Household Sorting, Local Wealth, and the Market for Homes Near Diverse Schools

Kenneth Whaley
Clemson University*

November 20, 2019

Job Market Paper

Abstract

I analyze mortgage loans and home values from 2000-2014 to study housing markets and household sorting in neighborhoods near racially diverse schools. Using a national sample of over 3,600 middle schools, I construct a measure of school demographic diversity and estimate how home values and mortgage loan amounts change as diversity increases. For identification I isolate variation in school demographics associated with the quasi-random timing of rental housing development under the Low-Income Housing Tax Credit (LIHTC) program. I find that mortgage values rise by nearly 6.7%, holding median home values constant and controlling for changes in local income levels and home buyer socioeconomic characteristics. The magnitude of the effect is consistent for white, black, and Hispanic home buyers, and coincides with a decrease in home values of 2.5% near diversifying schools. The effect is reversed for white home buyers near diverse schools in low-income areas, who borrow less for housing holding prices constant. I present two explanations for these findings, both which shed light on neighborhood wealth, down payment ability, and the consequences of household sorting over local amenities.

JEL Classification: R31, H4, R58

Keywords: Racial Diversity, Neighborhood Choice, Housing Demand

---

*Email: kdwhale@g.clemson.edu.

I have received helpful comments from F. Andrew Hanssen, Peter Blair, Robert Fleck, Christy Zhou, William Dougan, Mike Makowsky, Jhacova Williams, Jorge Garcia, Donna Feir, Abigail Wozniak, Jim Schmidt, Elora Raymond, Yuxi Yao, Matthew Harvey, Xi Bai, David Agrawal, and the Clemson Public Economics Workshop. All errors are my own.

1 Introduction

Classical theories of neighborhood choice beginning with Schelling (1971) argue that observed patterns of household sorting by race reflect a range of preferences for the racial demographics of neighbors and other neighborhood attributes. Empirical analogues of the Schelling model have shown preferences similar to those for neighborhood racial composition affect the decision to enroll children in a local public school (Caetano and Maheshri 2017). In this paper I study how household preferences for school characteristics affect the market for homes near schools that become more racially diverse.

Schools and other neighborhood features that affect the demand for housing in an area are commonly referred to as local amenities (Baum-Snow and Ferreira 2015). Changes to amenities affect home values as households trade-off neighborhood attributes and physical housing characteristics when when choosing where to live. Down payments and closing costs, however, are a fraction of the home value and impose a threshold of wealth necessary to purchase a home. Thus a home buyer may have preferences for certain physical housing and neighborhood features, but market demand and household wealth both play a role in shaping the location choice of home buyers. Given that the remaining balance owed on a home (net of up-front payments) is typically paid over the course of a 30-year home mortgage, the loan amount relative to home value can be informative about the financial traits of a home buyer. I estimate the relationship between these financial measures and exogenous changes to school racial diversity to understand how preferences over public school characteristics affect sorting and local economic outcomes.

I employ a national panel of over 3,600 schools from 2000-2014, using middle school attendance zones as geographic boundaries when calculating my two housing finance measures. I measure school diversity by the extent to which the school student body is comprised of one race or spread equally across five observable groups.[1] The initial distribution of the index

[1] I compute a scaled version of the Herfindahl-Hirschman Index that measures market concentration. My measure can take on values between zero and ten, with ten being a school with equal shares of all race types. Further discussion of the school and neighborhood diversity index is in section 2

in 2000 is bimodal, with a mass of low diversity schools and another mass comprised of high diversity schools. I observe schools on average become more racially diverse from 2000-2014. Estimating the causal effect of school diversity on housing markets requires disentangling the relationship between racial demographics and neighborhood attributes unobserved in the data but correlated with home values. In their seminal paper, Bayer et al. (2007) show that cross-sectional differences in home values that appear associated with race are largely a product of the relative quality of other amenities offered in minority versus white neighborhoods. My empirical design is a two-stage least squares model that exploits the panel nature of the data to relate plausibly exogenous changes in school diversity to changes in mortgage values and local home prices.

I first classify each school as low or high diversity by estimating a finite mixture model, commonly used in urban economics to identify housing submarkets based on observable neighborhood characteristics (Ugarte et al. 2004, Belasco et al. 2012). My identifying assumption is that the bimodal distribution of schools by diversity in the year 2000 reflects two latent school types. I then exploit differences in the timing of rental housing development under the Low Income Housing Tax Credit (LIHTC) program using an event-study model to identify quasi-random shifts in student enrollment counts by race and ethnic group. I plot the event-time coefficients as visual evidence that LIHTC-induced changes to white and Hispanic enrollment vary by neighborhood diversity type. I find little evidence of changes in black and Asian enrollment relative to the year LIHTC units become available. The model-predicted enrollment counts are used to calculate exogenous growth rates of student counts for each race and ethnic group.

To construct a shift-share instrument for my preferred two-stage least squares model, the predicted enrollment growth rates from the LIHTC event-study are applied to the initial school enrollment shares for each race and ethnic group. I use the resulting exogenous enrollment shares to recalculate the diversity measure for each school by year observation. The reconstructed diversity levels are used to instrument for changes in the observed levels of school diversity. This approach borrows from Boustan (2010) and Derenoncourt (2018) who study various outcomes of demographic shifts during the Great Northern Migration of black households from 1940-1970. In my main model I estimate the effect of school diversity on mortgage values, holding home values constant and conditioning on school zone and year fixed effects.

I find that the median value of newly originated mortgage loans within an attendance zone rises as schools become more diverse, holding median home values constant and controlling for other school and borrower characteristics that could also explain loan amounts. My results show that a one standard deviation diversity increase raises mortgage values between 4.8% and 6.7%.[2] This main result is consistent for white, black and Hispanic home buyers. Alternatively, I find that median mortgage values for white home buyers decrease near diversifying schools in low-income areas, holding home values and all other observable factors constant.

Controlling for changes in home values sharpens my identification and allows me to frame my results in terms of the loan-to-value ratio (LTV), a widely used but imperfect measure of the financial risk associated with a home purchase (Bian et al. 2018). I compute an aggregate index of mortgage loan values for home purchases within a school attendance zone from 2000-2014, using transactions data for provided under the Home Mortgage Disclosure Act (HMDA). My second financial measure is a constant-quality house price index provided by the Federal Housing Finance Agency, and aggregated to the school zone level. The two measures approximate school zone level variation in mortgage loan amounts and home values relative to the base year of 2000. When paired with school zone fixed effects in my main regressions, my estimates capture the covariation between school diversity and mortgage values within a school zone over time, conditional on home value appreciation used as a control.

Holding home values constant is necessary to control for potential spillovers from LI-

[2] Recall that the diversity measure takes values from 0-10. A one standard deviation increase in diversity is about 2.4, compared to a mean diversity level of 4.64 in 2000 and 5.96 in 2014.

HTC affecting mortgage values through omitted channels correlated with changes in school diversity. In an exhaustive analysis of the neighborhood effects of LIHTC, Diamond and McQuade (2019) show that house prices are predicted to rise in low-income areas following LIHTC as a signal neighborhood quality improvements. Alternatively, prices will fall in high-income areas once LIHTC development is introduced. Studies have also shown that LIHTC development is associated with decreases in violent crimes in low-income areas (Freedman and Owens 2011), and has the largest effect on property values in gentrifying areas (Baum-Snow and Marion 2009). For completeness I test for the effect of school diversity on my home value measure directly. The results of this exercise suggest that home values fall when LIHTC generates changes to school diversity.

I consider two potential explanations for rising mortgage values occurring simultaneously with falling home values. It is possible that as schools become more diverse, new home buyers select a higher quantity of housing as compared to the quantity of housing measured by the house price index, which approximates price changes of physically comparable homes over time. This implies that rising mortgage values are associated with higher wealth households moving into neighborhoods near diversifying schools and purchases larger/higher quality homes. An alternative explanation is decreased demand for housing near diversifying schools reduces home values as some households with preferences against diverse schools exit. Boustan (2010) argue that this behavior contributed to increases in black home ownership from 1950-1980, when home values in the central city decreased as white households moved to suburban areas. In my case mortgage values are rising as lower wealth households sort into neighborhoods near diverse schools and buy homes at lower values, but borrow more in terms of real mortgage amounts as wealth constrains down payment ability.

The central contribution of this paper is to advance the study of consequences associated with household sorting. Neighborhood schools are a unique amenity as households may sort over student demographic characteristics directly, or use those demographics as a weak signal of school quality. I create an environment where school demographic changes arise from a housing shock where the timing is independent of initial levels of school quality. The subsequent sorting behavior is important to understand if changes to neighborhood wealth emerge as a result. These results have implications that link housing and education policy given the connection between neighborhood wealth and the provision of public goods and quality schools.

2 Race and The Economics of Neighborhood Choice

Economic models of neighborhood choice feature households with preferences over characteristics of the individual housing unit (square footage, age, number of bedrooms and bathrooms) and attributes of the neighborhood (density, racial composition and average education attainment of neighbors). Equilibrium patterns of segregation over race, income or other observable attributes result from the optimization calculus of all households in a particular market (Rothenberg et al. 1991). The complex behavior that emerges in this theoretical setting is generally referred to in the literature as sorting, a powerful market outcome that suggests the equilibrium composition of a neighborhood can appear segregated, mixed, or in transition between the two. Sorting and the stable equilibrium outcome in neighborhoods is the fundamental idea of Schelling (1971) and more recent studies that have formalized the equilibrium sorting result.

Endogenously formed school demographic composition becomes an amenity that affects willingness to pay for housing in the area. School demographics are a margin of interest for home buyers for multiple reasons. If the household has school aged children, it follows that parents care about the formation of peer networks that affect educational outcomes. It matters less if peers materially affect student outcomes and more that parents have preferences for a certain type of peer. Caetano and Maheshri (2017) use data for Los Angeles County schools and show that once the share of minority students reach a certain level, the share of white students decreases at an accelerating rate until the school population reaches a high minority equilibrium. Falling demand for local public school services could coincide with decreased demand for housing in the immediate area.

Websites designed to provide home buyers with neighborhood characteristics provide detailed information about the school composition along and link the user to sites specifically describing school demographics.[3] The proliferation of home search sites reduces the cost of information used by households to form an opinion about the underlying racial composition of a neighborhood when choosing where to live. If homeowners associate property tax payments with the funding of local public schools, then preferences over school demographics will be reflected in the demand for housing in a neighborhood regardless of the presence of school aged children in the home.

I hypothesize that preferences over school demographics affect the demand price for housing as a byproduct of sorting. Cutler et al. (1999) assert that white households will pay more to live near other white households in the absence of laws upholding segregation. Further, the correlation between race and household wealth imply that sorting over racial attributes can lead to differences in the spatial distribution of wealth across a broader housing market. Therefore changes in the loan-to-value ratio for home purchases in a school attendance zone can reflect home buyer wealth changes by capturing variation in the down payment made for the average home purchase. If mortgage values decrease relative to prices, I infer that larger down payments are being made as the home buyers sorting into the market are wealthier. If mortgage values rise relative to prices, it is because home buyers with lower wealth sort into the neighborhood and borrow more for housing after making smaller down payments.

I test my hypothesis empirically by first constructing a measure of school and neighborhood diversity in section 2.1. My estimation strategy uses subsidized development under LIHTC to create exogenous variation in the demographics of schools, as described in section 5. Diamond and McQuade (2019) argue that LIHTC diversifies neighborhoods along various margins that increases home prices in low income areas and decreases prices in high

[3]Zillow, Redfin, and homes.com all provide both a snapshot of school demographics and a link to greatschools.com

income areas. In my main model I hold price changes constant and estimate the relationship between LIHTC induced demographic changes and mortgage values, taking the timing of the housing development as random. I observe demographic changes in a single measure of school diversity, constructed from the enrollment shares of different race and ethnic groups in a school.

2.1 A Measure of School Racial Diversity

My goal is to measure the extent to which the students in a neighborhood public school are concentrated amongst one race group or diversely distributed across several. My measure is a linear transformation of the Herfindahl-Hirschman Index (HHI), which is typically used to analyze firm market concentration.[4] Using a measure of concentration instead of the fraction of black or Hispanic residents expands the potential for inference within the empirical framework of sections 4 and 5. The school diversity index, SDI_{it} measures the concentration of the student population in neighborhood i across five subgroups $r \in \{$white, Hispanic, black, Asian, and 'other'$\}$.

Denote s_{rit} as the percentage of group r in school i year t, such that $s_{rit} \in [0, 100]$. The standard measure of HHI is the sum of the squared race shares, or

$$HHI_{it} = \sum_{r=1}^{5} s_{rit}^2. \tag{1}$$

With five groups, $max[HHI] = 10,000$ if the student population is from one group, and $min[HHI] = 2,000$ when each group comprises an equal 20% share. For interpretation I compute $HHI_{it}^{'}$ to assign a lower value to low diversity schools,

$$HHI_{it}^{'} = max[HHI] - HHI_{it}. \tag{2}$$

[4]The HHI gained popularity as a way to study the effects of mergers on the distribution of total market share across firms (Rhoades 1993), with higher values representing higher market power for a single firm. For interpretation I transform the measure so that neighborhoods with higher levels of diversity have higher index values.

