DIGITIZATION AND DIVERGENCE: ONLINE SCHOOL RATINGS AND SEGREGATION IN AMERICA

Sharique Hasan
Duke University
Fuqua School of Business
Durham, NC
sharique.hasan@duke.edu

Anuj Kumar
University of Florida
Warrington College of Business
Gainesville, FL
akumar1@ufl.edu

November 8, 2018

Abstract

We analyze whether widespread online access to school quality information affected economic and social segregation in America. We leverage the staged rollout of GreatSchools.org school ratings from 2006-2015 to answer this question. Across a range of outcomes and specifications, we find that the mass availability of school ratings has accelerated divergence in housing values, income distributions, education levels, as well as the racial and ethnic composition across communities. Affluent and more educated families were better positioned to leverage this new information to capture educational opportunities in communities with the best schools. An unintended consequence of better information was less, rather than more, equity in education.

JEL Classification: I21, I3

1 Introduction

Digitization and especially the Internet are transforming social and economic life (Brynjolfsson and McAfee, 2014; DiMaggio and Bonikowski, 2008; Jorgenson, 2001). Because of these advances, individuals today can access extraordinary amounts of information to help them make important decisions. Job seekers, for example, can readily find ratings of workplaces; patients, ratings of hospitals; and parents, ratings of public schools. Although we know much more about how this information affects individual choices (e.g., Santos, Gravelle and Propper, 2017; Luca and Smith, 2013; Salganik, Dodds and Watts, 2006), our understanding of the broader social consequences of this mass availability of information remains limited.

Among the most critical decisions an American family makes is choosing *where* to live (Kane, Riegg and Staiger, 2006). For many, a crucial input to this decision is the quality of a community's public schools (Gibbons, Machin and Silva, 2013; Nguyen-Hoang and Yinger, 2011). Families have generally learned about school performance informally or inconsistently through social networks, real-estate agents, and other sources (Mikulecky and Christie, 2014; Figlio and Lucas, 2004). As such, families deciding where to live face substantial uncertainty due to the lack of consistent and accessible measures of school performance. Despite having limited information, families' choices have still led neighborhoods to diverge economically (Owens, Reardon and Jencks, 2016). Today, however, parents have access to a substantial amount of school quality information online. In this article, we ask whether the widespread access to school performance information has accelerated social and economic divergence.

We answer this question by leveraging zip code–year variation in the nationwide expansion of GreatSchools.org ratings. GreatSchools, a non-profit based in Oakland, California, provides detailed information about school quality for over 100,000 public schools across America. The mission of GreatSchools is to empower parents by providing detailed information about public schools. In 2003, GreatSchools expanded its ratings beyond its original state of California. Our data shows that coverage increased from 4,643 zip codes across five states in 2006 to 20,551 zip codes across all 50 states in 2012. We use the rollout of school ratings to test whether this mass increase in school information accelerated or slowed the inequality in the home prices as well as the economic character of communities.

Across a range of specifications, we find that widespread access to school performance ratings accelerated divergence across zip codes. In our most conservative models, we find that housing prices for zip codes that are 2-sd apart in school performance diverge by an additional $2,249 after 1 year of rating availability and $8,996 after four years of availability. Further, we link rating availability to between a 0.11% and a 0.4% change in the proportion of top income earners in a zip code. Finally, we find greater adjustment in the White and Asian population within communities in response to rating availability, with the proportion of these racial groups increasing in better school districts. We find no such effect for African Americans and a stronger negative relationship between school quality and the Hispanic population when ratings are available online.

Our results speak to several streams of research in the social sciences. First, our study is one of the first to propose and test a novel mechanism for the increasing economic divergence across American communities (Reardon and Bischoff, 2011). We show that broader access to information increased segregation because high-income families could more readily leverage school ratings to move to neighborhoods with better schools. In this case, knowledge was indeed power, but only for the powerful. Second, our results speak to the growing literature on the social and economic impacts of digitization (Brynjolfsson and McAfee, 2014). Our research shows that the widespread availability of information enabled by the internet can have society-wide, and often unintended, effects. Finally, our results broaden the scope of the emerging research on online rankings by showing how they affect the outcomes of entire communities, not just consumers (Shore et al., 2015; Luca and Smith, 2013; Sauder and Espeland, 2009; Espeland and Sauder, 2007).

2 Literature Review

Today, an uneasy tension exists between two trends in American society. On the one hand, income and wealth inequality is on the rise (Piketty and Saez, 2006, 2003). Growing inequities across individuals and families have led to both greater income segregation in American communities (Reardon and Bischoff, 2011) and divergence in access to economic opportunities (Owens, Reardon and Jencks, 2016). On the other hand, more Americans than ever have access to the Internet and vast amounts of information to aid in their decision making (Pew, 2018). This widespread access promises to democratize knowledge and give all citizens the ability to find and take advantage of better opportunities (Brynjolfsson and McAfee, 2014).

One area where this tension is increasingly unfolding today is in the link between information, inequality and the access to an essential public good: schools. Although research indicates that access to good schools is highly unequal, parents today have unprecedented access to quantified school performance information online (Mikulecky and Christie, 2014). This information, some argue, gives parents—even lower-income ones—an essential tool for improving their children's education. At the same time, it may be the high-income families who best leverage this information to find and capture the best opportunities (Reeves, 2017). A fundamental question is whether access to school performance information online has helped to slow economic and social divergence, or to accelerate it.

2.1 Income Segregation and Inequality in Schooling

The rapid increase in top incomes, combined with the relative stagnation of wages for lower- and middle-income households, has reshaped many aspects of life in American communities. Reardon and Bischoff (2011), for instance, found that as income inequality increased from the 1970s into the 2000s, spatial segregation based on income grew as well. Furthermore, as wealthy families became geographically concentrated, patterns of income and racial segregation further accelerated (Reardon, Fox and Townsend, 2015; Jargowsky, 1996).

A consequence of this rise in income segregation was the effect that it had on American public education—a primary engine of economic opportunity (Coleman et al., 1966). Owens, Reardon and Jencks (2016), for instance, find that the effect of segregation led to a dramatic shift in the composition of school districts, with high-income families having disproportionate access to better quality schools (see, also: Reeves, 2017). One implication of this segregation is that without access to good schools, the educational achievement of lower-income students is diminished (Ziol-Guest and Lee, 2016; Reardon, 2011; Mayer, 2002). As Quillian (2014) documents, income segregation negatively affected not just the high school graduation rates of poor students, but also longer-term college attendance and graduation rates (see also, Mayer, 2002). In addition to the educational divergence between poor and non-poor households, income segregation appears to have exacerbated existing racial gaps in educational access and outcomes (Quillian, 2014; Logan, 2011; Sampson, Sharkey and Raudenbush, 2008; Rumberger and Palardy, 2005).

2.2 Accountability and School Performance

A fundamental question asked by both policymakers and scholars is whether these performance gaps can be reduced. Some policymakers have taken the view that holding schools accountable for student outcomes may be one mechanism to improve school performance and reduce gaps (Harris and Herrington, 2006). Although interventions around the issue of accountability are somewhat diverse, two primary accountability mechanisms are widespread: (1) the administration of standardized testing via *No Child Left Behind* and (2) public availability of quantified performance measures for schools based on test results (e.g., Figlio and Rouse, 2006). Efforts at accountability are multifaceted and include both public and private initiatives. The State of Florida, for instance, assigned letter grades to schools in a 'Report Card' about performance (Figlio and Lucas, 2004). In addition to public efforts, organizations such as GreatSchools.org, city-data.com, and 50Can.org also collect performance data and publish measures of school quality for use by parents and others (Mikulecky and Christie, 2014).

In theory, these quantified and widely accessible school quality measures should serve as an important tool for parents in improving their children's educational options. First, when parents have more information about the quality of their children's schools, they can be more informed advocates for improving the schools their child attends.[1] Second, with more information about the quality of other schools outside their current district, parents can relocate to better neighborhoods with higher performing public schools. Finally, rankings and ratings should ostensibly cause educational organizations to change in response to being evaluated (Shore et al., 2015; Espeland and Sauder, 2007; Sauder and Espeland, 2009; Espeland and Stevens, 1998).

