Estimating the Value of Diverse Neighbors: The Relationship Between Public School Demographics and Local Mortgage Values

Kenneth Whaley
Clemson University*

November 4, 2019

Job Market Paper

Abstract

Home buyers observe neighborhood demographics along with a host of other attributes that affect willingness to pay for housing in a particular location. This paper presents empirical evidence that mortgage values for home purchases respond to changes in neighborhood racial composition. I link annual data describing the demographics of over 7,000 public schools to mortgage values for home purchases within the school attendance zone each year from 2000-2014. Constructing a measure of school diversity I find that mortgage values increase roughly 1% to 2.4% as neighborhoods become more diverse. A shift-share approach is employed to identify exogenous changes in neighborhood demographics associated with the timing of rental housing development under the Low-Income Housing Tax Credit (LIHTC) program. Categorizing neighborhoods based on pre-existing diversity levels, I find suggestive evidence that the effect of changes in neighborhood demographics on home values varies by neighborhood type. The results provide evidence for how much households value living near a diverse set of neighbors.

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

Home buyers observe the racial demographics of a neighborhood prior to choosing where to live. Along with touring neighborhoods and first-hand interactions with realtors, home search assistance websites like Zillow, Redfin, and others provide a low-cost way to learn demographic details of potential neighbors. In this paper I examine how neighborhood racial diversity affects mortgage values for the purchase of single-family homes, to understand household preferences for the demographic characteristics of the people that live nearby.

Households sort into neighborhoods for a variety of reasons commonly associated with neighborhood amenities – benefits attained from residing in a particular location, aside from physical housing. Economists have sought to understand resident demographics as a neighborhood amenity given persistent residential segregation by race and income over US history (Logan and Parman 2017). Classical theories of neighborhood choice beginning with Schelling (1971), argue that the observed patterns of household sorting over race reflect a range of preferences for the racial demographics of neighbors. The central motivation of this paper is to test empirically how preferences for diverse neighbors affect the willingness to pay for a home.

Using public school attendance zones as geographical neighborhood boundaries, I define racial diversity as the extent which the student body of the local school is distributed over several race and ethnic groups or comprised mostly of students from one background. I observe a balanced panel of over 7,000 neighborhoods from 2000-2014, and on average schools in my sample have become more racially diverse over this time period. I combine the school demographic data with publicly available data describing the mortgage value and location of newly purchased homes during the sample period, using the median mortgage value in a neighborhood as a proxy for the median home value. The panel nature of the data allows me to employ several empirical tests to disentangle the effect of changing racial demographics on local home values.

The identification challenge arises from correlation between resident demographics and other neighborhood amenities unobserved in the data and correlated with home values. In their seminal paper, Bayer et al. (2007) show that cross sectional differences in home values appear associated with race, but are largely a product of the relative quality of other amenities offered in minority versus white neighborhoods. Using micro data from San Francisco, the authors show that differences in home values on opposite sides of a school attendance boundary are explained by relative school quality and not racial differences. My empirical design is a two stage least squares model that exploits the panel nature of the data to relate plausibly exogenous changes in neighborhood diversity to changes in mortgage values. My results show on average mortgage values rise between 1% and 2.5% when schools and neighborhoods become more diverse.

I exploit differences in the timing of rental housing development under the Low Income Housing Tax Credit (LIHTC) program to identify quasi-random shifts in neighborhood demographics. One of the largest federal government housing programs, the goal of LIHTC is to increase the supply of housing by providing tax credits for developers to sell in exchange for working capital. In an exhaustive analysis of LIHTC spillovers, Diamond and McQuade (2019) suggest the introduction of diverse neighbors as a potential mechanism the rise in house prices following LIHTC in certain neighborhoods. Davis et al. (2018) model household preferences over neighborhood demographics to understand the long run effects of LIHTC on equilibrium neighborhood characteristics. I use LIHTC development as the demographic shifter in a broader analysis of neighborhood demographics and home values.

The first part of my estimation strategy is a flexible event study design that isolates the effect of LIHTC on the racial composition of schools in my sample. Schools in the sample are restricted to those with no LIHTC prior to 2000, and at most one new LIHTC development during the sample period. I find substantive changes in the enrollment count of white, black and Hispanic students following LIHTC. Using the model predicted enrollment growth for each demographic group, I test how neighborhood demographics affect prices via a shift-share instrument for neighborhood diversity that takes the timing of LIHTC shocks as random. My methodology borrows from Boustan (2010) and Derenoncourt (2018) who study various outcomes of demographic shifts during the Great Northern Migration from 1940-1970.

To approximate neighborhood diversity I construct an index that measures the extent to which the local school student body is comprised of one race or spread equally across five observable groups.[1] The initial distribution of the index in 2000 is bimodal, with a mass of neighborhoods in a low diversity equilibrium and another mass comprised of high diversity schools. The shape of the distribution aligns with the predicted sorting result of Schelling (1971) and more recently Card et al. (2008), who use decennial neighborhood data to empirically test the predictions of classical neighborhood choice models. Caetano and Maheshri (2017) extends sorting analysis using data from LA County schools to model the role of parental preferences in shaping the racial mix of students. The common theme of these papers is the existence of multiple equilibrium defined by the minority share of a given population. While my measure has a separate interpretation from minority share, I use the diversity levels from the year 2000 to distinguish neighborhood as low or high diversity, adding context to my empirical findings.

The initial neighborhood categorization allows me to test for heterogeneity throughout the analysis. The idea is that pre-existing demographic characteristics reflect a sorting equilibrium that signal the preferences of incumbent residents to potential home buyers. Thus, there is reason to believe the effect of increased diversity can be different by neighborhood type. Low diversity neighborhoods experience increases in white population following LIHTC, while high diversity neighborhoods experience increases in black and Hispanic populations following LIHTC. By tracing the effects over time relative to the

[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

event year, I show that the demographic changes are not entirely driven by the immediate influx of LIHTC residents. I find suggestive evidence that the effect of diversity on prices varies by neighborhood type.

One contribution of this paper is to expand the analysis of an inherently local problem to a national sample of neighborhoods. Observing 7,233 neighborhoods annually over 15 years provides statistical power for several tests less suited for decennial census microdata. Conducting the analysis across different neighborhood types allow me to frame the results in the context of theories on sorting and de facto neighborhood segregation. Cutler et al. (1999) posit that in the absence of laws upholding segregation, some white households will be willing to pay more to live in neighborhoods with higher shares of white residents. My estimates provided quantitative evidence for the magnitude of these types of theoretical predictions.