Now, $max[HHI^{'}] = 8,000$ if the school student population is equally distributed across groups and 0 if all students are of one race. For ease of interpretation I scale my final measure between 0 and 10, such that

$$SDI_{it} = \frac{HHI_{it}^{'}}{max[HHI']} \times 10. \tag{3}$$

When $SDI_{it} = 0$ there is zero diversity (all students are from one r group) and $SDI_{it} = 10$ represents equal distribution across all five r groups. In context of this paper SDI_{it} is representative of both the students in the school and the residents of the neighborhood. The initial distribution of SDI_{it} is bimodal and supporting the notion that equilibrium sorting outcomes persist in US schools and neighborhoods. I present estimates for the distribution of SDI_{it} and discuss my specification that relates mortgage values and home prices to SDI_{it} in section 5.

2.2 Traditional Measures of Neighborhood Demographics

Diversity differs in both construction and interpretation from other demographic measures used in the analysis of neighborhood demographics, such as the dissimilarity and isolation indices. These measures are used to study segregation and racial disparities in economic outcomes for blacks and whites. My diversity measure is more flexible in the sense that it accommodates more than two categories of race and ethnic type. The trade off is that neighborhoods generally classified as low diversity can be majority white, black, Hispanic or Asian. In this section I discuss why my measure is suitable for the analysis of this paper and how it differs from other measures used in the neighborhood choice literature.

Traditional studies of neighborhood segregation within cities are concerned with the distribution and clustering of residents by race across a city or metropolitan areas. The dissimilarity index measures how evenly distributed the total population of black and white residents are spread across different neighborhoods in a larger geography. The isolation index measures exposure, the extent to which black residents come in contact with white residents of a city. The two traditional measures have been used to study the effects of segregation on income inequality (Cutler et al. 1999), and educational outcomes (Reardon 2016).

Logan and Parman (2017) provide a detailed discussion of the shortcomings of the two traditional measures. By construction, the two measures are concerned with how the population of black and white residents in a large geographical area are distributed across smaller neighborhood units. Thus the dissimilarity and isolation indices are attractive for studies of aggregate city level economic outcomes between two groups. The central question of this paper concerns the effects of within neighborhood demographic changes and the effect of home values, and less about how aggregate home values in a city are related to segregation across neighborhoods. Thus the two traditional measures are not appropriate for my analysis.

Traditional measures are also limited to the comparison of two groups, most often to understand differences between black and white resident outcomes. The composite diversity index SDI_{it} is flexible enough to describe population changes of several race and ethnic groups. This allows for a broader analysis of neighborhoods and combines several underlying forces that produce sorting in a modern context. Couture and Handbury (2017) describe the urban revitalization as a process attracting white working professionals to urban areas from 2000-2010, and Diamond (2016) analyzes neighborhood level spillovers from this influx of college-educated workers. Epstein (2008) study the effect of existing social networks on the location choice of Hispanic immigrants. SDI_{it} allows me to understand how the complex residential decisions of white, black, Hispanic, and Asian households as a whole affect home values.

2.3 Neighborhood Data

I observe racial demographics, and thus SDI_{it}, for a balanced panel of 3,661 schools from 2000-2014. This allows for rich analysis of neighborhood demographic changes over time and without the gaps between periods encountered when using census data. The population shares for race group r in the calculation of SDI_{it} are the fraction of total school enrollment belonging to each group. Summary statistics for school demographic variables are shown in table 1, column 1 for the full sample. The average school in my sample across all years is 15% black, 24% Hispanic, 6% Asian. Aud (2011) estimates the K-12 population for the US to be 17% black, 21% Hispanic, and 5% Asian using NCES data from the 2007-2008 school year. At the mean, my sample is roughly representative of the average US public school.

One of the two ways I measure neighborhood income in the data is the fraction of total enrollment receiving free lunch subsidies. To qualify for free lunch subsidies household income must be less than or equal to 130% of US poverty level, and Aud (2011) estimate 48% of US students in 2008 received the subsidy. In my full sample, 40% of the students live in households with income levels low enough to qualify for fully subsidized lunch. I include this variable in my analysis to control for changes in the lower tail of the income distribution within a neighborhood. I describe my other income measure, the average income of new home buyers, in section 3.1.

The school demographic data are publicly available through NCES and include schools with 6 grade students from the lower 48 states. Middle school attendance zones make up the large share of these schools, and figure 1 shows the national distribution of middle school zones in my data. Defining the neighborhood as a middle school attendance zone is attractive primarily because the geography is on average larger than elementary school zones but smaller than high school zones. Smaller boundaries are conducive to neighborhood fixed effects but reduce the power of my empirical model by decreasing the number of home mortgage transactions mapped to a neighborhood in a given year. Larger boundaries yield richer within neighborhood variation over time but increase the potential for measurement error when aggregating mortgages over an increased spatial area. As neighborhood radius increases, the immediate areas around homes in different parts of the neighborhood have less in common.

2.4 The Distribution of Schools by Diversity Level

I estimate the underlying distribution of SDI_{it} in figure 2a and 2b for the years 2000 and 2014, respectively. The two kernel density plots show that schools follow a bimodal distribution that supports long run equilibrium theories of racial sorting discussed in the beginning of section 2. Given the empirical distribution of school demographics I categorize neighborhoods into two groups. Schools in the left peak are low diversity and those on the right peak are high diversity, with some in equilibrium and others in transition between the two potential steady states. Figure 3 shows this sorting pattern is consistent across income groups, albeit much weaker in high income areas.

To formalize this categorization, I hypothesize that the distribution of SDI_{it} is a combination of the two normally distributed equilibrium types. The true type is unobserved, but the kernel densities show the potential for two distinct latent subgroups. I estimate the mean and standard deviation of the two latent distributions using a finite mixture model and maximum likelihood estimation. Studies in the urban literature use mixture models to identify housing submarkets, where observed physical characteristics of homes determine groups of relatively close substitute housing units (Ugarte et al. 2004, Belasco et al. 2012). The result of this procedure are estimates of the mean and standard deviation of each underlying distribution, along with a predicted probability of each school being of a certain type, conditional on the observed level of diversity SDI_{it}.

The density function $f(SDI|X)$ can be expressed as a linear combination of the $g \in \{low, high\}$ diversity group densities, with group specific parameters $\theta_g \in \{\mu_g, \sigma_g^2\}$. The data generating process is

$$f(SDI|X,\theta) = \rho f_{low}(SDI|\theta_{low}) + (1-\rho) f_{high}(SDI|\theta_{high}), \tag{4}$$

where ρ is the proportion of all schools in the low diversity group. The group-specific mean, μ_g, and standard deviation ,σ_g, of SDI can be estimated by maximizing the log-likelihood function

$$\max_{\theta_{low},\theta_{high}} log\ L = log[\rho f_{low}(SDI|\theta_{low}) + (1-\rho) f_{high}(SDI|\theta_{high})]. \tag{5}$$

The mixture model has an additional benefit of recovering estimated probabilities that each school is either high or low diversity type. Following Deb and Trivedi (2013), I use the observed value of SDI_{it} to estimate the posterior probability that school i is a low or high diversity type in the year 2000. Since type is a binary outcome, the model predicts the logit probabilities, ρ_i, of each school being categorized as a low diversity type. The binary assignment rule for a school diversity type follows

$$LowDiversity_i = \begin{cases} 1, & \text{if } \rho_i \geq 0.5 \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

School i is assigned to the low diversity group if the model predicts a greater than 50% probability that i is in the latent low diversity distribution. Otherwise the school is characterized as high diversity. This strategy yields a roughly equal share of neighborhoods classified as high or low diversity (51% vs 49%, respectively).

Columns two and three of table 1 show that schools in diverse neighborhoods tend to be larger, have a larger share of Hispanic and black students, and a larger share of free lunch students. In my main specification I also categorize schools as small or large urban, small or large suburban, or rural to account for the underlying association between racial composition and urban neighborhoods. In terms of geographic categorization, diverse schools are 40% urban, 27% suburban, and 33% rural; while low diversity schools are 21% urban, 22% suburban, and 57% rural. Geographic controls are included in my main regression and account for cross-sectional differences in mortgage values that average across market type.

To frame the results that follow, neighborhoods can be further described as majority white, black, or Hispanic. Of the 3,661 neighborhoods in my sample most (78%) have population that is greater than 50% white. Roughly 7% and 8% are majority black and

Hispanic, respectively. Nearly 90% of the low diversity neighborhoods are majority white, in contrast to high diversity areas where 60% of the neighborhoods are majority white, 8% are majority black, 13% are majority Hispanic, and 19% have no clear majority group.

3 Empirical Strategy

There are three parts to my estimation strategy. First, I estimate the effect of higher levels of neighborhood diversity on mortgage values and prices using a fixed effects model and observed levels of school diversity. Next, I illustrate the exogenous effect of LIHTC on neighborhood socioeconomic outcomes using a flexible event study design that allows for heterogeneity across neighborhood type. I show visual evidence highlighting the average effect of new subsidized housing development on neighborhood race and income characteristics, along with other school level inputs such as teacher counts and total student enrollment. Lastly, I estimate the effect of increased diversity on mortgage values by constructing a shift-share type instrument based on the growth rates of each demographic group as predicted by the event study model. The setup of this strategy follows closely from Boustan (2010) and Derenoncourt (2018) who explore how the Great Migration of 1940-1970 affected the residential decisions of white households, and the upward mobility of black households, respectively.

I control for local home prices in each of my main models to sharpen my analysis in two ways. For LIHTC to be a valid instrument for school demographic shifts, one assumption is that LIHTC only affects mortgage values through changes in school demographics. Holding prices constant absorbs the effects of other spillovers from LIHTC described in the literature. It follows that variation in new mortgages associated with changes in school diversity results from sorting based on school demographics. I also use prices to frame my estimates in terms of the loan-to-value ratio. When mortgage values changes holding median home values constant, the LTV changes and inference can be made about the financial standing of new home buyers in the neighborhood.

3.1 Neighborhood Price and Mortgage Data

I observe the value of mortgage contracts for single family home purchases along with the income of the home buyer from 2000-2014. Using public data available under the Home Mortgage Disclosure Act (HMDA) that provides the census tract of the home purchase, I map each observation to a school zone and use the median mortgage value in a given year as my outcome measure of neighborhood price. Median home buyer income is the second of two neighborhood income measures, the other described in section 2.3 as the fraction of students receiving free lunch. I weight all of specifications by the mortgage transaction count in a neighborhood for the year 2000.[5]

I measure home value appreciation using the FHFA house price index, or HPI. The HPI measures the appreciation rate of local house prices relative to the base year 2000. HPI is observed at the census tract level and the aggregate median is used to measure price appreciation at the school zone level. For comparability, I compute an aggregate index of mortgage loan values for home purchases within a school attendance zone from 2000-2014. The two measures approximate school zone level variation in mortgage loan amounts and home values relative to the base year of 2000. When paired with school zone fixed effects in my main regressions, my estimates capture the covariation between school diversity and mortgage values within a school zone over time, conditional on home value appreciation used as a control.

Descriptive statistics are presented in table 1 for mortgage values and home buyer income. Both measures are deflated to 2014-dollar values using CPI inflation factors from Oregon State University (Sahr 2014). The median home loan in 2014 prices is $237,094 and the median home buyer earns roughly $87,039 annually. Mortgage values in high diversity neighborhoods are 13% higher than low diversity neighborhoods, with essentially no

[5]The average neighborhood has 485 mortgage transactions in the year 2000.

difference in home buyer incomes.

igure 4 presents a binned scatter plot of the raw data for neighborhood SDI and mortgage values. The positive correlation between SDI and mortgage values across all observations is likely related to the propensity of diverse neighborhoods to be in cities with higher average mortgage values overall, not preferences for a diverse set of neighbors. The positive correlation could also be a byproduct of diverse neighborhoods being on average located closer to the city center, and thus systematically related to higher prices by proximity and not preferences. In the empirical tests to follow I employ school zone and year fixed effects, and thus estimate the effects of within neighborhood changes in diversity levels on mortgage values across time. This eliminates spurious cross sectional variation that would bias estimated effects of neighborhood diversity on mortgage values.

3.2 Fixed Effects Estimation

A naive fixed effects model relating mortgage values to neighborhood diversity can be expressed as

$$M_{it} = \alpha_1 SDI_{it} + \alpha_2 P_{it} + X_{it}^{'}\beta + \gamma_i + \gamma_t + \epsilon_{it}, \tag{7}$$

where M_{it} is the log of the mortgage value index for school zone i year t. $\P_{it}$ is the log of the FHFA house price index. Both school zone (γ_i) and year (γ_t) fixed effects are included in the regression along with a vector of other school and geographical characteristics X_{it}. Descriptive neighborhood covariates include the two income measures, total enrollment counts, and teacher counts. I estimate the average effect of diversity and test for heterogeneous effects across several neighborhood types.