A principal argument in the accountability narrative above is that individuals make better choices with more information. The *information* mechanism has growing support in the literature. In a variety of settings, researchers have found that individuals do respond dramatically to quality information. Individuals are significantly more likely to select better ranked or rated options over lower rated ones (Salganik and Watts, 2008; Salganik, Dodds and Watts, 2006; Chevalier and Mayzlin, 2006). In health care, for instance, Santos, Gravelle and Propper (2017) find that public information on doctor quality led to an increase in demand for high-quality physicians. Varkevisser, van der Geest and Schut (2012) find similar results

[1]GreatSchools.org, for instance, describes itself as "the leading national nonprofit empowering parents to unlock educational opportunities for their child." (GreatSchools About Page) *Accessed: August 28, 2010 2:08pm EST.*

for patients selecting cardiologists. Similarly, Pope (2009) finds that hospitals that improved in 'America's Best Hospitals' rankings saw significantly increased demand. Rankings also have a profound effect on the demand for educational institutions. Luca and Smith (2013), for instance, find that colleges ranked higher in *U.S. News and World Report College Rankings* received more applications.

Furthermore, research also suggests that this information is particularly valuable for disadvantaged students who may have gaps in their knowledge about where opportunities exist. Jensen (2010), for instance, finds that providing basic information about the financial returns from schooling increases educational persistence. Hoxby and Turner (2013) find that low-income students have limited knowledge about elite colleges, and simple mailers can dramatically increase their likelihood of applying to and attending these schools. These results, however, reflect 'partial' treatment effects and ignore the systemic effects of providing more information more broadly to both disadvantaged and wealthy families.

In contrast to these findings, it is possible that the benefits of quality information may accrue mostly to those who can act on it. Wealthier households are often better able to take advantage of new information about school quality. Figlio and Lucas (2004) in a study of one large school district in Florida, found that when schools were assigned performance 'grades,' housing prices adjusted to reflect school quality. Wealthier families in this district sorted into more expensive neighborhoods with better schools after ratings became available. While their study is valuable in that it can link the availability of performance measures to choices, substantial literature has long found that school quality is capitalized in housing values (Gibbons, Machin and Silva, 2013; Fack and Grenet, 2010; Kane, Riegg and Staiger, 2006; Brasington and Haurin, 2006). High-income households are willing to pay a premium for homes with better quality schools—approximately 4% more for a 1-standard deviation better school (see for a review, Nguyen-Hoang and Yinger, 2011).[2] As a consequence,

[2]Estimates from other authors approximates this effect size. Brasington and Haurin (2006) finds

school quality information should affect the choices of wealthier families more dramatically than lower-income families. Thus, in aggregate, the availability of school quality information should cause high-income families to leave zip codes with low-performing schools and move to higher-performing communities.

The gradual availability of online ratings of public schools by GreatSchools provides an unusual opportunity to estimate the effect of providing mass information about school performance on neighborhood composition and divergence across America. There is ample evidence that home prices and school quality are highly correlated (Nguyen-Hoang and Yinger, 2011).

As online ratings become available, we predict an upward shift in home prices for communities with better performing schools and a downward shift for lower-performing ones. This effect is reflected in the increased slope in the School Performance and Housing Value relationship. Due to ratings availability, we expect an additional decline in home prices for lower performing districts and an equivalent rise in prices for better-performing ones.

Furthermore, this shift in home values should also affect the economic and demographic composition of the affected communities. After rating availability, we predict based on prior research (e.g., Quillian, 2014) that higher performing school districts should see an increase in higher income and college educated households, with more residents who are White or Asian, relative to those who are African American or Black and Latino (Logan, 2011). Finally, we should expect higher rates of in-migration for communities where ratings are available for desirable, high-performing schools.

a 7.1% increase in housing prices for a 1-sd increase in school quality; Fack and Grenet (2010) in a study in France finds that housing values increase by 1.4 to 2.4% for a 1-sd increase in quality.

3 Empirical Strategy

Our empirical analysis estimates the effect of the availability of school ratings via the Internet (using the gradual availability of GreatSchools.org ratings as our proxy) on the changing economic and social character of American communities. Towards this goal, we combine several data sources. Our data are at the zip code-year level and include information on: (1) GreatSchools.org rating availability and average school quality; (2) housing prices; (3) proportion of high-income households; (4) racial and ethnic composition; (5) migration patterns; and (6) pre-GS data from 19 states' Departments of Education. Below we describe our data sources, the construction of our variables, and estimated models.

3.1 Data

GreatSchools.org: GreatSchools.org (hereafter, GS) is a national educational non-profit based in Oakland, California. It develops and disseminates quantitative ratings of thousands of American public schools based on the standardized test performance of their students. According to its current website, GS provides:

> *…easy-to-understand information on K-12 schools, including ratings, information on school resources and student outcomes, and reviews.*

GS computes ratings using government-administered standardized test scores in subjects including mathematics, reading, and science. Although the actual test scores used to compute the GS ratings differ in their content and measurement, GS normalizes these ratings into a decile scale, ranging from 1 through 10. The ratings are also color-coded to reflect quality differences, with green, orange, and red reflecting high, medium, and low performance, respectively. Figure 1 depicts examples of schools with ratings on the GS website and on the real estate website, Zillow.com.

[Figure 1 about here.]

Our analysis leverages the school performance data available to GreatSchools beginning in 2006. At that time, the GS database included data on five states and 4,643 zip codes. By 2012, GS covered all 50 states and about 20,551 zip codes in its database. From 2013 to 2015, GS maintained information on nearly 80,000 schools. Table 1 presents the increase in coverage of GS data with respect to the number of states, zip codes, and schools from 2006 to 2012.

[Table 1 about here.]

Home Prices: Zillow: Zillow.com is an online real estate platform and database. We acquired zip code level housing value data from Zillow.com's research database. Our primary dependent variable, *Housing Prices*, is derived from an aggregate measure of the value of all homes in a zip code called the Zillow Home Value Index or (ZHVI). The ZHVI, like the Case-Shiller Index, uses deed data for single-family homes, but also estimates sales prices for each home in a geographic area based on the characteristics of the home, tax assessment, sales transactions, and location using a hedonic approach (Dorsey et al., 2010). Prior research has found that the ZHVI is highly correlated to other standard home price indices (e.g., CSI), with $\rho = .96$ (Guerrieri, Hartley and Hurst, 2013), and has more comprehensive coverage as well (Damianov and Escobari, 2016). The measurements of the ZHVI are in dollars and are provided for each month beginning in 1997 until 2016, with scope increasing from 14,276 to 15,417 zip codes.

Household AGI Categories: Internal Revenue Service: The Internal Revenue Service (IRS) publishes an annual database of Individual Income Tax Statistics at the zip code level.[3] We use the tax statistics database available through the National Bureau of Economic Research that includes information on the number of tax returns in each zip code, returns by adjusted gross income (AGI), exemptions, and

[3] This data can be found Here

other tax return items. Most relevant to our analysis are the number of households at each of the following 6 AGI levels: (1) \$1 to under \$25,000, (2) \$25,000 to under \$50,000, (3) \$50,000 to under \$75,000 , (4) \$75,000 to under \$100,000, (5) \$100,000 to under \$200,000, (6) \$200,000 or more. The IRS data cover the years beginning in 2005 until 2015, with zip code coverage ranging from 38,499 in 2005 to 27,680 in 2015.

Racial and Ethnic Composition, Education & Migration: American Community Survey: We use the American Community Survey (ACS) data product from the US Census Bureau to gather estimates of the racial and ethnic composition of zip codes. The ACS provides estimates of a zip code's total population, as well as population size by race and ethnicity (White, Black or African American, Asian, and Hispanic and Latino). This data was obtained from ACS Demographic and Housing estimates available here on the American factFinder product of US Census Bureau. We further obtained data on educational attainment of populations at the zip code level from the ACS Educational Attainment estimates available here and data on migration into the zip codes from the ACS Selected Social Characteristics estimates available at here.