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). Facing a budget constraint and trading off between housing and the consumption of a composite numeraire good, households choose a neighborhood that maximizes utility over the level of housing and neighborhood attributes. 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.

For new home buyers, endogenously formed neighborhood racial composition becomes an amenity that affects willingness to pay for housing. The demographics of students in the local public school 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.

Local school demographics affect willingness to pay of households without school aged children. 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.[2] 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.

If endogenous sorting produces distinct neighborhood types, categorized by minority share, I hypothesize that exogenous changes to school and neighborhood racial demographics will change the price new home buyers will pay for housing. While sorting in theory represents heterogeneous preferences for the demographics of neighbors, I posit

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

that the equilibrium outcomes also serve as a signal for potential incoming residents. Cutler et al. (1999) assert that white households will pay more for majority white neighborhoods in the absence of laws upholding racial segregation. Home values in majority white neighborhoods could then fall as the population becomes more diverse. An alternative explanation lies in search costs for households choosing where to live. The long-term nature of housing investment implies that new home buyers demand neighborhoods in a stable equilibrium of either type, purely to reduce uncertainty around future home price appreciation. Exogenous shocks to racial composition may signal a neighborhood in transition, increasing uncertainty and decreasing the demand for owner occupied housing.

I test empirically the hypothesis that changes to racial composition of a neighborhood will affect prices by first constructing a measure of school and neighborhood diversity in section 2.1. I discuss the LIHTC program and the assumptions required for new construction to serve as an exogenous shifter to neighborhood diversity in section 5. Diamond and McQuade (2019) argue that LIHTC diversifies neighborhoods in a way that increases home prices in low income areas and decreases prices in high income areas. My analysis instead focuses on heterogenous effects brought about by the initial racial sorting equilibrium of neighborhoods and, to further understand how deviation from these initial states affect the willingness to pay of new home buyers.

2.1 A Measure of School and Neighborhood Diversity, SDI_{it}

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.[3] Using a measure of concentration

[3]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.

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,

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

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 to SDI_{it} in section 5.

2.2 Neighborhood Data

I observe racial demographics, and thus SDI_{it}, for a balanced panel of 7,233 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.

2.3 Neighborhood Type and the Distribution of SDI_{it}

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 neighborhoods follow a bimodal distribution that supports long run equilibrium theories of neighborhood composition discussed in the beginning of section 2. Given the empirical distribution of school demographics I categorize neighborhoods into two groups. Neighborhoods in the left peak are low diversity and those on the right peak are high diversity, with some neighborhoods 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 formally categorize neighborhoods, I assume the observed distribution of SDI_{it} is a combination of the two normally distributed equilibrium types. The true neighborhood 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 neighborhood 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 neighborhoods in the low diversity group. The mean and standard deviation of SDI for the low and high diversity groups 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 neighborhood 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 neighborhood i is a low or high diversity type in the year 2000. Since the neighborhood type is a binary outcome, the model predicts the logit probabilities of each neighborhood being categorized into either the low or high diversity underlying densities. A given neighborhood 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 neighborhood 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 7,233 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 The Relationship Between Demographics and Prices

There are three parts to my estimation strategy. First, I estimate the effect of higher levels of neighborhood diversity on mortgage values using a fixed effects model that accounts for time invariant neighborhood characteristics influencing this relationship. 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 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.

3.1 Neighborhood 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 2.2 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.[4]

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 $228,298 and the median home buyer earns roughly $85,232 annually. Mortgage values in high diversity neighborhoods are 8% higher than low diversity neighborhoods, with essentially no difference in home buyer incomes.

Figure 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 probably 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 neighborhood level 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.

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

3.2 Fixed Effects Estimation

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

$$Log(V_{it}) = \alpha_1 SDI_{it} + \alpha_2 SDI_{it} \times I(LowDiversity = 1) + X_{it}^{'}\beta_1 + \gamma_i + \gamma_t + \epsilon_{it}, \tag{6}$$

where V_{it} is the median mortgage value in neighborhood i year t. Both neighborhood (γ_i) and year (γ_t) fixed effects are included in the regression along with a vector of neighborhood covariates X_{it}. Descriptive neighborhood covariates include the two income measures, total enrollment counts, and teacher counts. I estimate the average effect of diversity and allow for heterogeneous effects by neighborhood type. Neighborhood choice models suggest the neighborhood types observed in the data reflect preferences for low or high levels of diversity, and thus the empirical effect of diversity on mortgages could vary across neighborhood type.

Estimates of α_1 and α_2 from equation 6 are presented in table 2. Each specification other than model one includes neighborhood and year fixed effects, and columns four and five include the estimates for the interaction term in equation 6. Across specifications I vary the inclusion of the neighborhood income measures, to test how the relationship between diversity levels and mortgage values is explained by neighborhood income differences. The estimated effect of model one is essentially the partial slope of the correlation between SDI_{it} and mortgage values show in figure 4. Comparing the results of the fixed effects models, the effects 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. The correlation between neighborhood diversity and prices is significantly weaker when neighborhood level fixed effects are included, thus when schooling differences associated with racial demographics are held constant the effect of neighborhood diversity disappears. Eliminating cross sectional differences between neighborhood diversity and mortgage values is an advantage of fixed effects estimation but without exogenous, time varying changes to neighborhood 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 2012 survey also shows that 38% of households renting in the program have children under 18. If there is no discernible difference in the racial composition of local LIHTC residents with children and those without, then the change in SDI_{it} following LIHTC development will reflect changes in both the school and neighborhood. I provide support for the empirical assumptions necessary to describe LIHTC as an exogenous shifter to SDI_{it} in section 5. Though the primary motivation of this paper is the relationship between SDI_{it} and mortgage values, establishing demographic diversity as the primary channel in which LIHTC affects neighborhood prices is an additional contribution. I rule out income changes correlated with the timing of LIHTC as the source of identification in section 5.

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 7,238 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 1,520 projects were put into service. 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 X_{it} as a function of exogenous exposure to a LIHTC shock, Z_{it} such that

$$X_{it} = Z_{it}'\pi + \gamma_i + \gamma_t + \xi_{it}. \tag{7}$$

The coefficient π in equation 7 is an estimate of the average effect of new LIHTC development on neighborhood diversity, conditional on neighborhood and year fixed effects γ. To capture the variation in outcome X_{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

$$Z_{it} = \sum_{g=1}^{2} \sum_{k=-6}^{10} (I_{ig} \times D_{ik}). \tag{8}$$

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 X_{it} response as predicted by the pre-treatment neighborhood type. Substituting equation 8 into equation 7, I express the event study model as

$$X_{it} = \sum_{g=1}^{2} \sum_{k=-6}^{10} (I_{ig} \times D_{ik})\pi_{gk} + \gamma_i + \gamma_t + \xi_{it}. \tag{9}$$

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 X_{it}. Included in my estimation of equation 9 are census division specific linear time trends to capture general demographic shifts unique to different regions of the country.