Estimates of α_1 and α_2 from equation 7 are presented in table 2. Each specification includes neighborhood and year fixed effects. Across specifications I vary the inclusion of the income and home buyer demographic measures, to test how the relationship between diversity levels and mortgage values is explained by changes in the composition of home buyers. Comparing the results of the fixed effects models, the effects of diversity in table 2 are small and imprecisely estimated.

The empirical predictions of Bayer et al. (2007) appear to hold similarly in the fixed effects estimation of table 2. When fixed effects are included, neighborhood quality attributes associated with racial demographics are held constant, and the effect of school diversity is nil. Eliminating cross sectional differences between school diversity and mortgage values is an advantage of fixed effects estimation but without exogenous, time varying changes to school demographics the model lacks identifying power.

4 LIHTC Development as Exogenous Shocks to SDI_{it}

Created as part of the Tax Reform Act of 1986 and managed by the Internal Revenue Service, the goal of the LIHTC program is to increase the supply of rental housing in the US. Federal tax credits are allocated to state and municipal housing authorities that distribute the credits to developers in a competitive application process. The Internal Revenue Service requires these agencies to release annually a detailed plan of how developer applications for LIHTC funding are scored and ranked for approval. Developers of approved projects sell the credits to passive investors in exchange for operational cash flow at an estimated 75 cents for every dollar of tax credit, and investor tax benefits are realized over a ten-year period post investment (Eriksen 2009). The location of LIHTC development in my sample neighborhoods are shown in figure 5.

Program guidelines require either 20% of tenants earn less than 50% of the metro area median income or at least 40% of tenants earn less than 60% area median income. Although developers are not required to rent the most units to income qualified tenants, the amount of tax credits received increases as the percentage of units occupied by low income residents goes up. In the data I find on average the share of units reserved for low income residents far exceeds the 20% and 40% thresholds. The rent limit for these units is 30% of the income level

(50% or 60% of area median) required to satisfy the resident income criteria. In practice, if 90% of the total units are leased to low income tenants, the rent charged on these specific units is $0.3 \times 0.6 = 0.18$ or 18% of the monthly median income for the area, adjusted for household size. An area of concern is the actual affordability of LIHTC units, as the rent is determined by median income of an entire metropolitan area, while locally a neighborhood that receives LIHTC could be very poor. Housing studies have shown that the share of LIHTC units that exceed fair market rents can exceed 30% in many cities (Cummings and DiPasquale 1999).

The direct effect of LIHTC on a demographic measure like SDI_{it} in principle depends on several factors. The results of a 2012 survey of LIHTC housing units shows that in terms of race, who lives in LIHTC depends largely on the region of the country (Hollar 2014). In southern states like South Carolina, Mississippi, Louisiana and Georgia many tenants are black, contrasting with parts of Appalachia like Kentucky or West Virginia where the majority of LIHTC tenants are white. Davis et al. (2018) relates the households that take up LIHTC housing as those in the bottom tercile of the housing market income distribution. Given the potential for LIHTC units to be priced at or above market rents, it is difficult to conclude that residents served are those of highest need (McClure 2010).

The exclusion restriction requires the LIHTC shock to only affect mortgages through changes in school racial diversity, conditional on prices and other neighborhood characteristics. The price index serves to absorb the price effects of variation in neighborhood amenities following LIHTC development discussed by Diamond and McQuade (2019). The income controls capture the changes in economic composition of neighbors over time, thus if the exclusion restriction holds I am able to interpret my estimates as the effect of school diversity on the median loan to value ratio.

4.1 LIHTC Data

I merge data for the timing and location of LIHTC builds from 2000-2014 to the student demographic data to form the full panel of data. Recall that I restrict the data to school zones with no LIHTC activity prior to 2000, and either one or zero new LIHTC builds during the years of observation. In the final sample of 3,661 neighborhoods, 21% received new subsidized housing during the sample period.

The LIHTC data come from HUD and describe the year the units are available for rent, the number of units put in service, and the physical location of the LIHTC complex. Two-thirds of the construction in the sample became active in 2007 or prior with peak construction in 2006 when 10% of the total projects in my sample became available to rent. Table 1 shows that on average LIHTC builds in the sample have about 288 units available for rent, representing roughly 2.9% of the total housing in the respective school zone. The typical LIHTC complex in the sample has roughly 260 of the 288 units held aside for low income tenants. The low income tenant share is well above either of the thresholds required to receive tax credits. From the HUD data I use the physical address of the LIHTC complex to map units to a school zone. The map in figure 5 shows the spatial distribution of LIHTC activity in my sample.

4.2 Empirical Effects of LIHTC Development on Neighborhood Demographics

In this section I estimate the effects of LIHTC on neighborhood demographic and economic outcomes using a dynamic event study model with two-way fixed effects. I predict changes to a given socioeconomic outcome Y_{it} as a function of exogenous exposure to a LIHTC shock, Z_{it} such that

$$Y_{it} = Z_{it}^{'}\pi + \delta P_{it} + \gamma_i + \gamma_t + \xi_{it}. \tag{8}$$

The coefficient π in equation 8 is an estimate of the average effect of new LIHTC development on neighborhood diversity, conditional on neighborhood and year fixed effects γ. P_{it} is the log of the FHFA house price index, and ξ_{it} is the exogenous error term. To capture the variation in outcome Y_{it} that is only related to exogenous effects of LIHTC on the racial composition of schools, I construct Z_{it} as the interaction of a set of event time indicators with indicators for whether the school was low or high diversity in 2000

$$Z_{it} = \sum_{g=1}^{2} \sum_{k=-6}^{10} (I_{ig} \times D_{ik}). \tag{9}$$

Each of the group indicators I_{ig} are interacted with a dummy $D_{ik} = 1$ if school zone i is k years pre or post LIHTC development and 0 otherwise. School zones that are never treated have the property $D_{ik} = 0$ for each panel observation. The interacted event study instruments approximate neighborhood exposure to LIHTC as a function of both the time since units become available for rent and the intensity of the Y_{it} response as predicted by the pre-treatment neighborhood type. Substituting equation 9 into equation 8, I express the event study model as

$$Y_{it} = \sum_{g=1}^{2} \sum_{k=-6}^{10} (I_{ig} \times D_{ik})\pi_{gk} + \delta P_{it} + \gamma_i + \gamma_t + \xi_{it}. \tag{10}$$

π_{gk} is now a set of $g \times k$ coefficients that map the estimated average treatment effect of LIHTC on school diversity for each k = 6 years pre to 10 years post development, by initial neighborhood type g. With the two way fixed effects my identification stems from changes within a school zone over time, controlling for time invariant neighborhood characteristics and year specific changes to Y_{it}. Included in my estimation of equation 10 are census division specific linear time trends to capture general demographic shifts unique to different regions of the country.

I estimate equation 10 separately for white, black, Hispanic, and Asian enrollment counts and two income measures - home buyer income and the fraction of students receiving free lunch. The heterogeneous event study coefficients π_{gk} are plotted in figure 6 through figure 11. On the horizontal axis is the time (in years) k since the new LIHTC units became available at $k = 0$. On the vertical axis is the magnitude of the effect on each Y_{it} outcome. Each point represents the average effect of LIHTC k years away from the in service year, relative to one year prior at $k = -1$. Following the literature, I exclude observations for neighborhoods receiving LIHTC in the year 2000 as there are no pre-treatment observations for this initial cohort. These estimates trace out the average effect of the subsidized housing development on my socioeconomic measures over time, by neighborhood type.

The effect of LIHTC on white enrollment counts shown in figure 6 imply a substantial increase of white residents in low diversity areas, with a smaller magnitude decrease in high diversity areas. The positive effect in low diversity areas is gradual, with no immediate jump in the year units become available or one year post development. Instead, beginning in year two white enrollment rises consistently until peaking at year five. The nature of this effect implies that white residents are moving in to low diversity neighborhoods not because of LIHTC take up, but in a way consistent with gentrification and improved neighborhood quality following new LIHTC development as proposed by Diamond and McQuade (2019). The gradual decrease in high diversity areas is (non-exhaustive) evidence that white households that choose to live near development have preferences over the type of neighborhood where the development takes place.

The estimated effects for black and Hispanic enrollment are presented in figures 7 and 8. Again, I find heterogeneous effects by initial neighborhood type, albeit an opposite pattern from the behavior of whites. Black and Hispanic enrollments rise in high diversity areas and fall in low diversity areas. The small initial bump in high diversity areas for years zero and one suggest the at least part of this change in driven by black and Hispanic residents of the new LIHTC development. After this initial increase, black enrollment growth declines and Hispanic enrollment growth remains flat for the large part of the event window. In low diversity areas the enrollment of both groups decreases, and although the evidence is not conclusive this is potentially a product of rising prices following the in migration of white residents in these neighborhoods. For completeness, figure 9 shows that Asian student enrollment decreases over time in high diversity neighborhoods following LIHTC, with a very small predicted increase in low diversity areas.

Two measures of neighborhood income are tested in the event study framework and plotted in figures 10 and 11. Home buyer income and the number of students receiving free lunch subsidies address two different regions of the income distribution. The former is more closely related to the upper end of the income distribution while the latter is more related to poverty levels in the neighborhood. The covariance between the two is important- if home buyer income is increasing as the number of students receiving free lunch increases, it is an indication of rising neighborhood inequality. Figures 10 and 11 show that LIHTC development does not fundamentally change home buyer income or the fraction of students receiving free lunch, holding home values constant. Given that both of these estimates are smooth and continuous across the event-time threshold, I argue that my event-study model identifies exogenous changes in school demographics around the timing of LIHTC development.

5 Instrumenting for Changes to Neighborhood Diversity

To estimate the causal effect of school demographic changes on neighborhood mortgage values, I construct a shift-share instrument for SDI_{it} based on predicte demographic changes shown in section 4.2. Starting with the observed enrollment counts used to construct initial values of SDI_{it} in the year 2000, I apply the growth rate of the predicted counts from the event study model to project exogenous growth of school enrollment shares for each racial group. The result of this procedure is an exogenously determined shock to neighborhood diversity, $\widetilde{SDI}_{it}$. My use of the shift share instrument most closely follows from Derenoncourt

(2018) and Boustan (2010), however these measures have been employed widely in various urban and regional growth settings (Goldsmith-Pinkham et al. 2018).

The predicted enrollment growth rate of each race group r for neighborhood i in year t can be expressed as

$$\hat{g}_{rit} = \frac{\hat{n}_{rit} - \hat{n}_{rit-1}}{\hat{n}_{rit-1}}, \tag{11}$$

where $\hat{n}_{rit}$ is the fitted value population count generated by model 10. Recall that SDI_{it} in the year 2000, used to categorize neighborhoods into initial types, is generated from the year 2000 enrollment shares as in equation 1. Behind the year 2000 shares are the observed population counts from that year, n_{ri00}. One year after this initial period I calculate

$$\widetilde{n}_{rit} = n_{ri00} \times (1 + \hat{g}_{rit}). \tag{12}$$

For each year to follow, we have

$$\widetilde{n}_{rit+1} = \widetilde{n}_{rit} \times (1 + \hat{g}_{rit+1}). \tag{13}$$

These predicted growth adjusted population counts for each of the five race groups are used to construct the growth adjusted shares

$$\widetilde{s}_{rit} = \frac{\widetilde{n}_{rit}}{\sum_{r=1}^{5} \widetilde{n}_{rit}}. \tag{14}$$

I then calculate $\widetilde{SDI}_{it}$ using the growth adjusted shares in 14 and the procedure of section 2.1. To estimate the effects of diversity on mortgage values, I specify a two stage least squares model where I instrument for diversity using $\widetilde{SDI}_{it}$.

5.1 2SLS Effect of SDI_{it} on Mortgage Values

I model the log of median mortgage values in neighborhood i, year t as a function of neighborhood diversity and other neighborhood characteristics in a two stage least squares specification. The first stage is

$$SDI_{it} = \delta_1 \widetilde{SDI}_{it} + \delta_2 P_{it} + X_{it}^{'}\beta + \gamma_i + \gamma_t + \xi_{it}, \tag{15}$$

and the main estimating equation takes the form

$$M_{it} = \lambda_1 \widehat{SDI}_{it} + \lambda_2 P_{it} + X_{it}^{'}\beta + \gamma_i + \gamma_t + \epsilon_{it}. \tag{16}$$

Neighborhood covariates include mortgage borrower income and the fraction of subsidized lunch students; along with the log of total enrollment, and the log of full time teacher headcounts to control for changes to school level inputs. Additionally I include separate dummy variables for a neighborhood in a large city, a small city, or large suburban area. Each of these categorical variables are interacted with an indicator that equals one if the observation is post LIHTC shock. I include these interacted geographical controls to account for supply and demand effects in the market for available land following new LIHTC development. The assumption is that supply and demand effects of land supply associated with new housing development is similar across neighborhood locales.