3.2 Empirical Model

Our analysis examines whether the availability of GreatSchools.org ratings for a zip code i at time $t-1$ affected its economic and social composition in time t. Further, we hypothesize that the effect of online rating availability had an asymmetric effect, depending on the quality of the schools in that zip code i. When ratings went online for high-performing schools, home prices increased, communities become wealthier and additional White, Asian, and educated residents move in. In contrast, when ratings became available for low performing schools, home prices decreased, and wealthier, White and Asian families left. Below we describe the general specification of our empirical models, issues around the identification of our results, and the construction of the independent and dependent variables used to examine these effects.

Our basic model is of the following form:

$$\begin{gathered} Y_{it} = \beta_1 SP_{i(t-1)} + \beta_2 YearsAvail_{it} + \\ \beta_3 (SP_{i(t-1)} \times YearsAvail_{it}) + \alpha_x + \epsilon_{it} \end{gathered} \tag{1}$$

Equation 3 is a panel model which exploits the variation in zip code level characteristics over several years of our period of observation. In this model, variable Y_{it} denotes the dependent variables including housing values, high-income household share, and ethnic and racial composition in zip code i in a year t. The variable SP_{it-1} reflects the standardized performance of schools in a zip code i for the prior year, $t-1$. Finally, the variable $YearsAvail_{it}$ denotes the number of years for which GreatSchools data have been available for the schools in zip code i by year t. We code the year in which the score is introduced as '0' with subsequent years iterating by +1. Finally, the main coefficient of interest in our model is β_3, which estimates the interaction effect of school performance and the exposure effect of rating availability.

In this specification, we use previous year school performance as an independent variable for two reasons. First, the information about school performance in standardized tests in an academic year is generally available at the end of the academic year. Second, it reduces the possibility of reverse causality - the dependent variable (say a higher percentage of high-income households) in a year causing better school performance in that year (prior literature suggests that children of high-income families perform better due to the availability of higher resources). However, the dependent variable in year t cannot affect the school performance in a year $t-1$.

Nevertheless, our model is still susceptible to omitted variable bias which can make β_3, the effect of rating availability at each level of school performance, hard to interpret. A variety of geography and time-related factors may affect the choice to provide school ratings for a community earlier or later, thus biasing β_3, and thus our interpretation. In our models, we primarily deal with these unobserved selection factors and demographic trends using several demanding fixed-effects specifications, which are generically denoted by α_x in Equation 3. In our baseline specifications, we use zip code and year-fixed effects. The zip code fixed effects capture differences in the scale of housing prices across communities. The year-fixed effects capture differences in scale across years (e.g., the overall change in prices across the United States, including periods such as the financial crisis in 2008). In even more demanding specifications, we include state-year (α_{st}) and then county-year fixed effects (α_{ct}) that capture non-parametric trends in housing values and school performance across states or counties. Including these variables allow us to account for unobserved time-varying shocks such as changes in policies, investments, or business dynamics at a state or county level in a year.

Finally, as long as GreatSchools does not have a systematic selection process for the schools and zip codes for which it publishes ratings (e.g., rating availability ($YearsAvail_{it}$) is exogenous to school quality and or housing values), we can interpret the coefficient β_3 as unbiased.[4] The basic intuition for this assumption is that after controlling for the endogenous school quality as a covariate, the coefficient of its interaction with rating availability in the specification will be unbiased.

To get an unbiased estimate of β_3, the decision to make ratings available for a zip code i at a given time t should be unrelated to a zip code's pre-rating levels or trends.

In aggregate, there does not seem to be a systematic bias in how GreatSchools chose the states or zip codes in which it made its ratings available. By 2007, the

[4]In Equation 3, it is worthwhile noting that the coefficient of SQ_{it} may remain biased because a variety of factors may simultaneously affect the school test scores and the house prices in a zip-code. For example, due to the entry of a large firm, more jobs, or other economic shocks, new, more highly educated workers may settle in that area and school performance may shift. However, we do not anticipate that such economic dynamics should affect whether GS is more or less likely to provide ratings for that zip code.

earliest years of our data, GS had already published test scores of five states: CA, CO, IL, NC, and TX. By 2009, test scores for five additional states (IA, NM, WV, MI, and GA) were published. By 2012, the website had ratings available for schools in all 50 states. In supplementary analysis[5], we compare the total number of schools (school districts) in a state for which test results are available on the website of the state Department of Education with the number of schools for which GS published the test scores in the year of introduction of that state on the GS website. We find that GS publishes the performance of all (or close to all) schools in a state in the first year and does not 'cherry pick' either the schools or geographical areas in the state for which to publish test results/ratings.

Next, we test whether there is a systematic bias in the timing of when GS ratings are made available for certain zip codes. To test this, we regress zip code characteristics on the duration of rating availability with the following equation:

$$Y_{it} = \alpha_s + \alpha_t + \beta_1 YearsAvail_{it} + \epsilon_{it} \tag{2}$$

In Equation 2 the variables have the same interpretation as in Equation 3. In our models, we include state- (α_s) and year- (α_t) fixed effects to account for the state and year level unobserved factors. These fixed effects account for time trends and large-scale state-level geographic dispersion in zip code characteristics. The sources of unobserved time-varying heterogeneity include the impact of changes in laws, the financial crisis, and other macroeconomic factors that would be likely to impact heterogeneity and be correlated with time at a national scale.

Accounting for these factors, the coefficient β_1 should be significantly different from '0' if GreatSchools systematically made ratings available for zip codes with lower performing schools earlier (or later) in their history. For instance, $\beta_1 > 0$ would indicate that ratings were systematically available earlier for zip codes with lower performing schools; a $\beta_1 < 0$ would indicate that ratings were available earlier

[5] Available upon request.

for zip codes with higher performing schools. Most importantly, an insignificant coefficient β_1 would indicate that availability is not systematically related to pre-treatment zip code characteristics. In Table 2, we find no systematic evidence that rating availability is related to any of several different socio-economic zip code level characteristics. The only dependent variable for which $YearsAvail_{it}$ has a significant coefficient is % Asian. In robustness tests, we find that including % Asian as a predictor in our regressions does not change the magnitude or significance of our coefficient of interest β_3 in Equation 3, the coefficient for the interaction term of rating availability and lagged school performance.

[Table 2 about here.]

3.3 Variables

Our analysis examines whether the availability of GS ratings for a zip code i at time $t-1$ affected its economic and demographic composition in t. Further, we hypothesize that the effect of online rating availability had a divergent effect, depending on the quality of the schools in that zip-code i. When ratings went online for high-performing schools, home prices increased, communities became more expensive and attracted wealthier, Whiter, more Asian and educated residents. In contrast, when ratings became available for low performing schools, home prices decreased, and wealthier, White, and Asian families left. Below we describe the construction of the independent and dependent variables used to examine these effects.

3.3.1 Independent Variables

To test our hypotheses, we construct two main independent variables, $YearsAvail_{it}$ and SP_{it-1}. We describe the construction and validation of these measures below.

Rating Availability: The primary treatment variable in our analysis is $YearsAvail_{it}$. This variable counts the number of years that at least one school rating has been available for a school in zip code i by year t. For instance, if the first year that GreatSchools has data about a zip code i is $t = 2005$, we code $YearsAvail_{i,2005} = 0$, $YearsAvail_{i,2006} = 1$, $YearsAvail_{i,2007} = 2$, etc. Thus β_2 in Equation 3, the coefficient on $YearsAvail_{it}$ captures the change in SP_{it-1} as a function of one additional year of rating availability.