I estimate equation 9 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 X_{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. Figure 10 shows that home buyer income rises briefly in low diversity areas but falls in the long run, and only falls in the long run for low diversity areas (with no initial increase). The pre-trends in 11 are not parallel thus it is difficult to make a claim about the immediate effect of LIHTC on free lunch students, but the long run effect in low diversity areas is a decrease in the number of subsidized lunch students.

5 Instrumenting for Changes to Neighborhood Diversity

To estimate the causal effect of demographic shifts on neighborhood mortgage values, I construct a shift share instrument for SDI_{it} based on predicted population changes shown in section 4.2. Starting with the observed population 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 neighborhood population 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

$$\hat{g}_{rit} = \frac{\hat{n}_{rit} - \hat{n}_{rit-1}}{\hat{n}_{rit-1}}, \tag{10}$$

where $\hat{n}_{rit}$ is the fitted value population count generated by model 9. 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

$$\widetilde{n}_{rit} = n_{ri00} \times (1 + \hat{g}_{rit}). \tag{11}$$

For each year to follow, we have

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

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

$$\widetilde{s}_{rit} = \frac{\widetilde{n}_{rit}}{\sum_{r=1}^{5} \widetilde{n}_{rit}}. \tag{13}$$

I then calculate $\widetilde{SDI}_{it}$ using the growth adjusted shares in 13 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

$$SDI_{it} = \alpha_1 \widetilde{SDI}_{it} + X_{it}^{'}\beta_1 + \gamma_i + \gamma_t + \xi_{it}, \tag{14}$$

and the main estimating equation takes the form

$$Log(V_{it}) = \alpha_2 SDI_{it} + X_{it}^{'}\beta_2 + \gamma_i + \gamma_t + \epsilon_{it}. \tag{15}$$

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 land supply in rural areas (the excluded category) is perfectly elastic but varies in a way that is common across the three other geographical groups.

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. On average, mortgage values in a neighborhood increase as the area becomes more diverse by nearly 2.4%. Comparing column 1 to columns 2 and 3, the effect is partly explained by changes in neighborhood income associated with higher levels of diversity. In columns 4 and 5 I allow for the effect to differ between high and low diversity neighborhoods. Controlling for income changes I find no statistically significant difference in the effect of diversity on mortgage values between neighborhood type. Controlling for income, increases in neighborhood diversity on average increase mortgage values by roughly 1%.

To frame the magnitude of these estimates I present the reduced form results in table

4. Compared to the 2SLS results across specifications I find the reduced form results are nearly twice the magnitude of the 2SLS results. In fact these estimates are of order comparable to the prior work describing the net effects of LIHTC on neighborhood prices such as Baum-Snow and Marion (2009) and Diamond and McQuade (2019), which claim home prices could increase up to 5% following subsidized housing development, depending on the income characteristics of incumbent residents. Thus what my 2SLS estimates are identifying is the treatment effect of LIHTC on home prices specific to changes in local demographics. My conclusion from this comparison is that racial composition changes explain about half of the overall effect on prices following LIHTC construction. In the sections to follow, I test how different incumbent neighborhood types respond to this demographic shift.

5.2 The Effect of SDI_{it} on Mortgages, by Incumbent Racial Majority

The categorization of neighborhood type by SDI_{it} is broad enough that the effect of diversity may differ whether neighborhood residents are majority white, black or Hispanic. In this section I estimate the effects on subsets of the sample as defined by majority group. Note that a neighborhood can be classified as low or high diversity even though a specific demographic group represents over 50% of the population. Recall from section 2.3 that 78% (5,642) of the total sample are majority white neighborhoods, 7% (506) are majority black and 8% (579) are majority Hispanic. The small sample of black and Hispanic neighborhoods limits the power of my model, however I form general conclusions based on the sign of the results.

I find that substantial differences in the effect of SDI_{it} on mortgages in this set of regressions. The OLS results in table 5 column 1 suggest that low diversity, majority white neighborhoods do not observe a statistically positive increase in mortgage values when SDI_{it} rises. However the OLS regressions in table 2 suggest that the positive effect of diversity on mortgage values will be understated by this model. The results in column

6 support this narrative. There is a net positive effect of increases in SDI_{it} in majority white neighborhoods of each type, however the effect is roughly 1% larger when the neighborhood is already diverse.

Table 6 shows the effect for majority black neighborhoods. Given the propensity for these neighborhood types to exist near urban cores, one can imagine that low diversity, majority black neighborhoods are those ripe for gentrification following LIHTC. Although the results of this analysis are not statistically significant, the magnitude of the effect in column 6 indicate the potential for the largest positive mortgage value increases as diversity levels rise.

6 Conclusion

In this paper I offer empirical tests theoretical predictions of neighborhood choice models. Using panel data describing neighborhood demographic characteristics I find that mortgage values rise as neighborhoods become more diverse. I exploit the quasi-random timing of subsidized housing development under the LIHTC program to generate demographic population shifts, allowing for initial neighborhood diversity levels to predict the household sorting response. Testing the model over various subsets of the data, I find suggestive evidence that the effect varies by neighborhood type. These results imply heterogeneity in the preferences for diverse neighbors and imply that the way households assemble across neighborhoods reflects underlying preferences for local racial diversity.

References

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

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.

BOUSTAN, L. P. (2010): "Was postwar suburbanization white flight? 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.

CARD, D., A. MAS, AND J. ROTHSTEIN (2008): "Tipping and the Dynamics of Segregation," *The Quarterly Journal of Economics*, 123, 177–218.

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. 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.

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

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.