I present the results of my preferred model in table 3. Each of the models shown include neighborhood and year fixed effects but vary in the inclusion of the explanatory controls for neighborhood income and home buyer demographics. On average, price-constant mortgage values in a neighborhood increase as the area becomes more diverse. Mortgage values are predicted to rise by roughly 5% for a one standard deviation increase of SDI_{it} of 2.56, controlling for median prices and attendance zone fixed effects. The estimate is robust to controlling for changes in neighborhood income and the fraction of non-white home buyers. This result suggest that as schools become more diverse, the average home buyer borrows more for a similarly priced home in real prices. In principle this implies a rising loan to value ratio for the median house near diversifying schools.

5.2 The Effect of SDI_{it} on Mortgages, by Home Buyer Race

Higher loan to value ratios as predicted by my model imply that home buyers near diversifying schools are making smaller down payments as a fraction of the sale price. Since I control for prices, the estimate is driven by variation in mortgage values not associated with changes to neighborhood price levels. If down payment ability is largely explained by wealth, the next step is to test how the main effect varies by race of the buyer. To do this I estimate equation 16 for the mortgage value index calculated separately for buyers of each race group. The results are shown in table 5.

The results in table 5 show the average effect of diversity on loan to value is consistent across borrower type. If borrowers near diverse schools have less wealth and make smaller down payments, this holds true for white, black, and Hispanic buyers. Each specification includes income controls, suggesting that higher wealth households within each race group tend to sort away from diverse schools. The model predicts that one implication of racial wealth gaps shows up in the borrowing levels of white households in diversifying, low-income neighborhoods. White buyers in these areas make larger down payments as a percentage of the sell price on average, decreasing the loan to value ratio relative to other areas.

5.3 2SLS Effect of SDI_{it} on Prices

The price constant analysis of mortgages in the prior sections suggest that home buyers near diversifying schools have lower levels of wealth, regardless of racial background. The price index measure has an important property a measure of constant-quality price levels in a given area. In theory, observed price changes are then driven by factors related to neighborhood amenities. I test for the effect of exogenous diversity on the house price index by estimating a version of equation 16 that takes the log of the house price index as the dependent variable.

The results are presented in table 6.

Columns one through four are fixed effects estimates that vary by inclusion of various explanatory variable types. The results are small but precisely measured negative effect of diversity on house prices. Column five is the 2SLS estimate, which suggests that when SDI_{it} increases by one standard deviation, or 2.54, prices fall by roughly 2.8%. This is consistent with theories of decentralized sorting that suggest majority white neighborhoods will demand higher prices when preferences over the race of neighborhoods exist. The result in column five is robust to controlling for changes in income and the racial characteristics of home buyers.

6 Discussion

It is possible for the effects of LIHTC correlated with racial diversity changes could bias the estimate of the effect of diversity on prices. The exclusion restriction holds, however, when prices are used as a control in the mortgage regressions. Any time varying changes in the mortgage index associated with the LIHTC instrument should be capture by the price index and income controls. Including the home buyer income and racial characteristics, neighborhood poverty levels, along with prices allows me to identify a causal effect changes in school diversity on mortgage borrowing levels. The results of this paper provide evidence that lower wealth home buyers sort into neighborhoods near diversifying schools. As diversity levels rise the median loan to value ratio rises, implying smaller down payments made for purchases in these areas.

If, as the model suggests, prices indeed decrease when schools become more diverse, home buyers benefit from lower prices but leverage a larger percentage of the value in a home loan. The literature has shown that higher loan to value ratios predict the probability of default on a loan (Floros and White 2016). This could contribute to poorer loan performance near diverse schools as a byproduct of lower wealth households sorting into these areas. My results suggest wealth based sorting is present for home buyers in each race group of my sample.

I posit two potential explanations for rising mortgage values occurring simultaneously with falling home values. It is possible that as schools become more diverse, new home buyers select a higher quantity of housing as compared to the quantity of housing measured by the house price index, which approximates price changes of physically comparable homes over time. This implies that rising mortgage values are associated with higher wealth households moving into neighborhoods near diversifying schools and purchases larger/higher quality homes. An alternative explanation is decreased demand for housing near diversifying schools reduces home values as some households with preferences against diverse schools exit. Boustan (2010) argue that this behavior contributed to increases in black home ownership from 1950-1980, when home values in the central city decreased as white households moved to suburban areas. In my case mortgage values are rising as lower wealth households sort into neighborhoods near diverse schools and buy homes at lower values, but borrow more in terms of real mortgage amounts as wealth constrains down payment ability.

My analysis also sheds light on a potential mechanism by which gentrification of low-income areas takes place. In low-income, diversifying areas white home buyers borrower have lower loan to value ratios which imply larger down payment percentages. This is a secondary effect of wealth as this result does not hold for buyers of other racial backgrounds. It follows that white home buyers with preferences for diverse schools can gain a higher quantity of housing services and incur less mortgage interest expense over the life of the loan. This makes low-income, diversifying areas attractive for white home buyers with moderate down payment ability.

7 Conclusion

In this paper I tests how theoretical predictions of neighborhood choice models affect mortgage markets as school racial diversity changes. Using panel data describing neighborhood demographic characteristics I find that mortgage values rise as neighborhoods become more diverse, holding prices constant. The effects holds for home buyers of each race and ethnic group in the data. I also find suggestive evidence that home values decline following school diversity shocks from LIHTC rental development. Two potential mechanisms are presented to explain changes to these two housing market measures, both implying changes to neighborhood wealth. Testing the model over various subsets of the data, I find suggestive evidence that the effect varies by neighborhood type. My results suggest that household sorting over school demographics can lead to changes in aggregate neighborhood wealth.

References

AUD, S. (2011): *Status and Trends in the Education of Racial and Ethnic Groups (2010)*, DIANE Publishing.

BAUM-SNOW, N. AND F. FERREIRA (2015): "Causal inference in urban and regional economics," *Handbook of regional and urban economics*, 5, 3–68.

BAUM-SNOW, N. AND J. MARION (2009): "The effects of low income housing tax credit developments on neighborhoods," *Journal of Public Economics*, 93, 654–666.

BAYER, P., F. FERREIRA, AND R. MCMILLAN (2007): "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 115, 588–638.

BELASCO, E., M. FARMER, AND C. LIPSCOMB (2012): "Using a finite mixture model of heterogeneous households to delineate housing submarkets," *Journal of Real Estate Research*, 34, 577–594.

BIAN, X., Z. LIN, AND Y. LIU (2018): "House price, loan-to-value ratio and credit risk," *Journal of Banking & Finance*, 92, 1–12.

BOUSTAN, L. P. (2010): "Was postwar suburbanization âĂIJwhite flightâĂİ? Evidence from the black migration," *The Quarterly Journal of Economics*, 125, 417–443.

CAETANO, G. AND V. MAHESHRI (2017): "School segregation and the identification of tipping behavior," *Journal of Public Economics*, 148, 115–135.

COUTURE, V. AND J. HANDBURY (2017): "Urban revival in America, 2000 to 2010," Tech. rep., National Bureau of Economic Research.

CUMMINGS, J. L. AND D. DIPASQUALE (1999): "The low-income housing tax credit an analysis of the first ten years," *Housing Policy Debate*, 10, 251–307.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor (1999): "The rise and decline of the American ghetto," *Journal of political economy*, 107, 455–506.

Davis, M. A., J. Gregory, and D. A. Hartley (2018): "The Long Run Effects of Low Income Housing on Neighborhood Composition," .

Deb, P. and P. K. Trivedi (2013): "Finite mixture for panels with fixed effects," *Journal of Econometric Methods*, 2, 35–51.

Derenoncourt, E. (2018): "Can you move to opportunity? Evidence from the Great Migration," Tech. rep., Harvard University, mimeo.

Diamond, R. (2016): "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 106, 479–524.

Diamond, R. and T. McQuade (2019): "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development," *Journal of Political Economy*, 127, 000–000.

Epstein, G. S. (2008): "Herd and network effects in migration decision-making," *Journal of Ethnic and Migration Studies*, 34, 567–583.

Eriksen, M. D. (2009): "The market price of low-income housing tax credits," *Journal of Urban Economics*, 66, 141–149.

Floros, I. and J. T. White (2016): "Qualified residential mortgages and default risk," *Journal of Banking & Finance*, 70, 86–104.

Freedman, M. and E. G. Owens (2011): "Low-income housing development and crime," *Journal of Urban Economics*, 70, 115–131.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018): "Bartik instruments: What, when, why, and how," Tech. rep., National Bureau of Economic Research.

HOLLAR, M. K. (2014): "Understanding whom the LIHTC program serves: Tenants in LIHTC units as of December 31, 2012," *Available at SSRN 2563151*.

LOGAN, T. D. AND J. M. PARMAN (2017): "The national rise in residential segregation," *The Journal of Economic History*, 77, 127–170.

MCCLURE, K. (2010): "Are Low-Income Housing Tax Credit Developments locating where there is a shortage of affordable units?" *Housing Policy Debate*, 20, 153–171.

REARDON, S. F. (2016): "School segregation and racial academic achievement gaps," *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2, 34–57.

RHOADES, S. A. (1993): "Efficiency effects of horizontal (in-market) bank mergers," *Journal of Banking & Finance*, 17, 411–422.

ROTHENBERG, J., G. C. GALSTER, R. V. BUTLER, AND J. R. PITKIN (1991): *The maze of urban housing markets: Theory, evidence, and policy*, University of Chicago Press.

SAHR, R. (2014): "Inflation Conversion Factors for Years 1774 to Estimated 2023," .

SCHELLING, T. C. (1971): "Dynamic models of segregation," *Journal of mathematical sociology*, 1, 143–186.

UGARTE, M., T. GOICOA, AND A. MILITINO (2004): "Searching for housing submarkets using mixtures of linear models," in *Spatial and Spatiotemporal Econometrics*, Emerald Group Publishing Limited, 259–276.

Tables and Figures

Table 1: Summary Statistics

	Full Sample	Low Diversity	High Diversity
Real mortgage value ($2014)	237,094.75	221,693.82	250,772.12
	(111153.9)	(89719.7)	(125650.8)
Real home buyer income	87,039.43	85,278.73	88,603.08
	(35326.8)	(31915.4)	(38042.2)
School Diversity Index, 2000	4.64	1.59	6.19
	(2.564)	(0.934)	(1.537)
School Diversity Index, 2014	5.96	4.51	7.24
	(2.334)	(2.199)	(1.581)
Fraction Black	0.15	0.11	0.19
	(0.196)	(0.201)	(0.185)
Fraction Hispanic	0.24	0.16	0.31
	(0.232)	(0.204)	(0.232)
Fraction Asian	0.06	0.05	0.07
	(0.0907)	(0.0826)	(0.0964)
School Enrollment	795.45	751.64	834.36
	(345.4)	(358.2)	(328.8)
Fraction Free Lunch	0.40	0.32	0.47
	(0.251)	(0.243)	(0.238)
Neighborhoods with LIHTC	0.25	0.23	0.27
	(0.432)	(0.420)	(0.442)
LIHTC units per build	288.28	278.98	294.99
	(168.62)	(165.92)	(170.89)
Neighborhood housing units	9,703.09	8,315.91	11,001.73
	(7830.2)	(5983.9)	(9039.9)
Middle Schools	3,661	1,905	1,756
School×Year Observations	54,915	27,645	29,502
Share of Total LIHTC		0.52	0.48

Notes: The school diversity index is described in the empirical strategy and takes a value between 0 and 10, with 10 being most diverse. The mean number of LIHTC units per build is calculated with only school zones that received the treatment. Neighborhood total housing units include estimates of all occupied and vacant housing of all types.