School Performance: Our measure of school performance SP_{it-1} is at the zip-code i and year $t-1$ levels. We construct our final variable using the mean score on the mathematics examination for all students in a given school s for a given grade g in year t: $SCORE_{sgt}$.[6]

To calculate our measure, we first standardize each score $SCORE_{sgt}$ into a Z-score for each school-grade-year observation: $ZSCORE_{sgt}$. Each school and grade's scores are normalized relative to all others whose students were tested using the *same* standardized examination in that year. Finally, we create an aggregate measure of school performance at the zip code level SP_{it} by calculating the mean of $ZSCORE_{sgt}$ for all schools in zip code i for a given year t.

3.3.2 Dependent Variables

Home Values: Our primary measure of home values is the Zillow Home Value Index ($ZHVI_{it}$). The ZHVI is a seasonally adjusted measure of the average dollar value of a home in a zip code. Since this data is provided on a monthly level for each year, we use data for April as it is the month with the most number of home sales nationally according to Zillow footnoteFound here. However, the correlation between monthly ZHVI indices across all months is $\rho > .99$, suggesting similar scal-

[6]Because different states administer different state-level standardized tests in Math during the period of our data, we have students' scores in Math in 105 different standardized tests. These test scores indicate the percentage of total tested students who meet or exceed the expected performance benchmark for that test as established by the state education department.

ing for all months of the ZHVI data within a year. We use this data to examine the effect of rating availability on changes to home values. In our sample, the average home value is \$217,842.6, with a minimum value of \$13,600 (Earle, AR) and a maximum \$5,442,900 (Atherton, CA).

Percent Top Income: We use the Internal Revenue Service's Individual Income Tax Statistics database to construct a variable calculating the percentage of households with Adjusted Gross Income over \$100,000 in a given zip code–year. We use the Urban Institute definition to define \$100,000 and above as the threshold for the upper-middle class (Rose, 2016). In our sample, the average zip code had 13.62% of households earning more than \$100k per year. There were 3,650 zip codes with '0' households earning more than \$100,000 per year.

Percent White, Black, Asian, and Hispanic: We use the US Census Bureau's American Community Survey (ACS) to construct our demographic variables. From 2010 to 2016, the Bureau publishes estimates for the number of White, Black, Asian and Hispanic residents in a zip code. The average demographic of a zip code was 77.4% White, 7.4% black 8.8% Hispanic and 1.93% Asian. Over this period, the percentage White population of all zip codes declined from 77.37% to 76.9%.

Percent College Educated: We use the ACS to calculate the proportion of college-educated residents in a zip code in a given year. Approximately 30% of residents in an average zip code had an Associates degree or higher.

Migration: Finally, we use the ACS to examine the degree of migration into a zip code for a given year. During our analysis period, we see that the average zip code had 12% in-migration, of which 9.9% were from within the state.

We present summary statistics in Table 3.

[Table 3 about here.]

4 Results

4.1 Model Free Evidence

We begin by examining the mean home values for zip codes with above and below-median quality schools at two points in time: one year before GS rating availability, and three years afterward. This analysis derives data from the 19 states for which we have school performance information pre- and post-GreatSchools Rating availability.[7] One year prior to GS rating availability, the mean value of a home in a zip code with above-median schools was $237,151 with a *95% CI* of [$231,366, $242,936]. The mean value of a home in a zip code with below-median schools was $172,142 with a *95% CI* of [$167,751, $176,533]. This represents a difference in home values between above- and below-median districts of $65,009 with a *95% CI* of [$57,346, $72,672]. That is, homes with above-median schools were valued 38% higher than those with below-median schools.

Three years after GreatSchools ratings were made available, the mean home value for zip codes with above-median schools was $249,641 with a *95% CI* of [$244,942, $254,341] and those with below-median schools was $153,116 with a *95% CI* of [$150,483, $155,750]. This represents a difference in home values between above- and below-median districts of $96,525 with a *95% CI* of [$90934, $102,115]. That is, homes with above-median schools were valued 63% higher than homes with below-median schools.

In summary, home values for zip codes with above-median schools increased by approximately $12,500. In contrast, home values in zip codes with below-median

[7] These states are AZ, DC, FL, IN, MA, MD, MI, MN, NJ, NM, NY, OH, OR, PA, SC, TN, VA, WA, WI

schools decreased by $19,000. As a result, buying a home in a zip code with above-median schools went from being 38% more expensive to 63% more expensive.

These unadjusted statistics provide preliminary evidence of an increasing gap in home values between treatment and control zip codes (e.g., those with and without GS ratings for their schools). Below, we formally test for the effect of rating availability on housing prices, the distribution of household income, demographics, and migration. For each of these dependent variables, we estimate Equation 3 with fixed effects at the zip code and year level separately, the state–year level, and county–year level. In each of our models, we cluster our standard errors at the zip code level.[8]

4.2 Rating Availability, Home Values, and High-Income Households

We begin our analysis by estimating Equation 3 using housing values, $log(ZHVI_{it})$, as our dependent variable.[9] We present these results in Table 4. Model 1 estimates the basic model without any fixed effects. The coefficient of interest is *School Performance* $*$ *Years Available*, which is positive and statistically significant. The coefficient estimate $\beta_3 = 0.014$ suggests that for every one standard deviation increase in school performance, every year of rating availability increases housing prices by 1.15%. Based on this estimate, after three years, a zip code at one standard deviation higher than the zip code of average school performance with ratings available will have homes priced at 3.49% higher than a zip-code with similar schools but no ratings available. Figlio and Rouse (2006) found that the home values increase by 6.7 percent over a three year period in areas of state assigned grade 'A' schools to that of grade 'B' school areas, and no difference in home values for grade 'B' school areas and grade 'C' school areas over three years. Kane, Riegg and Staiger (2006)

[8]We cluster correct standard errors at the zip-code level, to account for correlation in error terms across time at the smallest geographical unit in our data - zip-code level instead of against county or state.

[9]We take the logarithm of ZHVI to account for right skew in our data.

similarly found a 9.8 percent increase in home values with a 1-sd increase in school test scores than average in Mecklenburg County, North Carolina.

[Table 4 about here.]

In Model 2, we include zip code and year fixed effects separately to account for unobserved heterogeneity at the zip code level and yearly trends in housing prices. Again, we find a positive and significant coefficient for β_3. With these fixed effects, our coefficient drops to $\beta_3 = 0.008$, which amounts to 0.631% higher home values for a zip code with 1-sd better school, one year after availability. The next two models include State-Year fixed effects and County-Year fixed effects that account for different trends in housing values across states and counties, respectively. In these models, we formally account for the potential for each county (or state) to have differing trends in their housing prices based on changes to tax policy, crime, or other factors that may have changed during this period. Again, we find consistent coefficient estimates for β_3. In Model 4, which is our most conservative specification, we find $\beta_3 = 0.008$, which amounts to a 0.628% effect.

Regarding dollar values, housing prices for zip codes that are 2-sd apart in school performance diverge by an additional $3,513[10] after one year of rating availability and $8,996 after four years of availability in our most conservative models.

The change in housing prices also signals a potential change in the underlying demographics of zip codes where ratings became available. Next, we estimate Equation 3 using the % of high income households, *% 100k+*, as our dependent variable. These results are presented in Table 5.

[Table 5 about here.]

In Model 1, we present the results without any fixed effects. The coefficient $\beta_3 = .246$ is positive and statistically significant, which indicates that the percentage

[10] β_3=0.008 $\longrightarrow$ log(new/old)ZHVI=2*0..008=0.016. $\longrightarrow$ delta zhvi = exp(0.016)-1 * avg zhvi =0.016*217842 = 3513

of high-income households increases by 0.2% for a zip code with 1-sd better schools than average for each additional year of availability. With zip code and year fixed effects in Model 2, the coefficient is smaller at $\beta_3 = .070$, but still statistically significant. These controls reduce our estimate of the change in the percentage of high-income households to a more conservative 0.06% for a zip code with 1-sd better schools than average. In Models 3 and 4, we include state-year and county-year fixed effects. Our results remain within the general bounds of Models 1 and 2. Models 3 and 4 produce $\beta_3 = .261$ and $\beta_3 = .241$, respectively.