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 ($2014)	228,298.50	218,821.00	237,171.07
	(119216.2)	(108224.0)	(128041.5)
Real home buyer income	85,232.13	84,921.95	85,522.52
	(37705.6)	(36364.7)	(38922.5)
School Diversity Index, 2000	4.64	1.59	6.19
	(2.564)	(0.934)	(1.537)
School Diversity Index, 2014	5.96	4.66	7.18
	(2.415)	(2.440)	(1.627)
Fraction Black	0.15	0.11	0.19
	(0.196)	(0.201)	(0.185)
Fraction Hispanic	0.24	0.16	0.31
	(0.232)	(0.204)	(0.232)
Fraction Asian	0.06	0.05	0.07
	(0.0907)	(0.0826)	(0.0964)
School Enrollment	804.46	769.97	836.75
	(361.5)	(383.1)	(336.8)
Fraction Free Lunch	0.40	0.32	0.47
	(0.251)	(0.243)	(0.238)
Neighborhoods with LIHTC	0.25	0.23	0.27
	(0.432)	(0.420)	(0.442)
LIHTC units per build	288.28	278.98	294.99
	(168.62)	(165.92)	(170.89)
Neighborhood housing units	9,703.09	8,315.91	11,001.73
	(7830.2)	(5983.9)	(9039.9)
Middle Schools	7233	3585	3648
School$\times$Year Observations	108495	53775	54720
Share of Total LIHTC		0.48	0.52

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: The Effect of Neighborhood Diversity on Mortgage Values

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
SDI_it	0.0318***	0.00225	-0.000364	-0.00294*	0.00166	0.000521
	(0.00278)	(0.00260)	(0.00259)	(0.00124)	(0.00330)	(0.00153)
$SDI\_it * LowDiversity$					-0.00462	-0.00848***
					(0.00394)	(0.00225)
N	93771	93770	92056	87218	92056	87218
r2	0.0312	0.918	0.919	0.963	0.919	0.963
Fixed Effects		×	×	×	×	×

Notes: Standard errors in parentheses are clustered at the neighborhood level. The dependent variable in each regression is the log of neighborhood mortgage values, and models 2-6 include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: 2SLS Effect of Neighborhood Diversity on Mortgage Values

	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
SDI_{it}	0.0244***	0.0191**	0.00991**	0.0189**	0.00999**
	(0.00641)	(0.00644)	(0.00309)	(0.00652)	(0.00313)
$SDI_{it} * LowDiversity$				0.0160*	-0.000868
				(0.00771)	(0.00471)
N	85287	83765	79410	83765	79410
r2	0.0553	0.0805	0.532	0.0606	0.533
Fstat	713.7	680.9	743.5	343.5	374.9
Neighborhood Covariates		×	×	×	×
Neighborhood Income Controls			×		×

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and models 2-5 include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Reduced Form Effect of Neighborhood Diversity on Mortgage Values

Reduced Form	(1)	(2)	(3)	(4)	(5)
SDI_{it}	0.0489***	0.0370**	0.0204**	0.0239	0.0202**
	(0.0121)	(0.0121)	(0.00621)	(0.0128)	(0.00663)
$SDI_{it} * LowDiversity$				0.0740**	0.00115
				(0.0259)	(0.0140)
N	86016	83792	79410	83792	79410
r2	0.917	0.918	0.962	0.918	0.962
Neighborhood Covariates		×	×	×	×
Neighborhood Income Controls			×		×

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and models 2-5 include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: The Effect of Neighborhood Diversity on Mortgage Values: White Neighborhoods

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
SDI_{it}	0.0106***	0.117***	0.108***	0.0528***	0.106***	0.0522***
	(0.00204)	(0.0181)	(0.0179)	(0.00766)	(0.0155)	(0.00733)
$SDI_{it} * LowDiversity$	-0.0161***				-0.00470	-0.0136*
	(0.00243)				(0.00859)	(0.00542)
N	66731	63450	62411	58923	62411	58923
r2	0.964	0.273	0.235	0.454	0.211	0.469
Fstat		151.2	145.6	184.1	82.28	99.54
Neighborhood Covariates	×		×	×	×	×
Income Controls	×			×		×

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and all models include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: The Effect of Neighborhood Diversity on Mortgage Values: Black Neighborhoods

	(1) OLS	(2) 2LS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
SDI_{it}	-0.00461	-0.0678	-0.0175	0.0286	-0.301	-0.0776
	(0.00568)	(0.144)	(0.127)	(0.0598)	(0.607)	(0.139)
$SDI_{it} * LowDiversity$	0.0158*				0.560	0.210
	(0.00764)				(0.822)	(0.206)
N	6656	7011	6952	6656	6952	6656
r2	0.946	0.121	0.00154	0.516	0.565	0.346
Fstat		2.285	2.719	3.803	3.352	4.896
Neighborhood Covariates	×		×	×	×	×
Income Controls	×			×		×

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and all models include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: The Effect of Neighborhood Diversity on Mortgage Values: Hispanic Neighborhoods

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
SDI_{it}	0.00855***	-0.136	-0.124	-0.00748	-0.494	-0.0304
	(0.00202)	(0.0768)	(0.0658)	(0.0353)	(2.891)	(0.0603)
$SDI_{it} * LowDiversity$	-0.0166**				-7.729	-0.372
	(0.00526)				(55.20)	(0.415)
N	7165	7679	7479	7165	7479	7165
r2	0.845	0.398	0.324	0.604	0.45	0.281
Fstat		8.544	12.88	9.831	15.91	14.67
Neighborhood Covariates	×		×	×	×	×
Income Controls	×			×		×

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and all models include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: The Effect of Neighborhood Diversity on Mortgage Values: Low Income Neighborhoods

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
SDI_{it}	-0.00210	0.0372*	0.0375*	0.0112	0.0438**	0.0127
	(0.00601)	(0.0157)	(0.0158)	(0.0104)	(0.0145)	(0.00915)
$SDI_{it} * LowDiversity$					-0.0418	-0.00891
					(0.0449)	(0.0314)
N	17524	17218	16928	15991	16928	15991
r2	0.918	0.0162	0.0264	0.441	0.00412	0.441
Fstat		109.6	113.6	93.47	62.14	50.36

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and all models include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: The Effect of Neighborhood Diversity on Mortgage Values: Middle Income Neighborhoods

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
SDI_{it}	-0.00269	0.00848	0.00231	0.0105*	-0.00139	0.00773
	(0.00199)	(0.00903)	(0.00917)	(0.00446)	(0.00944)	(0.00469)
$SDI_{it} * LowDiversity$					0.0513***	0.0226**
					(0.0137)	(0.00869)
N	44588	44103	43275	40958	43275	40958
r2	0.951	0.0111	0.0321	0.531	0.0241	0.524
Fstat		341.1	312.7	309.4	156.3	155.6

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and all models include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: The Effect of Neighborhood Diversity on Mortgage Values: High Income Neighborhoods

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
SDI_{it}	-0.00428	0.0401**	0.0335*	0.00942	0.0329*	0.00982
	(0.00223)	(0.0136)	(0.0136)	(0.00564)	(0.0130)	(0.00558)
$SDI_{it} * LowDiversity$					-0.00258	-0.0107
					(0.00936)	(0.00620)
N	23914	22880	22502	21444	22502	21444
r2	0.967	0.0611	0.0394	0.566	0.0348	0.569
Fstat		165.2	159.2	220.8	90.35	114.5
Neighborhood Covariates	×		×	×	×	×
Income Controls	×			×		×