Table 2: OLS Effect of School Diversity on Local Mortgage Values

Dependent Variable : Log(Mortgage Value Index)	(1) OLS	(2) OLS	(3) OLS	(4) OLS
SDI	0.00154	0.00129	-0.00146	-0.00149
	(0.00157)	(0.00160)	(0.00111)	(0.00111)
Log(House Price Index)	0.879***	0.876***	0.564***	0.566***
	(0.0110)	(0.0111)	(0.0118)	(0.0113)
N	48026	48026	48026	48026
r2	0.854	0.855	0.909	0.909
School Zone & Year FE	×	×	×	×
School and Neighborhood Covariates		×	×	×
Income Characteristics			×	×
Buyer Race Demographics				×

Standard errors are clustered at the school zone level and shown in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports OLS estimates of the effect of school diversity (SDI) on mortgage values. The mean and standard deviation of SDI are 4.64 and 2.56, respectively. neighborhood quality School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Table 3: 2SLS Effect of School Diversity on Local Mortgage Values

Dependent Variable: Log(Mortgage Value Index)	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
$\widehat{SDI}$	0.0283***	0.0286***	0.0201***	0.0202***
	(0.00515)	(0.00534)	(0.00389)	(0.00389)
Log(HPI)	0.888***	0.889***	0.583***	0.584***
	(0.0115)	(0.0115)	(0.0125)	(0.0119)
N	44492	43603	40979	40979
r2	0.564	0.560	0.714	0.716
Fstat	329.7	318.2	312.2	315.6
School Zone & Year FE	×	×	×	×
School and Neighborhood Covariates		×	×	×
Income Characteristics			×	×
Buyer Race Demographics				×

Standard errors are clustered at the school zone level and shown in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports two-stage least squares estimates of the effect of school diversity (SDI) on mortgage values. The instrument for SDI is the shift-share diversity levels as predicted by the LIHTC development shock. The mean and standard deviation of SDI are 4.64 and 2.56, respectively. neighborhood quality School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Table 4: Reduced Form Effect of Shift-Share Instrument on Local Mortgages

Dependent Variable: Log(Mortgage Value Index)	(1) OLS	(2) OLS	(3) OLS	(4) OLS
$\widetilde{SDI}$	0.0500***	0.0499***	0.0355***	0.0359***
	(0.00843)	(0.00864)	(0.00641)	(0.00643)
Log(HPI)	0.882***	0.881***	0.576***	0.578***
	(0.0112)	(0.0114)	(0.0122)	(0.0116)
N	44786	43618	40979	40979
r2	0.851	0.852	0.905	0.906
School Zone & Year FE	×	×	×	×
School and Neighborhood Covariates		×	×	×
Income Characteristics			×	×
Buyer Race Demographics				×

Standard errors are clustered at the school zone level and shown in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports reduced form estimates for the shift-share diversity levels regressed on the mortgage value index. School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Table 5: 2SLS Effect of School Diversity on Local Mortgages, by Race of Buyer

Dependent Variable: Log(Mortgage Value Index)	2SLS White	2SLS Black	2SLS Hispanic	2SLS Asian
$\widehat{SDI}$	0.0271***	0.0242*	0.0248**	0.0176
	(0.00454)	(0.0101)	(0.00768)	(0.0105)
$\widehat{SDI}$ × Low-Income	-0.0414***	0.00531	-0.0111	0.00404
	(0.0105)	(0.0268)	(0.0193)	(0.0340)
N	40258	22264	28618	24058
r2	0.430	0.127	0.187	0.172
Fstat	160.0	136.0	149.6	135.3
School Zone & Year FE	×	×	×	×
House Price Controls	×	×	×	×
Neighborhood and Buyer Controls	×	×	×	×

Standard errors are clustered at the school zone level and shown in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports two-stage least squares estimates for the effect of school diversity (SDI) on median mortgage values by the race of home owner. Low-income areas are those where school zone median income is in the bottom tercile of the state distribution. The house price index and all other home buyer and neighborhood controls are included in each specification. The instrument for SDI is the shift-share diversity levels as predicted by the LIHTC development shock. The mean and standard deviation of SDI are 4.64 and 2.56, respectively. School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Table 6: The Effect of School Diversity on Local House Prices

Log(House Price Index)	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS
$\widehat{SDI}$	-0.00187	-0.00425	-0.00650***	-0.00577**	-0.0112*
	(0.00240)	(0.00236)	(0.00177)	(0.00177)	(0.00527)
N	51570	50576	47430	47430	43790
r2	0.760	0.766	0.841	0.846	0.343
Fstat					307.0
School Zone & Year FE	×	×	×	×	×
School and Nbhood Controls		×	×	×	×
Income Characteristics			×	×	×
Buyer Race Demographics				×	×

Standard errors are clustered at the school zone level and shown in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports OLS and two-stage least squares estimates for the effect of school diversity (SDI) on the house price index. The instrument for SDI is the shift-share diversity levels as predicted by the LIHTC development shock. The mean and standard deviation of SDI are 4.64 and 2.56, respectively. School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Table 7: The Effect of School Diversity on Local Mortgages, by Neighborhood Type

Dependent Variable Log(Mortgage Value Index)	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
$\widehat{SDI}$	0.0164***	-0.0102	0.0217***	0.0306***	0.0207***
	(0.00473)	(0.0153)	(0.00409)	(0.00480)	(0.00398)
$\widehat{SDI}$ × High-Diversity	0.000986				
	(0.00480)				
$\widehat{SDI}$ ×Majority White		0.0559**			
		(0.0180)			
$\widehat{SDI}$ × Low-Income			-0.0188*		
			(0.00930)		
$\widehat{SDI}$ × Urban				-0.0200***	
				(0.00510)	
$\widehat{SDI}$ × Basis Boost					-0.0108
					(0.0124)
N	40979	40979	40437	40979	40979
r2	0.721	0.695	0.714	0.709	0.714
Fstat	196.0	104.0	164.1	162.1	157.7
School Zone & Year FE	×	×	×	×	×
House Price Controls	×	×	×	×	×
Neighborhood and Buyer Controls	×	×	×	×	×

Standard errors are clustered at the school zone level and shown in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Notes: This table reports two-stage least squares estimates for the effect of school diversity (SDI) on median mortgage values across different neighborhood types. Low-income areas are those where school zone median income is in the bottom tercile of the state distribution. Basis boost areas are those where LIHTC development will receive additional subsidies as outline by program guidelines. The house price index and all other home buyer and neighborhood controls are included in each specification. The instrument for SDI is the shift-share diversity levels as predicted by the LIHTC development shock. The mean and standard deviation of SDI are 4.64 and 2.56, respectively. School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Table 8: The Effect of School Diversity on Local House Prices, by Neighborhood Type

Dependent Variable: Log(House Price Index)	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
$\widehat{SDI}$	-0.0207*	-0.0221	-0.0118*	-0.00458	-0.0107*
	(0.00906)	(0.0154)	(0.00538)	(0.00632)	(0.00537)
$\widehat{SDI}$ × High-Diversity	0.00970				
	(0.00792)				
$\widehat{SDI}$ × Majority White		0.0240			
		(0.0200)			
$\widehat{SDI}$ × Low-Income			0.0103		
			(0.0148)		
$\widehat{SDI}$ × Urban				-0.00810	
				(0.00682)	
$\widehat{SDI}$ × Boost					-0.00865
					(0.0152)
N	43790	43790	43217	43790	43790
r2	0.354	0.360	0.346	0.349	0.344
Fstat	189.8	137.6	158.2	158.2	154.4
School Zone & Year FE	×	×	×	×	×
House Price Controls	×	×	×	×	×
Neighborhood and Buyer Controls	×	×	×	×	×

Standard errors are clustered at the school zone level and shown in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports two-stage least squares estimates for the effect of school diversity (SDI) on house prices across different neighborhood types. Low-income areas are those where school zone median income is in the bottom tercile of the state distribution. Basis boost areas are those where LIHTC development will receive additional subsidies as outline by program guidelines. The house price index and all other home buyer and neighborhood controls are included in each specification. The instrument for SDI is the shift-share diversity levels as predicted by the LIHTC development shock. The mean and standard deviation of SDI are 4.64 and 2.56, respectively. School and neighborhood controls include total enrollment, teacher counts and time trends interacted with indicators for urban and suburban school types. Income characteristics are median home buyer income and the fraction of students receiving free lunch subsidies in the school.

Figures

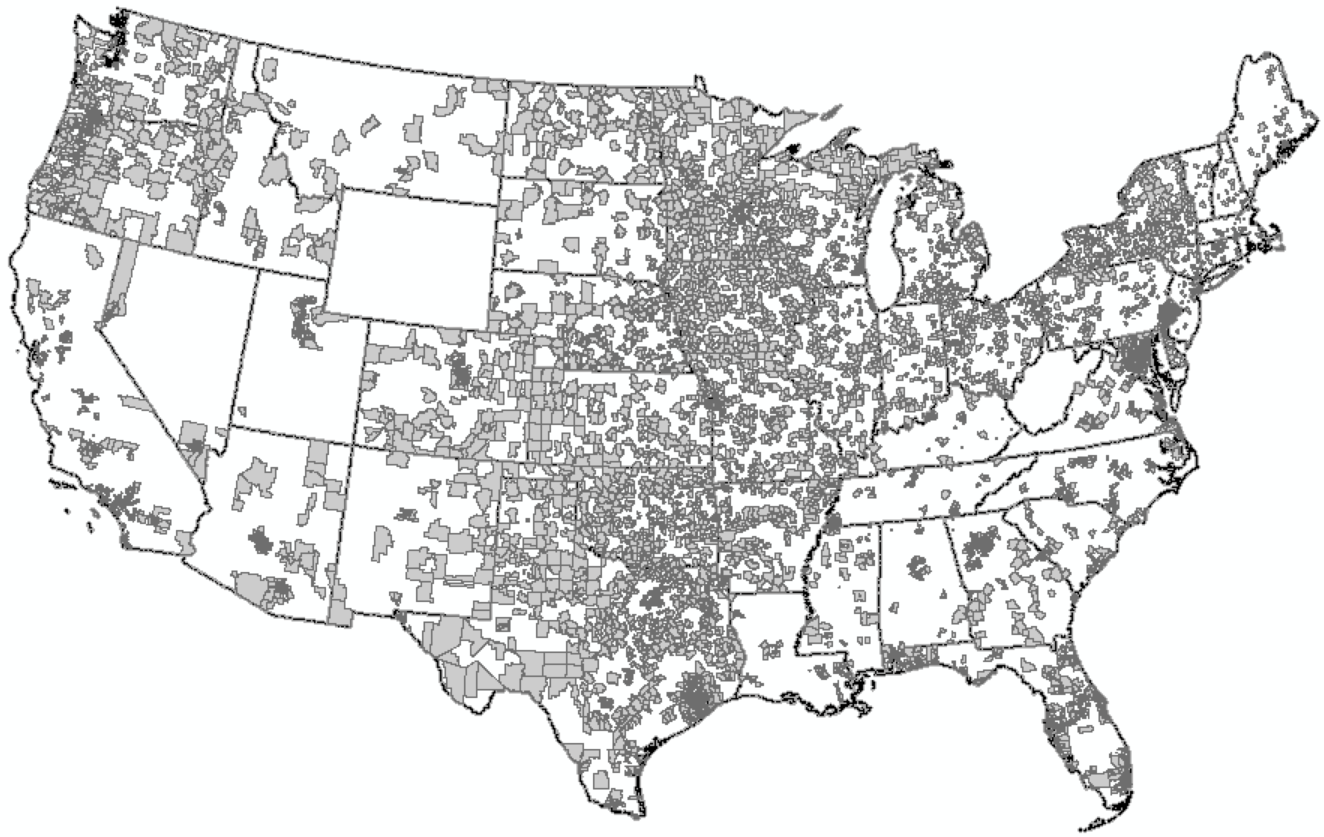
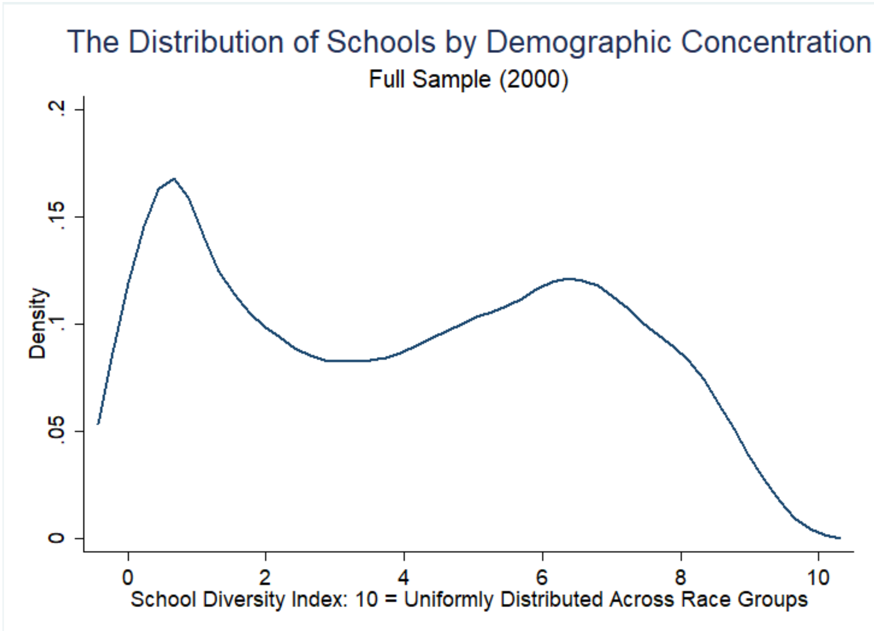
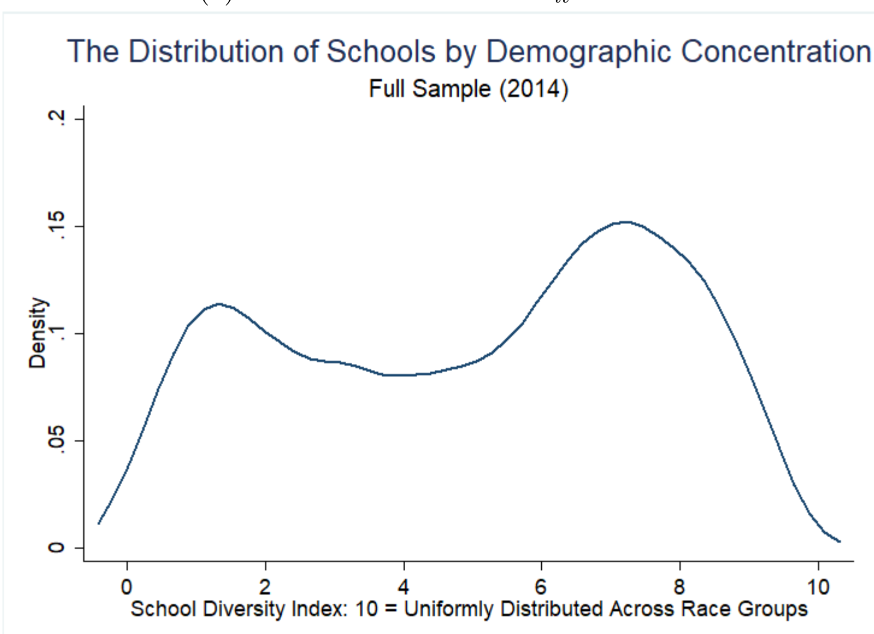


Figure 1: The shaded grey areas are middle school zones used as neighborhoods in my sample. Boundaries are fixed to 2009 attendance zones publicly available at https://www.sabinsdata.org/.

