do not account for unobserved differences between schools which may confound the effect of being located near a charter school. One way to reduce bias would be to obtain achievement data for numerous years and use a fixed-effect specification to control for unobserved, time-invariant factors.
2. **Poorly Fit Sampling Model** For the regression and hierarchical models, I assumed that school achievement rates are normally distributed. While the normality assumptions allows ease of computation, achievement rates may be better fit by a different distribution. Since the achievement rate is always a share of the total student population, it may not conform to a normal distribution. Ferrari and Cribari-Neto (2004) recommend beta regression for parameters that are continuous and bound to the interval (0,1), such as proportions. While they advocate maximum likelihood estimation for parameters, I would investigate ways to use beta regression within a Bayesian framework.
3. **Inadequate Hierarchical Structure** This analysis attempted to explicitly model the spatial structure of the data in order to consider the impact of this structure on estimation results. However, there are two issues with this approach. First, the hierarchical model includes only two levels: counties and schools. A more comprehensive model might include an additional level for school district or neighborhood. Second, the hierarchical model lacks controls for confounding factors at the school level. The hierarchical model can be adapted to a regression framework so that additional information can be incorporated at each level (Hoff, 2009).

In conclusion, this analysis makes an important contribution to the debates about the role of charter schools in education reform. Future studies should address not only the limitations of this analysis, but also the degree to which the results can be generalized to other settings.

References

- Bettinger, E. P. (2005). The effect of charter schools on charter students and public schools. *Economics of Education Review*, 24:133–147.
- Bifulco, R. and Ladd, H. F. (2004). *The impacts of charter schools on student achievement: Evidence from North Carolina*. Terry Sanford Institute of Public Policy, Duke University, Durham, NC.
- Bohte, J. (2004). Examining the impact of charter schools on performance in traditional public schools. *Policy Studies Journal*, 32(4):501–520.
- Booker, K., Gilpatric, S., Gronberg, T., and Jansen, D. (2008). The effect of charter schools on traditional public school students in Texas: Are children who stay behind left behind? *Journal of Urban Economics*, 64:123–145.
- Box, G. E. P. and Draper, N. (1987). *Empirical Model Building and Response Surfaces*. John Wiley & Sons, New York, NY.
- Ferrari, S. L. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31:799–815.
- Gamoran, A. and Long, D. (2006). *Equality of Educational Opportunity: A 40-Year Retrospective (WCER Working Paper No. 2006-9)*. University of WisconsinMadison, Wisconsin Center for Education Research, Madison.
- Hoff, P. D. (2009). *A First Course in Bayesian Statistical Methods*. Springer, New York, NY.
- Holmes, G. M., DeSimone, J., and Rupp, N. G. (2003). *Does school choice increase school quality?* National Bureau of Economic Research, Cambridge, MA.
- Hoxby, C. M. (2004). *Achievement in charter schools and regular public schools in the United States: Understanding the differences*. National Bureau of Economic Research, Cambridge, MA.
- Illinois Network of Charter Schools, I. (2012). Profiles of illinois charter public schools.
- Illinois State Board of Education, I. (2012). *2009-2010 and 2010-2011 Illinois Charter School Biennial Report*.
- Jencks, C. and Phillips, M. (1998). *The black-white test score gap*. Brookings Institution Press, Washington, D.C.
- Kamienski, A. (2011). Competition: Charter and public elementary schools in Chicago. *Journal of School Choice*, 5:161–181.
- Ni, Y. (2009). The impact of charter schools on the efficiency of traditional public schools: Evidence from Michigan. *Economics of Education Review*, 28:571–584.
- Orfield, G., Lee, C., and University), C. R. P. H. (2005). *Why segregation matters : poverty and educational inequality*. Civil Rights Project, Harvard University, Cambridge, Mass.
- Reardon, S. and Galindo, C. (2009). The hispanic-white achievement gap in math and reading in the elementary grades. *Am.Educ.Res.J.American Educational Research Journal*, 46(3):853–891. Journal Article.

- Sass, T. R. (2006). Charter schools and student achievement in Florida. *Education Finance and Policy*, 1:91–122.
- Zimmer, R. and Buddin, R. (2009). Is charter school competition in California improving the performance of traditional public schools? *Public Administration Review*, 69:831–845.