These results suggest a widening gap in the proportion of high-income households in zip codes with low-performing schools and those with high-performing schools. Regarding magnitude, the gap in the percentage of 100k+ income households between zip codes 1-sd above and below average increases by between an additional 0.11% and 0.4% one year after rating availability, and from 0.45% to 1.6% four years after rating availability.

4.2.1 Robustness with Pre- and Post- GreatSchools Data

A concern with our main specification is we do not have pre-GS school quality information for a zip code in the models above. To mitigate this concern, we further collected the test scores from 19 states where school performance data from before the entry of GreatSchools is available.[11]

Unfortunately, states provide this information in various formats, often spread over many websites. Indeed, many states provide this information only at the school district level and not at the school level. After an extensive search process, we collected school-level standardized tests scores for schools in 19 states and territories (e.g., Washington DC) before their availability on the GS website. We summarize this data in Table 6. Overall, we have school performance data for 33,985 schools in 9411 zip codes across 19 states.

[11] The states in our sample are AZ, DC, FL, IN, MA, MD, MI, MN, NJ, NM, NY, OH, OR, PA, SC, TN, VA, WA, and WI.

[Table 6 about here.]

With this data, we estimate the following model:

$$\begin{aligned} Y_{it} = \beta_1 SP_{i(t-1)} + \beta_2 YearsAvail_{it} + \\ \beta_3(SP_{i(t-1)} \times YearsAvail_{it}) + \alpha_{ct} + \epsilon_{it} \end{aligned} \tag{3}$$

In this model, $YearsAvail_{it}$ is the number of years since the introduction of school scores on the GS website. Therefore, the $YearsAvail_{it}$ variable will be equal to 0 for all years up to the year of its introduction on the GS website. Like our previous models, $SP_{i(t-1)}$ is our measure of standardized school performance for a given zip code–year observation. We again estimate this model using the dependent variables: the log of a zip code's Zillow house price index and percentage of households having an annual income greater than 100K in that zip code. β_3 is our coefficient of interest. Table 7 presents these estimations.

[Table 7 about here.]

Even in these models, we find that the coefficient of interest—the interaction between $SP_{i(t-1)}$ and $YearsAvail_{it}$—remains positive and highly significant for both dependent variables. The coefficient estimates are also similar to our findings with the full sample of zip codes in tables 4 and 5. These estimates provide further support for our main finding that the mass availability on online school ratings led to the divergence in house prices and the concentration of high-income households in zip codes with higher performing schools.

[Figure 2 about here.]

In Figure 2 we provide a lead-lag plot of the effect of GreatSchools rating availability on housing prices in nominal dollars. Before the availability of online ratings, zip codes with equivalent quality schools differed little in home values. However, after ratings became available for some zip codes, home prices began to diverge. Figures 3a and 3b more clearly show this dynamic. In this graph, we can see that by year three, the difference in home prices for treatment (GS Available) and control (GS Not available) was 5.8% or approximately $13,885.

[Figure 3 about here.]

4.3 Mechanism and Robustness checks

There is a strong correlation between income levels, race and ethnicity in American society Reardon, Fox and Townsend (2015). In the models presented in Table 8 we estimate the impact of rating availability on the changing composition of communities. Again, we estimate Equation 3 with county-year fixed effects. We cluster our standard errors at the zip code level.

In these models, our dependent variables are the percentage of White, Black, Hispanic and Asian residents in a zip code (Models 1 through 4). Furthermore, in Model 5, we present results for the proportion of individuals with associates degrees or higher. Broadly, we find the demographics of the communities in which ratings became available began to diverge. Qualitatively, we find that zip codes with high performing schools gained White ($\beta = 0.241\%$), Asian ($\beta = 0.165\%$) and college-educated residents ($\beta = 0.427\%$) with the availability of ratings. It appears availability did have a minor effect on the percentage of African American or Black residents, but significantly reduced the percentage of Hispanic residents ($\beta = -0.687\%$).[12]

Like the prior results, these estimates suggest that when ratings become available, the racial and ethnic composition of communities shifts. For example, regarding magnitude, the difference in the percentage of White & Asian residents in

[12]Note that the dependent variables such as the percentage of Whites in a zip code are estimates which have a margin of error clearly given in ACS tables. Such measurement error in our dependent variable in OLS does not bias the coefficient estimates but increases the standard error of estimates. We adjust all standard errors in Table **??** for reported margins of error in the ACS tables.

zip-codes 1-sd above and below average based on school performance increases by between an additional .66% one year after rating availability, and 2.64% four years after rating availability. The gap in college-educated residents increases by a similar magnitude. This divergence is *in addition* to that caused by other factors beyond the impact of rating availability.

[Table 8 about here.]

In our final set of models, we examine the effect of rating availability on migration into and out of the zip code. These models, like those above, are estimated using county-year fixed effects.

In Models 6 and 7, we examine % overall in-migration and % migration from within the state. We find that rating availability significantly affects migration into a zip code. A zip code with above average schools has overall higher in-migration ($\beta =$ 0.063%), which appears to be driven by migrations within a state ($\beta = 0.044\%$). Furthermore, we find that when ratings become available for a zip code with above average schools, fewer individuals in that zip code leave.

To summarize, we find evidence that rating availability accelerated the divergence across American communities. Specifically, the gap between zip codes with high-performing and low performing schools increased on several critical and related dimensions. First, housing prices began to diverge further—with zip codes containing better schools also having higher priced homes. Second, the ethnic composition of such communities also changed: White and Asian families increasingly moved into these communities, and the proportion of Hispanic residents declined. The change was also economic: zip codes with the better performing and more visible schools attracted college-educated residents with higher incomes. All these changes further widened the gap between the zip codes with low- and high-performing schools as identified in prior research.

5 Conclusion

Can greater access to online information help to bridge the rising inequality in American society? Using the gradual availability of online school ratings provided by GreatSchools.org, we ask whether the widespread access to quantified school performance information available today has minimized or accelerated this divergence. Across a range of specifications, we find that access to school performance ratings appeared to accelerate, rather than reduce, economic divergence across zip codes in the United States.

Regarding effect size, we find that after three years, a 1-standard deviation better zip code (measured by school performance) will have homes priced at 3.49% higher than a zip code with equivalent quality schools, but no ratings available. This significant change in housing prices is also related to economic and demographic divergence across zip codes. In most of our specifications, we find that neighborhoods with lower-performing schools lose high-income and college-educated residents as well as White and Asian residents. We find an asymmetric effect for high-performing zip codes. These results broadly support the thesis that widespread access to quantified school performance information accelerated, rather than minimized, social and economic divergence across American communities.

We also acknowledge several limitations of our approach. First, ours is an observational study that uses the time-varying availability of online ratings across communities. As a result, given that rating availability is not random, our estimates may still have some degree of bias. However, we can account for many possible sources of selection bias in our models using various fixed effects specifications. Furthermore, it is possible that our effect sizes have a potentially conservative bias. That is, if rating availability is related to the ease of access to the data for GreatSchools, then it is likely that this school quality information should have already been priced into homes, as school quality information can be accessed from other sources. This problem of ‘low-hanging fruit’ would lead to a conservative bias in our estimates. Nevertheless, we believe that this issue should still temper the interpretation of our results.

Another limitation of our analysis is that we focus on the effect of rating availability on community characteristics, but do not address the equally important question of how rating availability affects school performance. Prior research has shown that such accountability measures often affect school performance, but only through the types of composition effects we identify in this paper (e.g., Figlio and Lucas, 2004). Therefore, more research needs to be conducted on how parents use this information to influence schools and what rating availability means for individual student outcomes.

Finally, we have conducted our analysis at the zip code level. This approach allowed us to analyze the effect of rating availability on many outcomes at that level of analysis. However, this approach also introduces noise in our estimates because zip codes often, but not always, define the geographic units delineating school boundaries. Moreover, analyzing outcomes at such an aggregate level limits our ability to identify the effect of rating availability on the choices of individual households, and thus our ability to more neatly understand mechanisms.