Notes: Standard errors in parentheses are clustered at the neighborhood level. All models include neighborhood by year fixed effects. The dependent variable in each regression is the log of neighborhood mortgage values, and all models include controls for locale (large urban, small urban, large suburban) interacted with a post-treatment dummy to control for heterogenous effects of construction housing dynamics. Kleibergen-Paap F statistic for weak instruments is shown for each of the IV regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figures

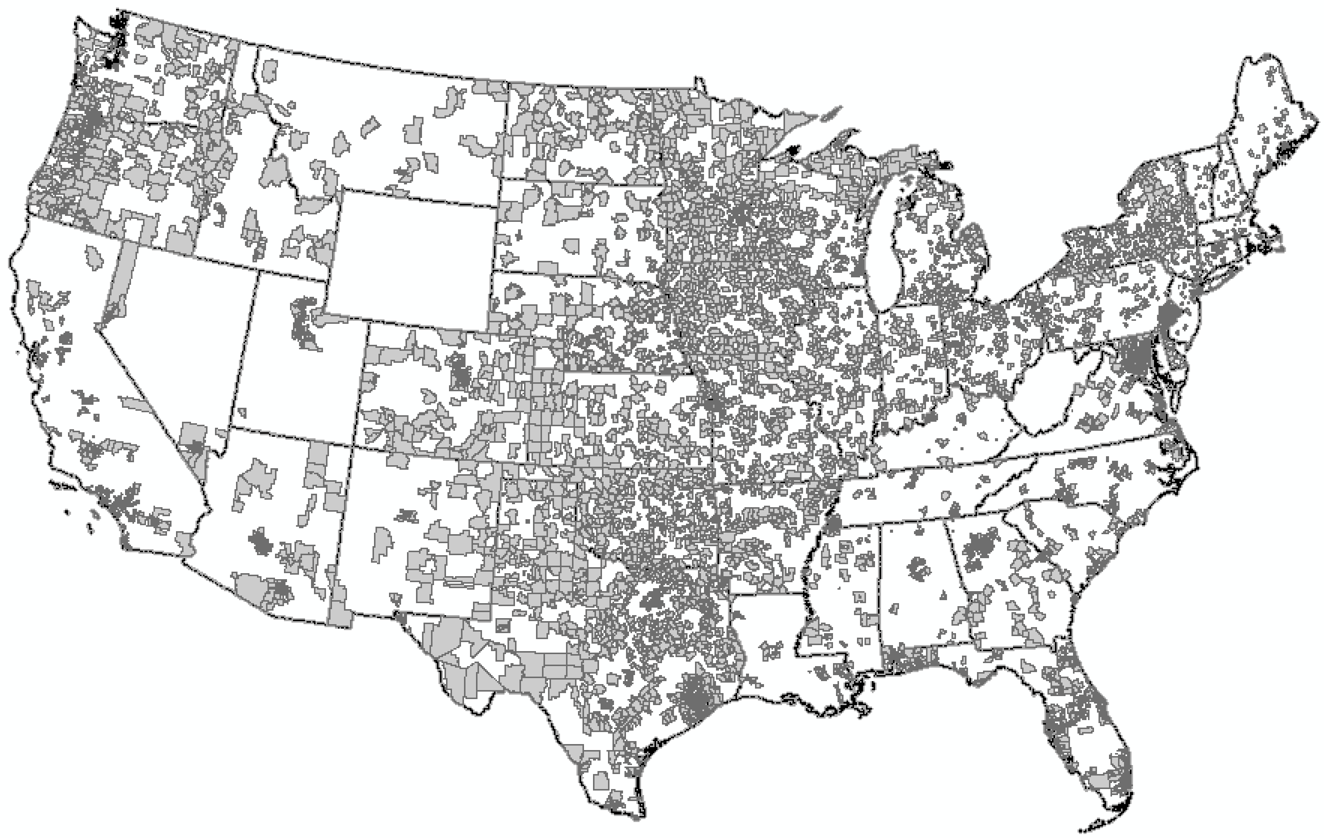
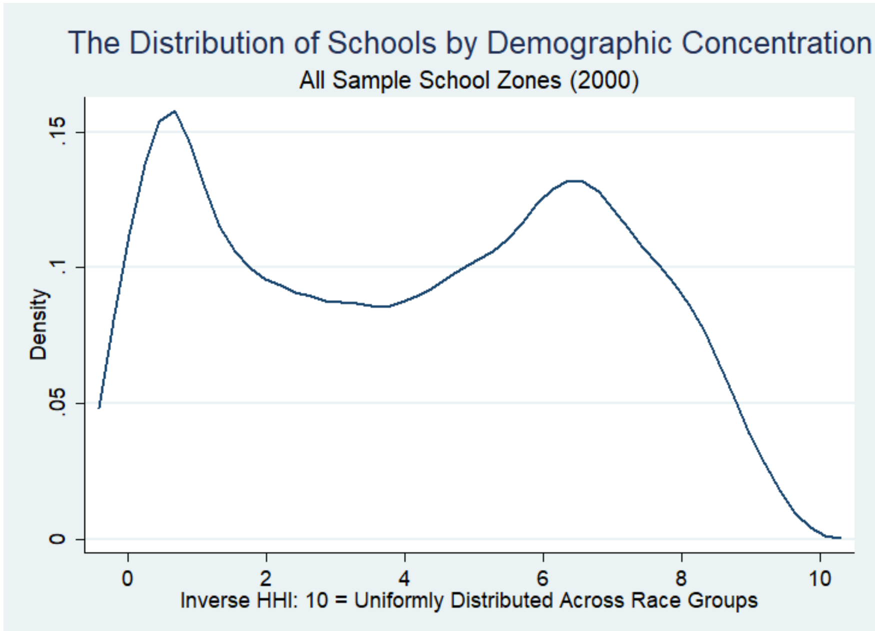
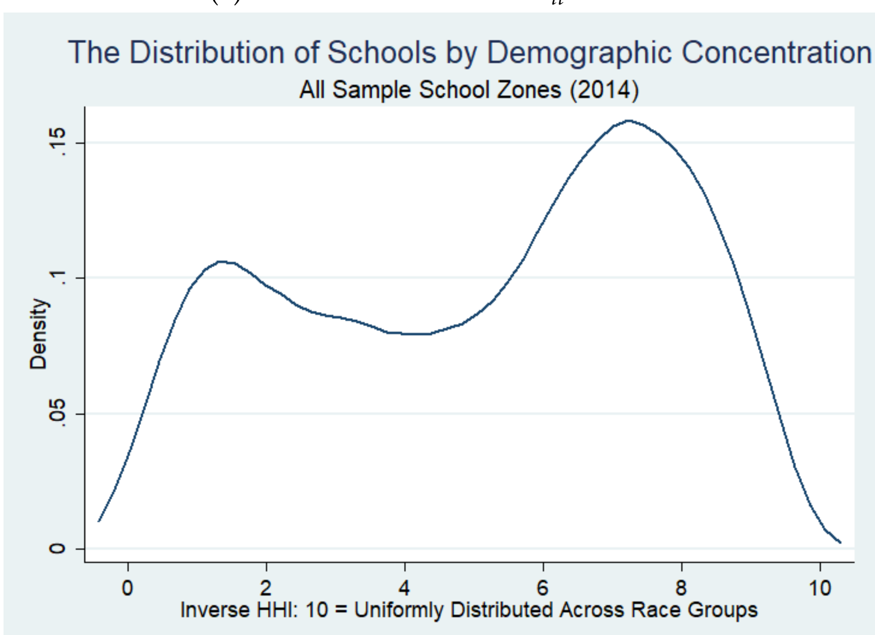