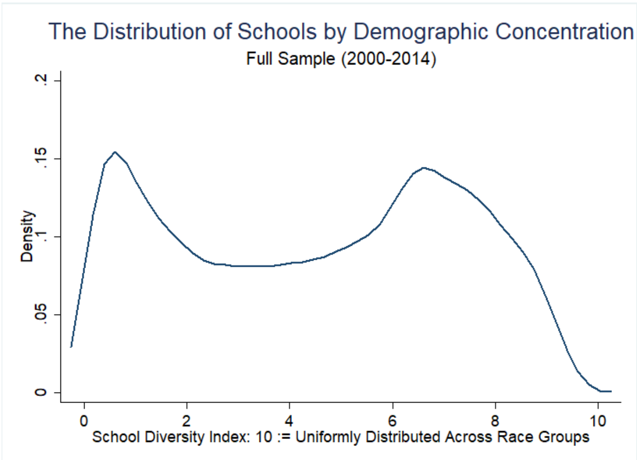


(a) The distribution of SDI_{it} in 2000.

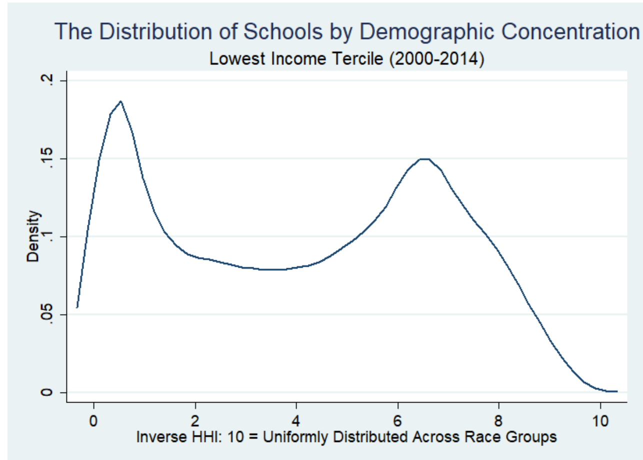


(b) The distribution of SDI_{it} in 2014.

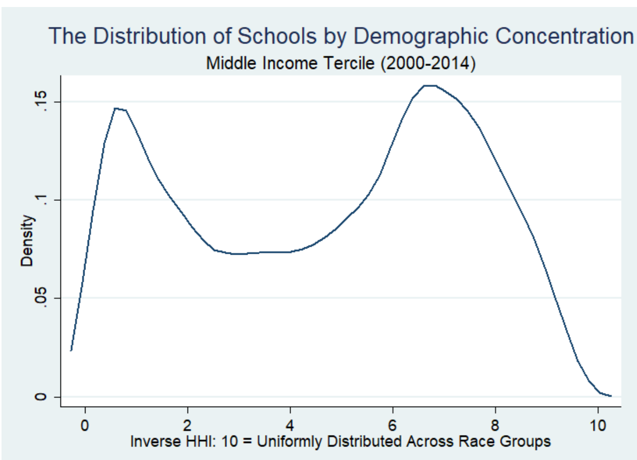
Figure 2: The distribution of schools by racial diversity as measured by SDI_{it}. Schools have generally become more diverse from 2000 to 2014 as evidenced by a larger mass on the right half of the distribution.



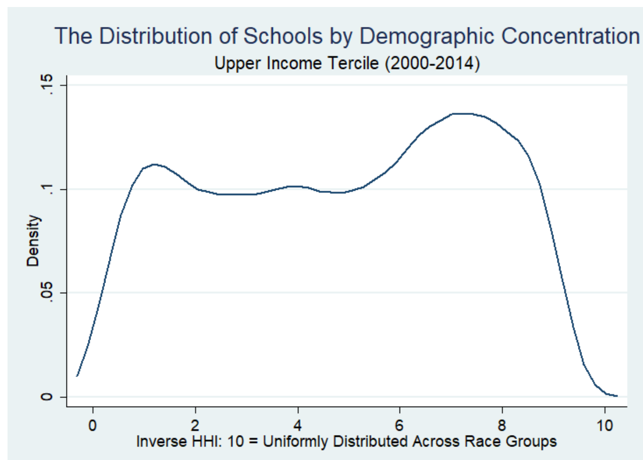
(a) Entire Sample



(b) Low Income Schools



(c) Middle Income Schools



(d) High Income Schools

Figure 3: The shape of the distribution of schools by racial diversity as measured by SDI_{it} is persistent across income levels.

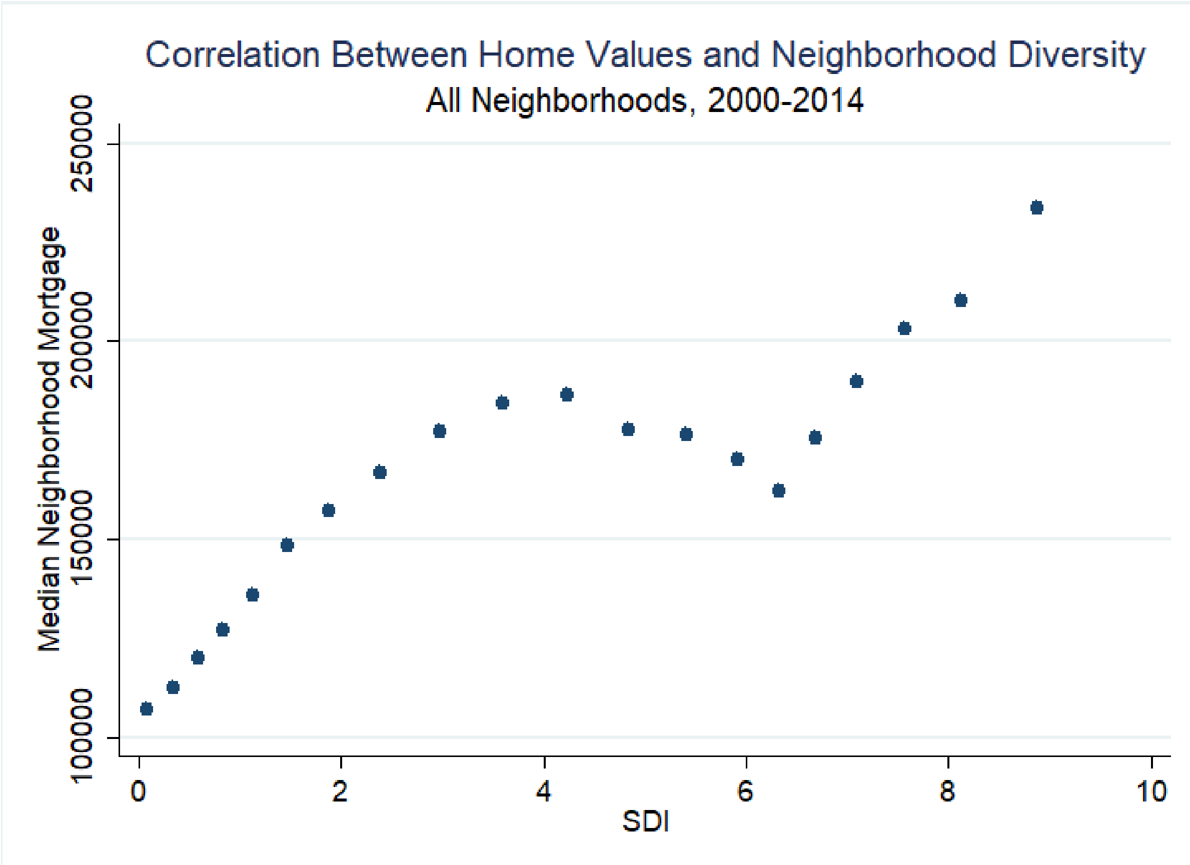


Figure 4: A binned scatter plot of the raw data shows a positive unconditional correlation between neighborhood diversity and mortgage values.

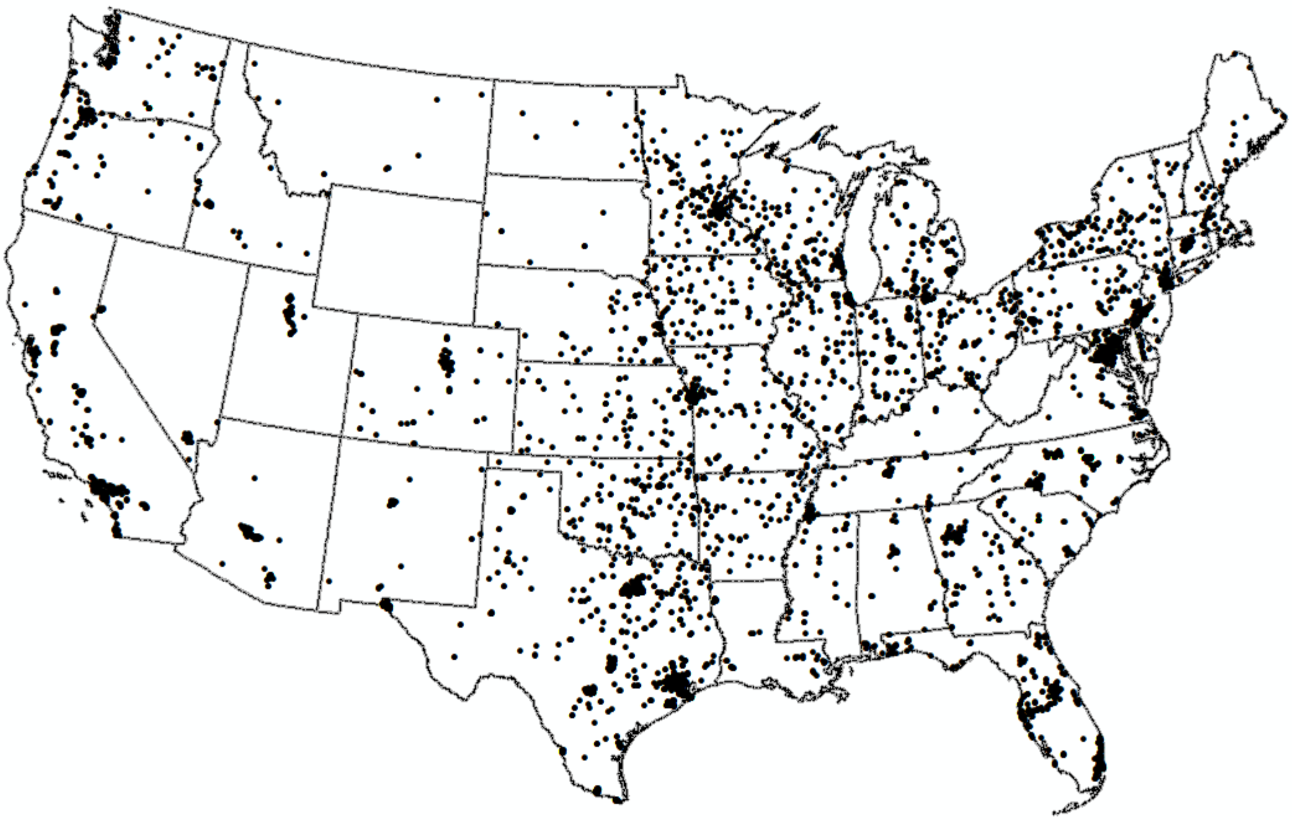


Figure 5: The dots represent individual LIHTC developments from 2000-2014 in my sample middle school zones.