We hope these results encourage new research on how large-scale access to information and resources through the internet are affecting critical social dynamics (Hargittai and Hinnant, 2008; DiMaggio et al., 2004; Cotten, Anderson and Tufekci, 2009). Research exploring the value of online informational interventions—and how to most effectively design them—has potential to inform policy and practice, especially as more individuals are using the internet to make important decisions about their economic and social well-being.

References

Brasington, David and Donald R Haurin. 2006. "Educational outcomes and house values: a test of the value added approach." *Journal of Regional Science* 46(2):245–268.

Brynjolfsson, Erik and Andrew McAfee. 2014. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Chevalier, Judith A and Dina Mayzlin. 2006. "The effect of word of mouth on sales: Online book reviews." *Journal of Marketing Research* 43(3):345–354.

Coleman, James S et al. 1966. "Equality of educational opportunity.".

Cotten, Shelia R, William A Anderson and Zeynep Tufekci. 2009. "Old wine in a new technology, or a different type of digital divide?" *New Media & Society* 11(7):1163–1186.

Damianov, Damian S and Diego Escobari. 2016. "Long-run equilibrium shift and short-run dynamics of US home price tiers during the housing bubble." *The Journal of Real Estate Finance and Economics* 53(1):1–28.

DiMaggio, Paul and Bart Bonikowski. 2008. "Make money surfing the web? The impact of Internet use on the earnings of US workers." *American Sociological Review* 73(2):227–250.

DiMaggio, Paul, Eszter Hargittai, Coral Celeste and Steven Shafer. 2004. "From unequal access to differentiated use: A literature review and agenda for research on digital inequality." *Social Inequality* pp. 355–400.

Dorsey, Robert E, Haixin Hu, Walter J Mayer and Hui-chen Wang. 2010. "Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle." *Journal of Housing Economics* 19(2):75–93.

Espeland, Wendy Nelson and Michael Sauder. 2007. "Rankings and reactivity: How public measures recreate social worlds." *American Journal of Sociology* 113(1):1–40.

Espeland, Wendy Nelson and Mitchell L Stevens. 1998. "Commensuration as a social process." *Annual Review of Sociology* 24(1):313–343.

Fack, Gabrielle and Julien Grenet. 2010. "When do better schools raise housing prices? Evidence from Paris public and private schools." *Journal of public Economics* 94(1-2):59–77.

Figlio, David N and Cecilia Elena Rouse. 2006. "Do accountability and voucher threats improve low-performing schools?" *Journal of Public Economics* 90(1-2):239–255.

Figlio, David N and Maurice E Lucas. 2004. "What's in a grade? School report cards and the housing market." *American Economic Review* 94(3):591–604.

Gibbons, Stephen, Stephen Machin and Olmo Silva. 2013. "Valuing school quality using boundary discontinuities." *Journal of Urban Economics* 75:15–28.

Guerrieri, Veronica, Daniel Hartley and Erik Hurst. 2013. "Endogenous gentrification and housing price dynamics." *Journal of Public Economics* 100:45–60.

Hargittai, Eszter and Amanda Hinnant. 2008. "Digital inequality: Differences in young adults' use of the Internet." *Communication Research* 35(5):602–621.

Harris, Douglas N and Carolyn D Herrington. 2006. "Accountability, standards, and the growing achievement gap: Lessons from the past half-century." *American Journal of Education* 112(2):209–238.

Hoxby, Caroline and Sarah Turner. 2013. "Expanding college opportunities for high-achieving, low income students." *Stanford Institute for Economic Policy Research Discussion Paper* (12-014).

Jargowsky, PA. 1996. "Take the money and run: economic segregation in US metropolitan areas." *American Sociological Review* 61(6):984–98.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics* 125(2):515–548.

Jorgenson, Dale W. 2001. "Information technology and the US economy." *American Economic Review* 91(1):1–32.

Kane, Thomas J, Stephanie K Riegg and Douglas O Staiger. 2006. "School quality, neighborhoods, and housing prices." *American Law and Economics Review* 8(2):183–212.

Logan, John R. 2011. "Separate and unequal: The neighborhood gap for blacks, Hispanics and Asians in Metropolitan America." *Project US2010 Report* pp. 1–22.

Luca, Michael and Jonathan Smith. 2013. "Salience in quality disclosure: evidence from the US News college rankings." *Journal of Economics & Management Strategy* 22(1):58–77.

Mayer, Susan E. 2002. "How economic segregation affects children's educational attainment." *Social Forces* 81(1):153–176.

Mikulecky, Marga and Kathy Christie. 2014. "Rating States, Grading Schools: What Parents and Experts Say States Should Consider to Make School Accountability Systems Meaningful." *Education Commission of the States* .

Nguyen-Hoang, Phuong and John Yinger. 2011. "The capitalization of school quality into house values: A review." *Journal of Housing Economics* 20(1):30–48.

Owens, Ann, Sean F Reardon and Christopher Jencks. 2016. "Income segregation between schools and school districts." *American Educational Research Journal* 53(4):1159–1197.

Piketty, Thomas and Emmanuel Saez. 2003. "Income inequality in the United States, 1913–1998." *The Quarterly Journal of Economics* 118(1):1–41.

Piketty, Thomas and Emmanuel Saez. 2006. "The evolution of top incomes: a historical and international perspective." *American Economic Review* 96(2):200–205.

Pope, Devin G. 2009. "Reacting to rankings: evidence from "America's Best Hospitals"." *Journal of Health Economics* 28(6):1154–1165.

Quillian, Lincoln. 2014. "Does segregation create winners and losers? Residential segregation and inequality in educational attainment." *Social Problems* 61(3):402–426.

Reardon, Sean F. 2011. "The widening academic achievement gap between the rich and the poor: New evidence and possible explanations." *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances* pp. 91–116.

Reardon, Sean F and Kendra Bischoff. 2011. "Income inequality and income segregation." *American Journal of Sociology* 116(4):1092–1153.

Reardon, Sean F, Lindsay Fox and Joseph Townsend. 2015. "Neighborhood income composition by household race and income, 1990–2009." *The Annals of the American Academy of Political and Social Science* 660(1):78–97.

Reeves, Richard V. 2017. *Dream hoarders: How the American upper middle class is leaving everyone else in the dust, why that is a problem, and what to do about it.* Brookings Institution Press.

Rose, Stephen. 2016. "The growing size and incomes of the upper middle class." *Urban Institute* .

Rumberger, Russell W and Gregory J Palardy. 2005. "Does segregation still matter? The impact of student composition on academic achievement in high school." *Teachers College Record* 107(9):1999.

Salganik, Matthew J and Duncan J Watts. 2008. "Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market." *Social Psychology Quarterly* 71(4):338–355.

Salganik, Matthew J, Peter Sheridan Dodds and Duncan J Watts. 2006. "Experimental study of inequality and unpredictability in an artificial cultural market." *Science* 311(5762):854–856.

Sampson, Robert J, Patrick Sharkey and Stephen W Raudenbush. 2008. "Durable effects of concentrated disadvantage on verbal ability among African-American children." *Proceedings of the National Academy of Sciences* 105(3):845–852.

Santos, Rita, Hugh Gravelle and Carol Propper. 2017. "Does quality affect patientsâĂŹ choice of doctor? Evidence from England." *The Economic Journal* 127(600):445–494.

Sauder, Michael and Wendy Nelson Espeland. 2009. "The discipline of rankings: Tight coupling and organizational change." *American Sociological Review* 74(1):63–82.

Shore, Cris, Susan Wright, Vered Amit, Judy Brown, Casper Bruun Jensen, Mark Maguire, Sally Engle Merry, Michael Sauder, Wendy Espeland, Gavin Smith et al. 2015. "Audit culture revisited: Rankings, ratings, and the reassembling of society." *Current Anthropology* 56(3):431–432.

Varkevisser, Marco, Stephanie A van der Geest and Frederik T Schut. 2012. "Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands." *Journal of Health Economics* 31(2):371–378.