Figure 1: The shaded grey areas are middle school zones used as neighborhoods in my sample. Boundaries are fixed to 2009 attendance zones publicly available at https://www.sabinsdata.org/.

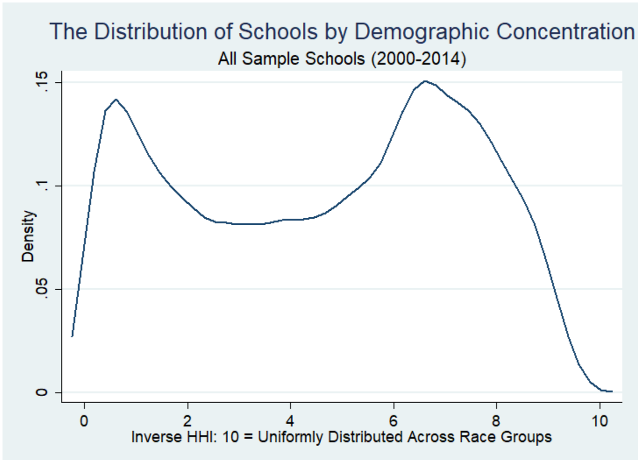


(a) The distribution of SDI_{it} in 2000.

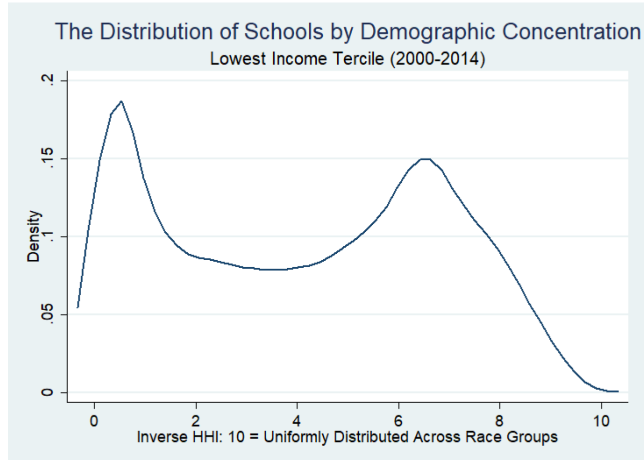


(b) The distribution of SDI_{it} in 2014.

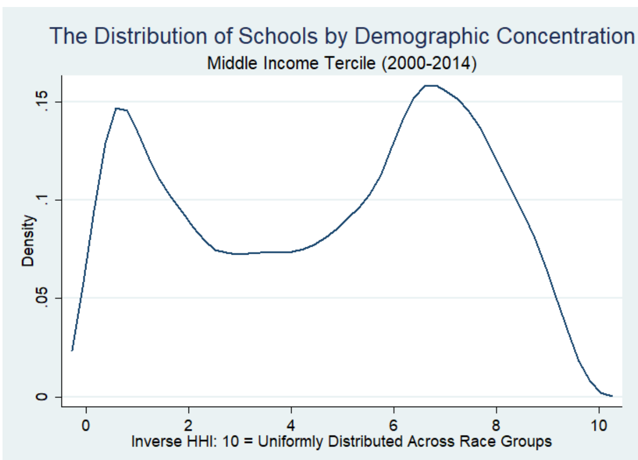
Figure 2: The distribution of schools by racial diversity as measured by SDI_{it}. Schools have generally become more diverse from 2000 to 2014 as evidenced by a larger mass on the right half of the distribution.



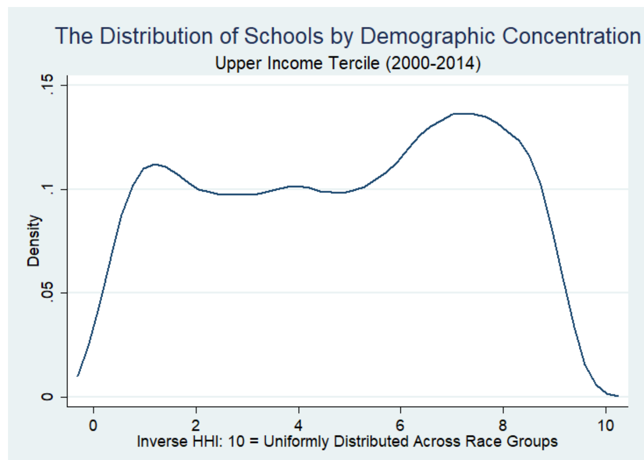
(a) Entire Sample



(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.