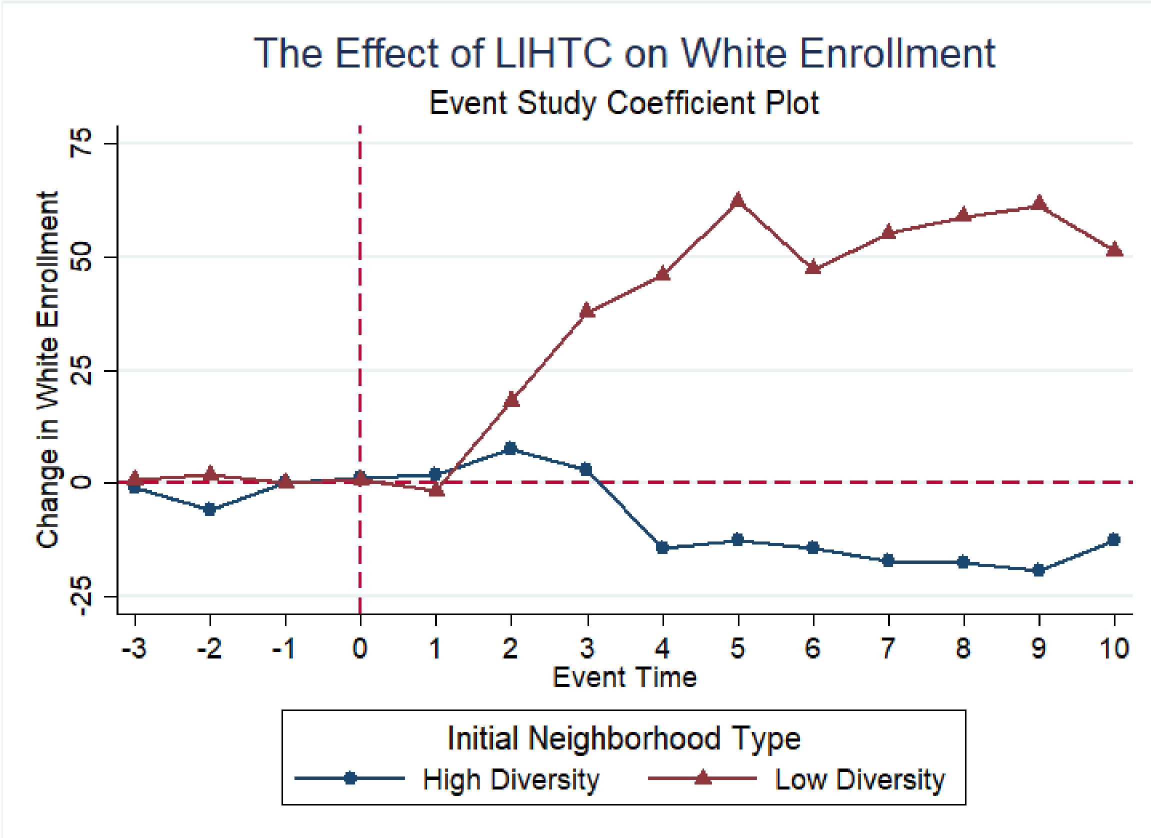


Figure 6: This diagram is a plot of the event study coefficients π_{gk} from regression equation 10 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

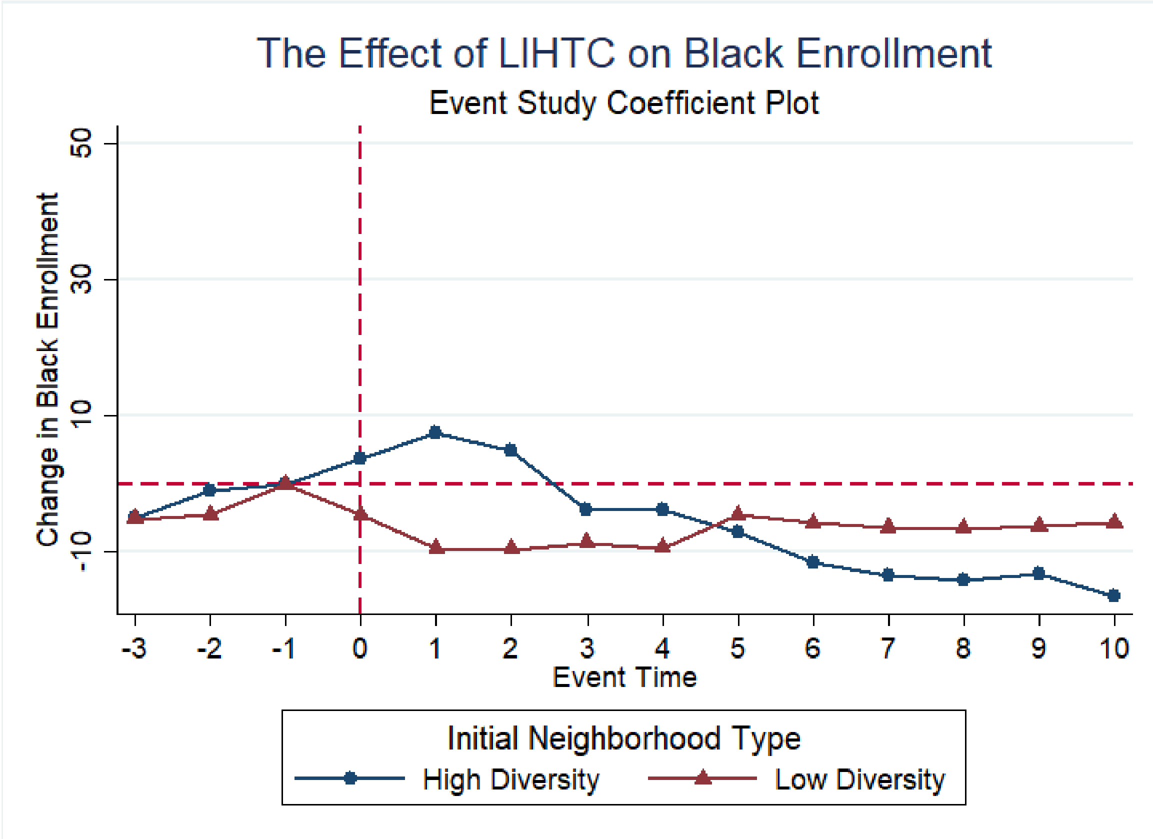


Figure 7: This diagram is a plot of the event study coefficients π_{gk} from regression equation 10 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

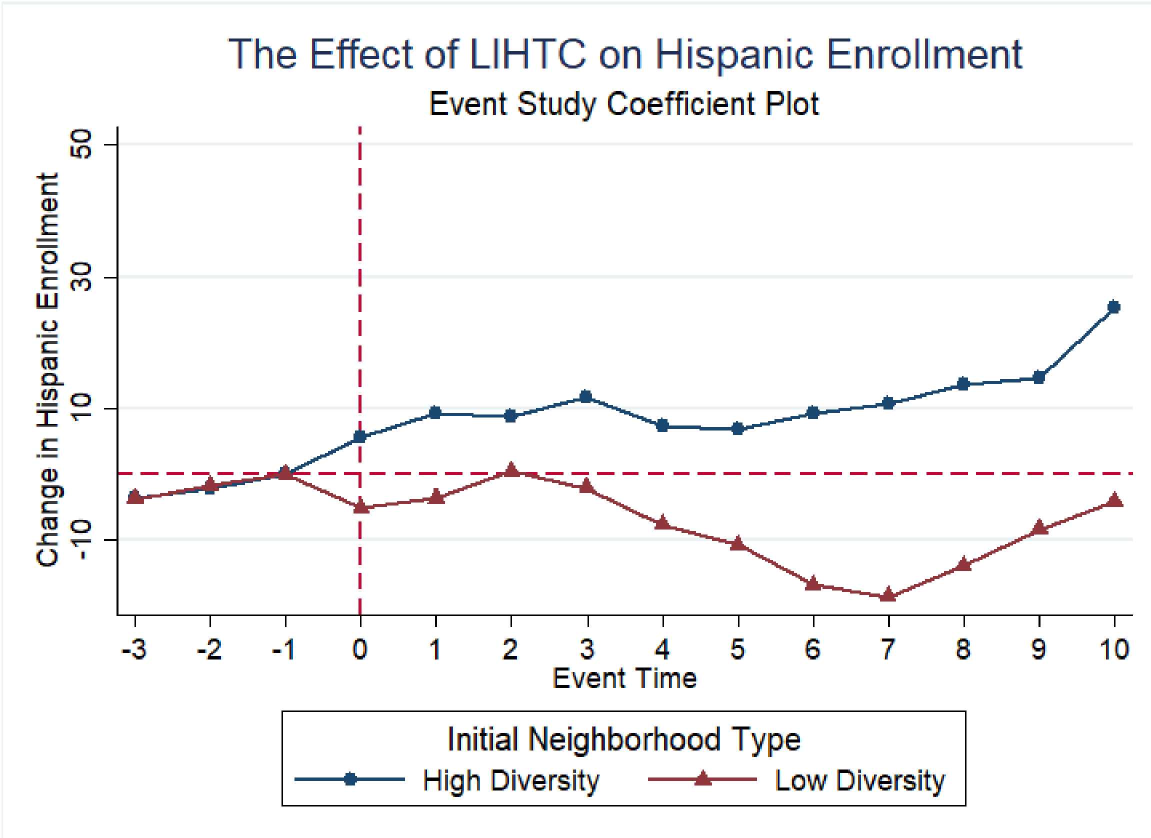


Figure 8: This diagram is a plot of the event study coefficients π_{gk} from regression equation 10 for g = 2, the high diversity group. The model predicts a decrease in diversity for these school types following LIHTC.

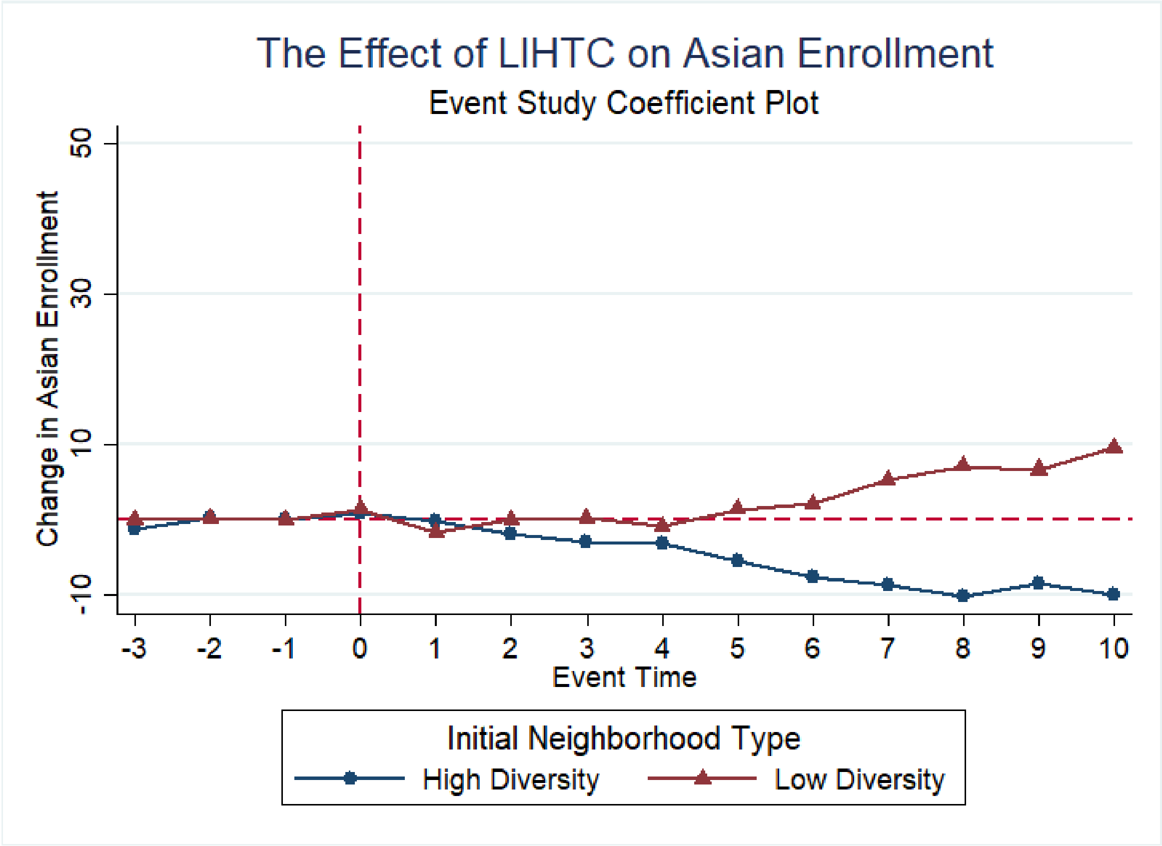


Figure 9: This diagram is a plot of the event study coefficients π_{gk} from regression equation 10 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