Ziol-Guest, Kathleen M and Kenneth TH Lee. 2016. "Parent Income–Based Gaps in Schooling: Cross-Cohort Trends in the NLSYs and the PSID." *AERA Open* 2(2):2332858416645834.

Figure 1: Example of GreatSchools ratings on main website (above) and through the Zillow website (below).

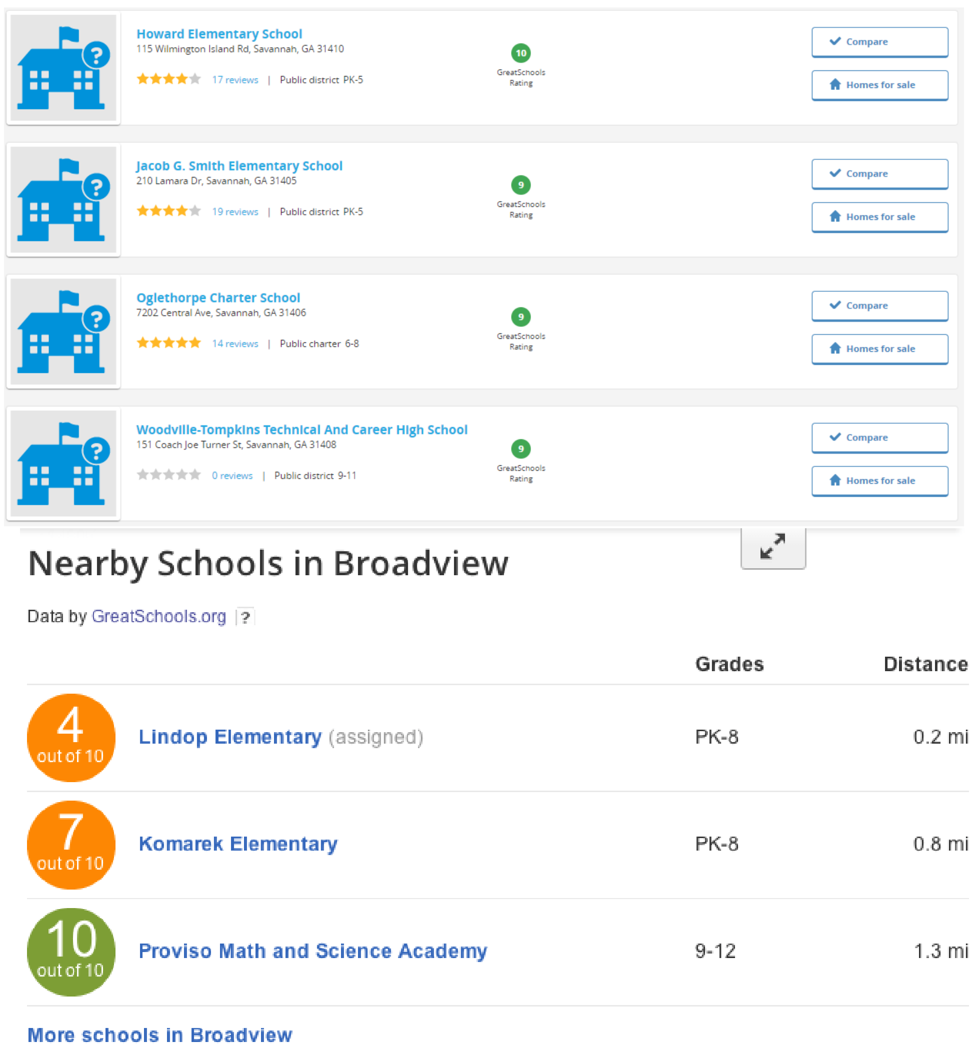
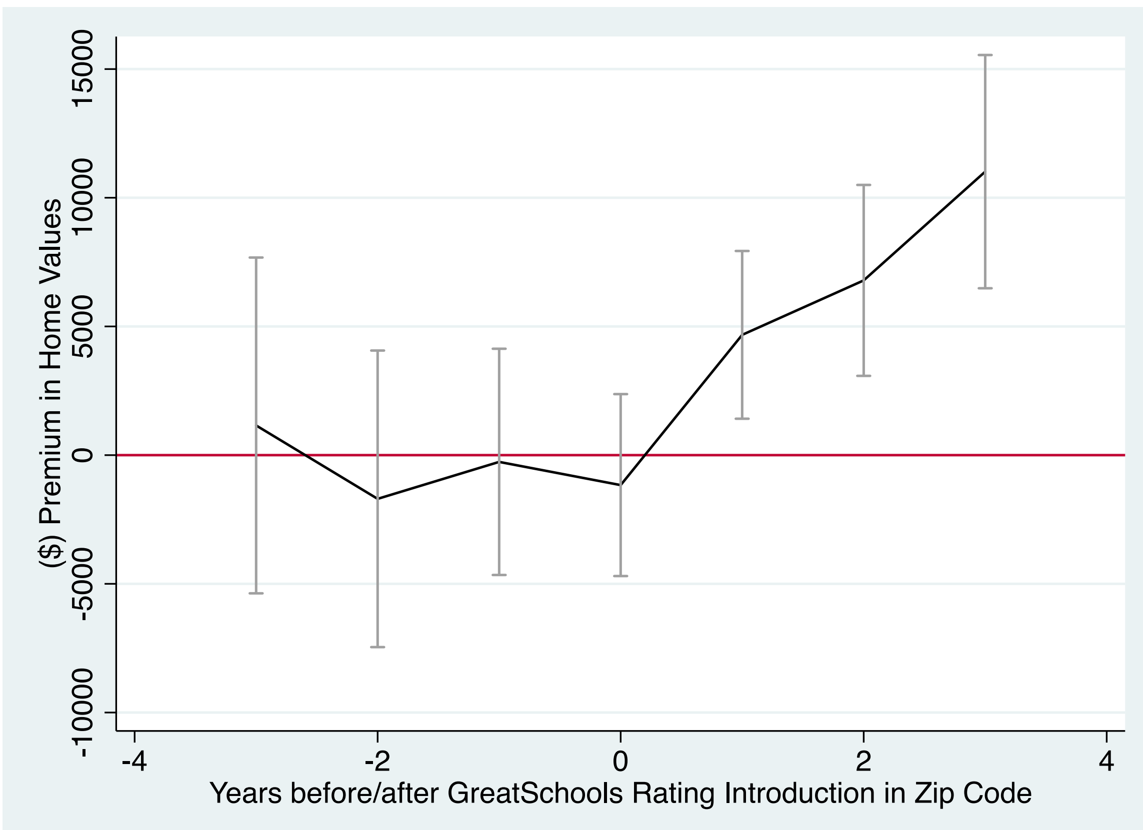
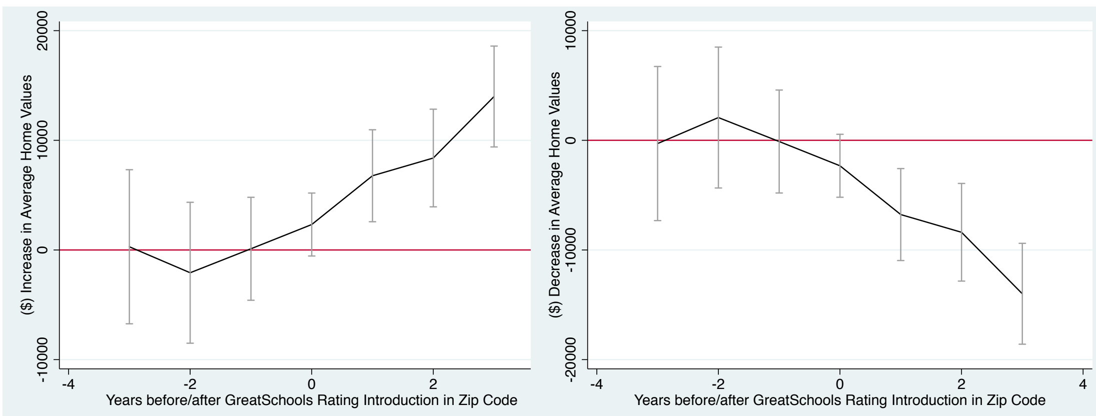


Figure 2: Plot of the difference in housing values between zip codes whose schools do/do not have GreatSchools ratings. The plot derives from models with county-year fixed effects and 19 states for which we have school performance information pre- and post- GreatSchools Rating availability (AZ, DC, FL, IN, MA, MD, MI, MN, NJ, NM, NY, OH, OR, PA, SC, TN, VA, WA, WI).