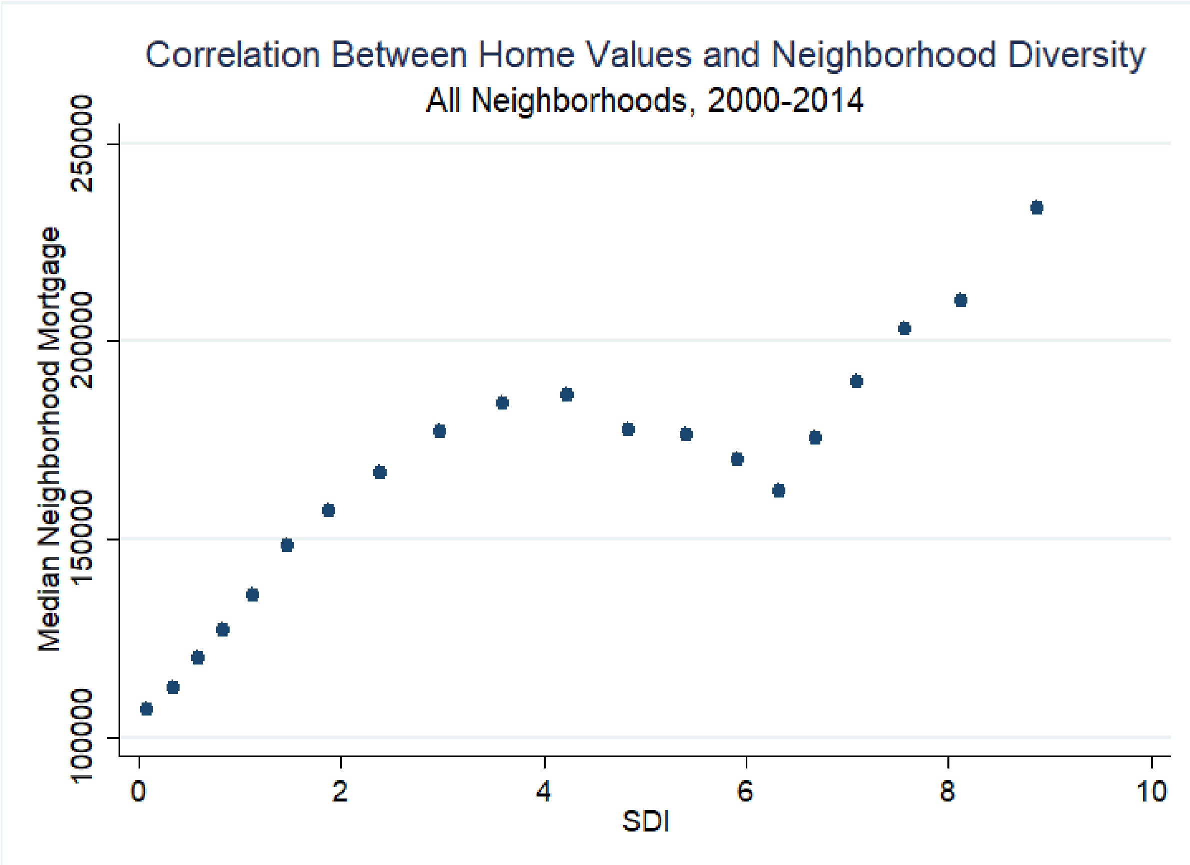


Figure 4: A binned scatter plot of the raw data shows a positive unconditional correlation between neighborhood diversity and mortgage values.

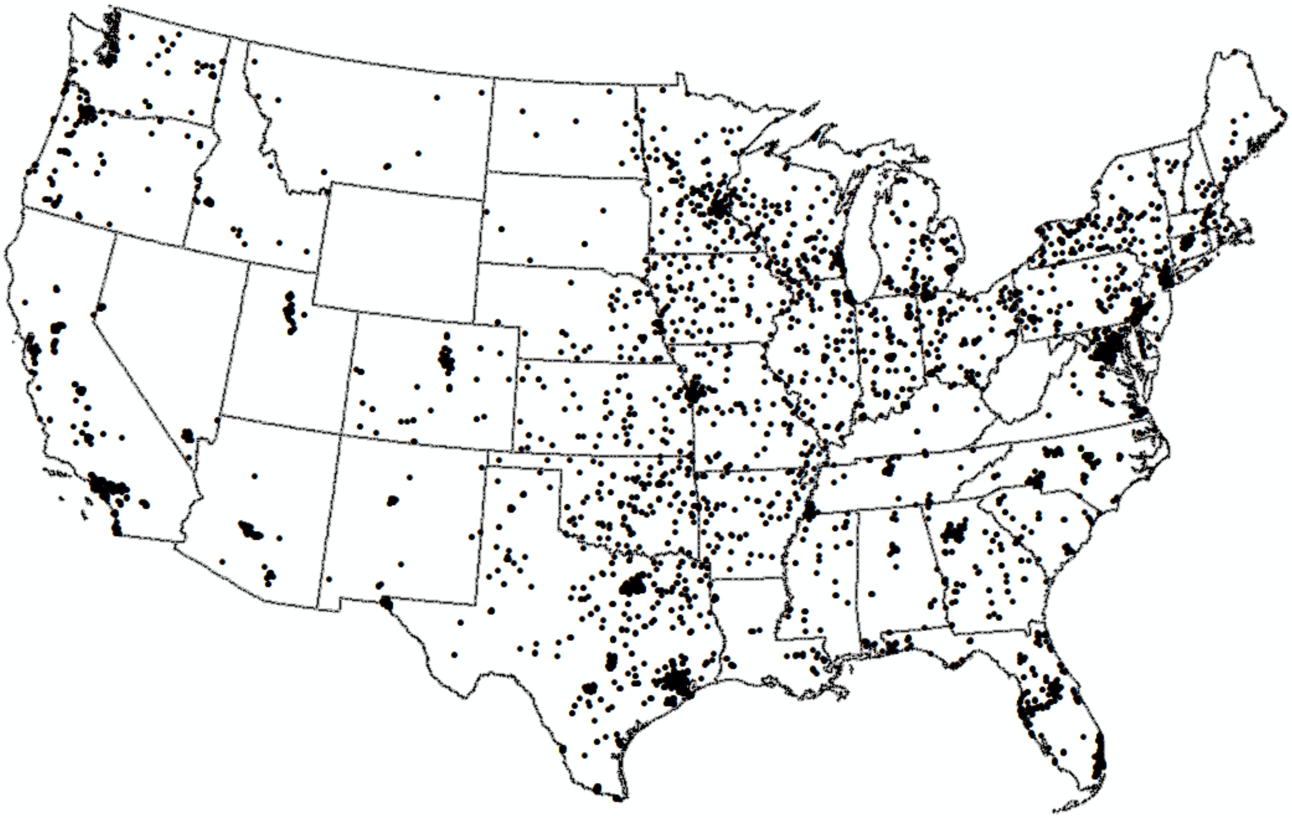


Figure 5: The dots represent individual LIHTC developments from 2000-2014 in my sample middle school zones.