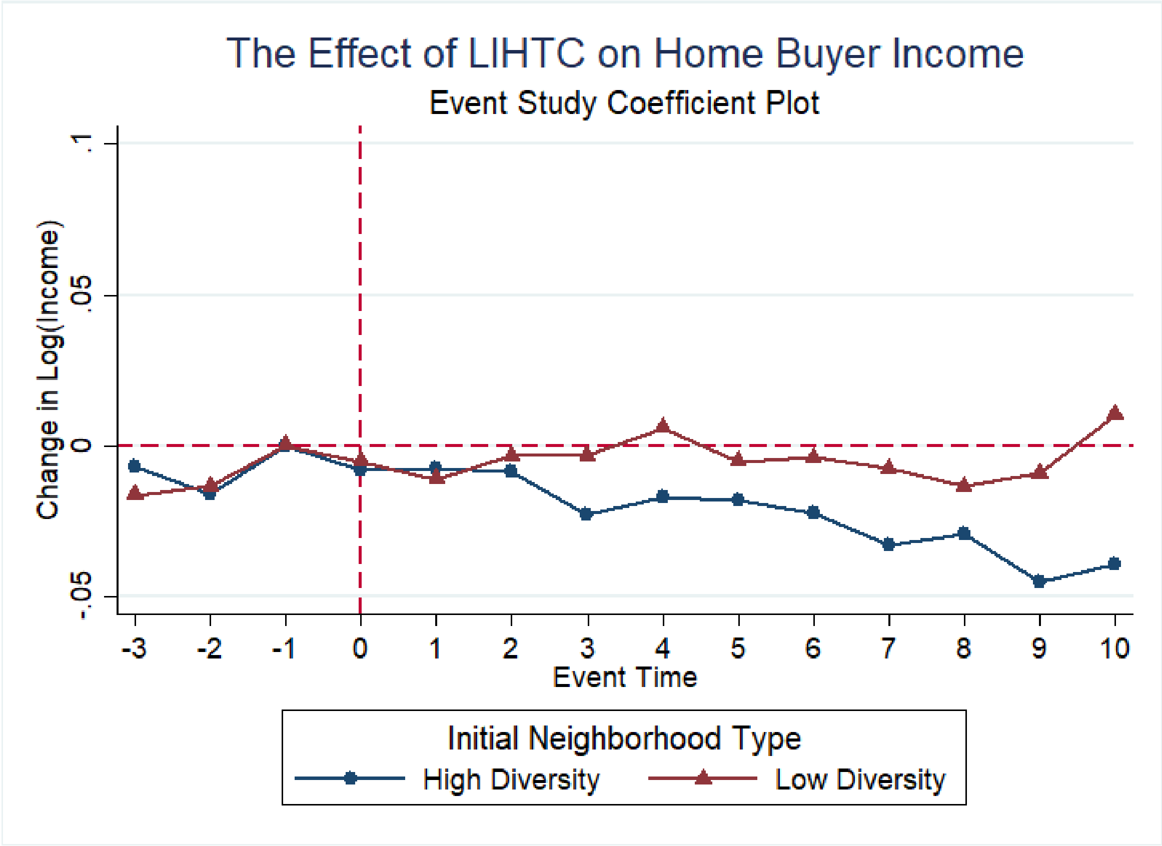


Figure 10: This is a plot of the event study coefficients from regressing log(applicant income) on the interacted event time dummies and the covariates included in first stage, for the initially low diversity school zones. For the low diversity school zones, the income of new home buyers did not change near the event time.

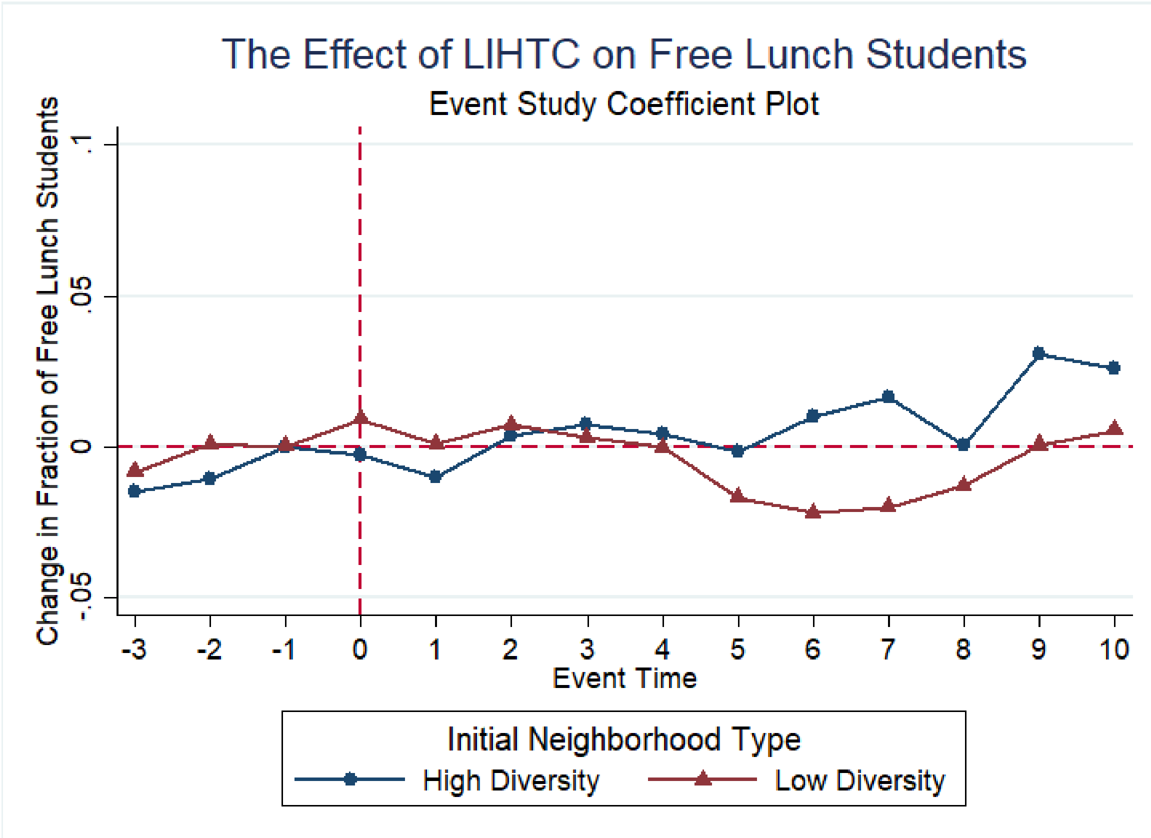


Figure 11: This is a plot of the event study coefficients from regressing log(applicant income) on the interacted event time dummies and the covariates included in first stage, for the initially high diversity school zones. For the high diversity school zones, the income of new home buyers did not change near the event time.

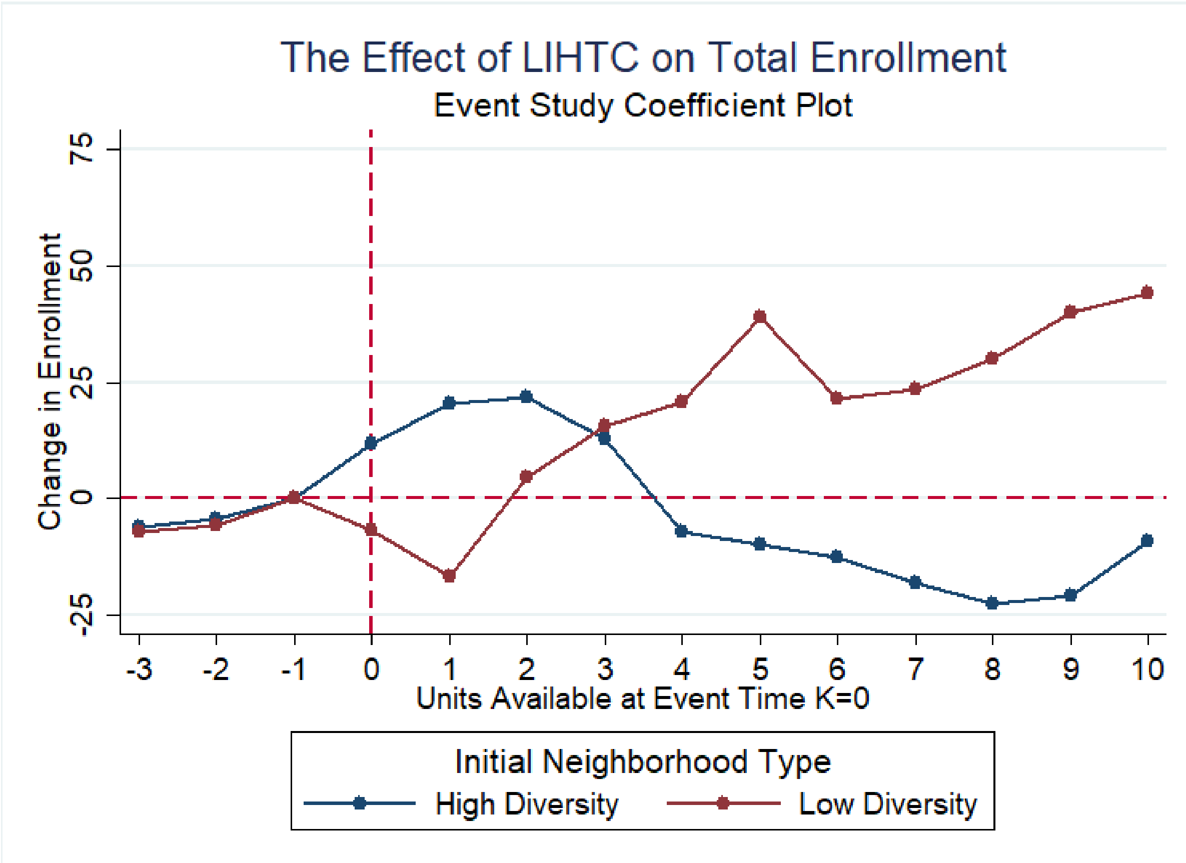


Figure 12: This diagram is a plot of the event study coefficients π_{gk} from regression equation 10 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

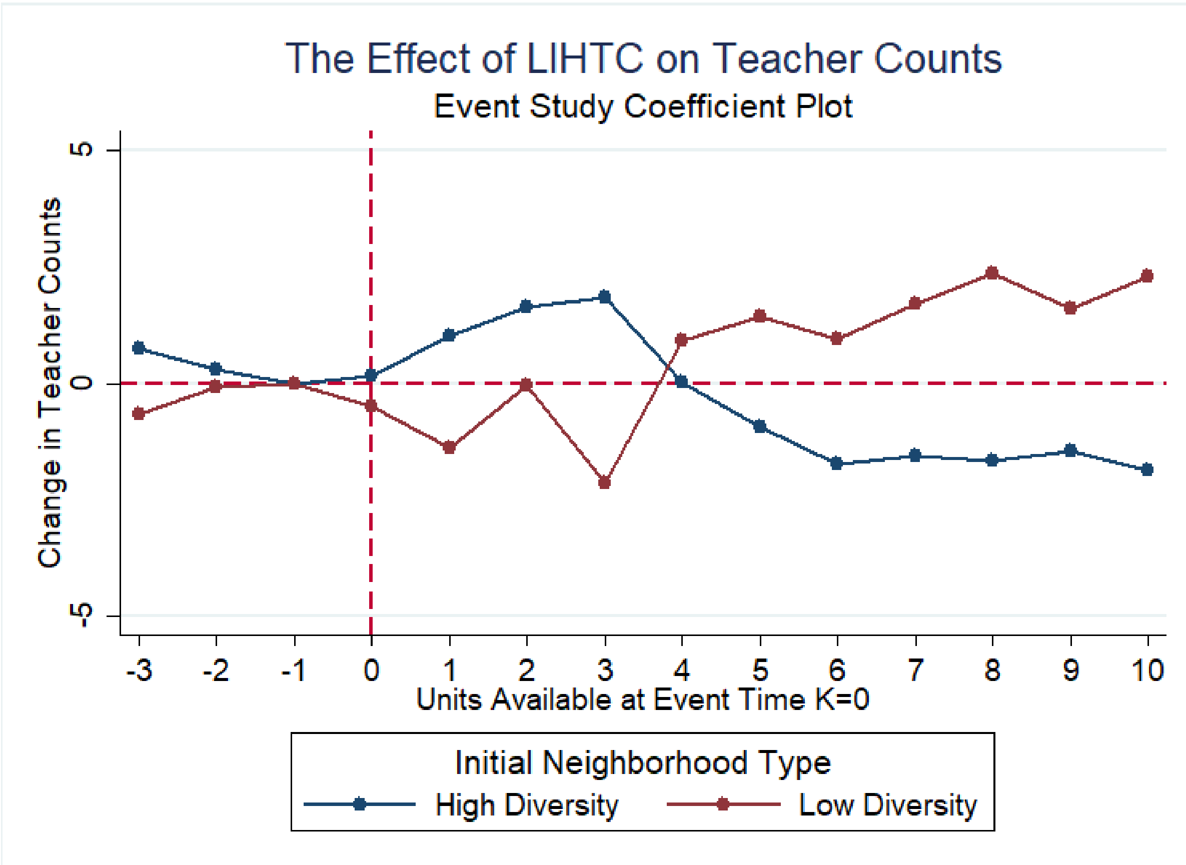


Figure 13: This diagram is a plot of the event study coefficients π_{gk} from regression equation 10 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.