(a) Effect of GreatSchool Rating availability for Above Median Schools

(b) Effect of GreatSchool Rating availability for Below Median Schools

Figure 3: Plot of the difference in housing values for zip codes with and without GreatSchools ratings for zip codes above- and below-median test scores.

Year	States	zip-codes	Schools
2006	5	4643	20298
2007	5	4658	20637
2008	8	5926	24898
2009	10	6543	26330
2010	16	8299	32052
2011	26	10856	42581
2012	50	20551	74087

Table 1: Coverage of Greatschools.org data from 2006 to 2015.

Table 2: Tests examining the relationship between rating exposure $YearsAvail_{it}$ and zip-code level characteristics.

	SP_{it-1}	Housing	% 100k+	% White	% Black	% Hisp	% Asian	% College
Years Available	0.008	-0.002	0.266	-1.099	0.523	0.672	0.387^{**}	0.620
	(0.025)	(0.018)	(0.239)	(0.796)	(0.535)	(0.496)	(0.137)	(0.407)
Constant	0.076	12.180^{***}	18.812^{***}	74.688^{***}	8.602^{***}	10.584^{***}	2.425^{***}	30.724^{***}
	(0.053)	(0.135)	(1.072)	(0.182)	(0.099)	(0.136)	(0.061)	(0.485)
Observations	126117	92975	119777	124427	124427	124427	124427	116112
R^2	0.010	0.427	0.175	0.321	0.254	0.394	0.260	0.162

Standard errors in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 3: Summary statistics for main variables used in our analysis.

	count	mean	sd	min	max
School Performance	126282	0.025	0.812	-7.335	4.610
Years Available	147787	3.376	2.691	0.000	10.000
Log(ZHVI+1)	165549	12.048	0.658	9.518	15.510
$ ZHVI	165549	217842.590	198219.159	13600.000	5442900.000
% 100k+ Households	279720	13.621	10.358	0.000	84.507
% White	231840	77.370	26.350	0.000	100.000
% Black	231840	7.402	15.748	0.000	100.000
% Hispanic	231840	8.800	16.377	0.000	100.000
% Asian	231840	1.937	5.161	0.000	100.000
% College Degree	195734	30.468	16.805	0.000	100.000
% Migriation	195961	12.077	9.558	0.000	100.000
% In-State Migration	195961	9.916	7.977	0.000	100.000
% No Migration	195961	87.555	9.880	0.000	100.000
Observations	349041				

Table 4: The effect of school rating availability on the relationship between school quality and housing prices.

	Log(ZHVI+1)	Log(ZHVI+1)	Log(ZHVI+1)	Log(ZHVI+1)
School Performance	0.293***	-0.030***	0.292***	0.262***
	(0.009)	(0.003)	(0.016)	(0.007)
Years Available	0.022***	-0.009***	-0.004	-0.023***
	(0.001)	(0.000)	(0.007)	(0.005)
School Performance * Years Available	0.014***	0.008***	0.014***	0.008***
	(0.002)	(0.000)	(0.004)	(0.002)
Constant	11.912***	12.188***	12.018***	12.099***
	(0.006)	(0.003)	(0.030)	(0.021)
Zip FE	No	Yes	No	No
Year FE	No	Yes	No	No
State-Year FE	No	No	Yes	No
County-Year FE	No	No	No	Yes
Observations	79718	79718	79718	79718
R^2	0.156	0.986	0.572	0.838

Standard errors in parentheses
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 5: The effect of school rating availability on the relationship between school quality and % of 100k+ income households in zip-code.

	% 100k+	% 100k+	% 100k+	% 100k+
School Performance	5.477***	-0.128***	5.277***	5.425***
	(0.113)	(0.034)	(0.478)	(0.227)
Years Available	0.117***	-0.340***	0.133	-0.219*
	(0.017)	(0.006)	(0.127)	(0.105)
School Performance * Years Available	0.246***	0.070***	0.261*	0.214***
	(0.029)	(0.006)	(0.117)	(0.058)
Constant	13.407***	17.246***	13.358***	14.608***
	(0.068)	(0.033)	(0.451)	(0.374)
Zip FE	No	Yes	No	No
Year FE	No	Yes	No	No
State-Year FE	No	No	Yes	No
County-Year FE	No	No	No	Yes
Observations	99339	99339	99339	99339
R^2	0.241	0.986	0.393	0.619

Standard errors in parentheses
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 6: Data description for State Department of Education Data and GreatSchools introduction.

State	DOE Years	Years	GS First Year	Schools	Zips
Arizona	2008-11	4	2012	1825	334
District of Columbia	2007-10	4	2011	116	19
Florida	2005-11	7	2012	2744	789
Indiana	2006-11	6	2012	1474	502
Massachusetts	2008-11	4	2012	1532	411
Maryland	2006-11	6	2012	1069	267
Michigan	2005-07	3	2008	2262	656
Minnesota	2006-09	4	2010	1593	535
New Jersey	2003-11	9	2012	1834	511
New Mexico	2005-11	7	2012	748	182
New York	2007-10	4	2011	3284	1032
Ohio	2006-11	6	2012	3371	828
Oregon	2004-11	8	2012	1228	308
Pennsylvania	2006-11	6	2012	2878	891
South Carolina	2009-11	3	2012	839	284
Tennessee	2010-11	2	2012	1138	415
Utah	2008-09	2	2010	890	
Virginia	2007-11	5	2012	1783	502
Washington	2009-12	4	2013	1515	409
Wisconsin	2012	1	2012	1862	536
	Total	95		33985	9411

Table 7: The effect of school rating availability on the relationship between school quality and home values as well as the % of 100k+ income households in zip-code. Estimates use data from 19 states for which we have school performance information pre- and post- GreatSchools Rating availability. All models include county-year fixed effects standard errors cluster corrected at the zip code level.

	Log(ZHVI+1)	%100k+
School Performance	0.278^{***}	6.871^{***}
	(0.007)	(0.144)
School Performance * Years Available	0.015^{***}	0.132^{**}
	(0.003)	(0.044)
Observations	55357	70292
R^2	0.832	0.688

Standard errors in parentheses
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 8: The effect of school rating availability on the relationship between school quality and zip-code demographics and migration patterns.

	% White	% Black	% Hisp	% Asian	% College	% In Migr.	% State Migr.
School Performance	9.034***	-6.490***	-1.204***	0.112	5.793***	-1.312***	-1.357***
	(0.306)	(0.254)	(0.177)	(0.066)	(0.220)	(0.112)	(0.094)
Years Available	-0.931**	0.164	0.660*	0.147	0.039	0.172	0.278*
	(0.334)	(0.239)	(0.259)	(0.081)	(0.218)	(0.133)	(0.110)
School Performance * Years Available	0.241***	0.116*	-0.687***	0.165***	0.427***	0.063**	0.044*
	(0.064)	(0.050)	(0.049)	(0.020)	(0.045)	(0.023)	(0.020)
Constant	76.939***	8.187***	8.941***	2.110***	32.329***	11.954***	9.486***
	(1.417)	(1.014)	(1.093)	(0.342)	(0.938)	(0.575)	(0.472)
County-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109685	109685	109685	109685	103143	103150	103150
R^2	0.737	0.637	0.718	0.561	0.611	0.314	0.322

Standard errors in parentheses
* $p < .05$, ** $p < .01$, *** $p < .001$