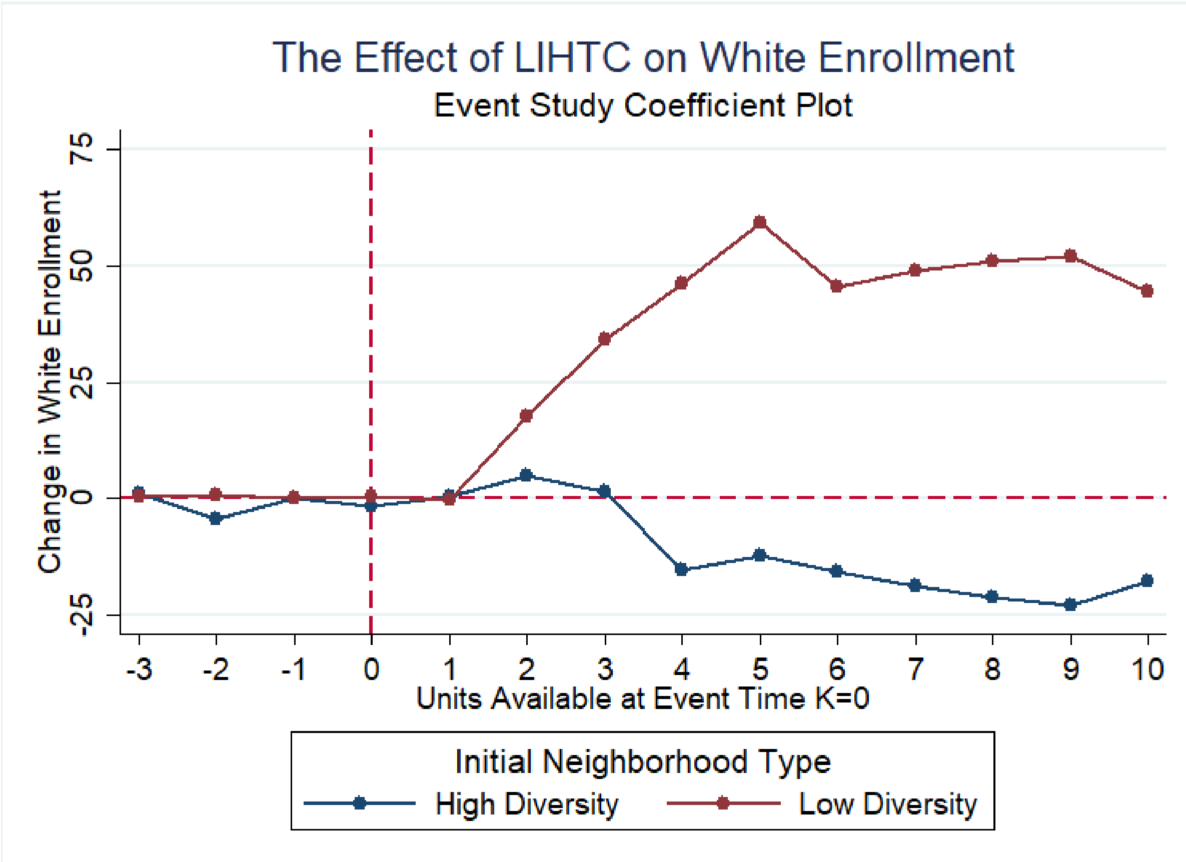


Figure 6: This diagram is a plot of the event study coefficients π_{gk} from regression equation 9 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

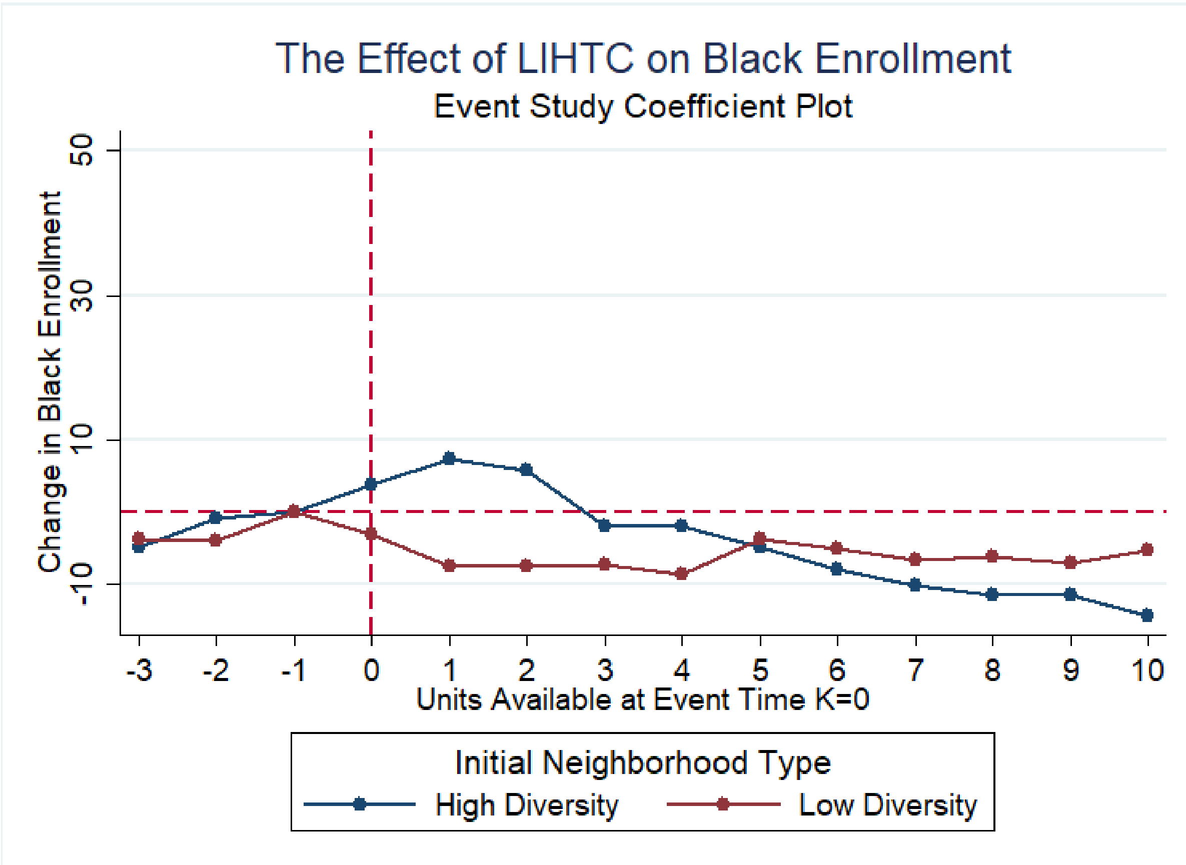


Figure 7: This diagram is a plot of the event study coefficients π_{gk} from regression equation 9 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

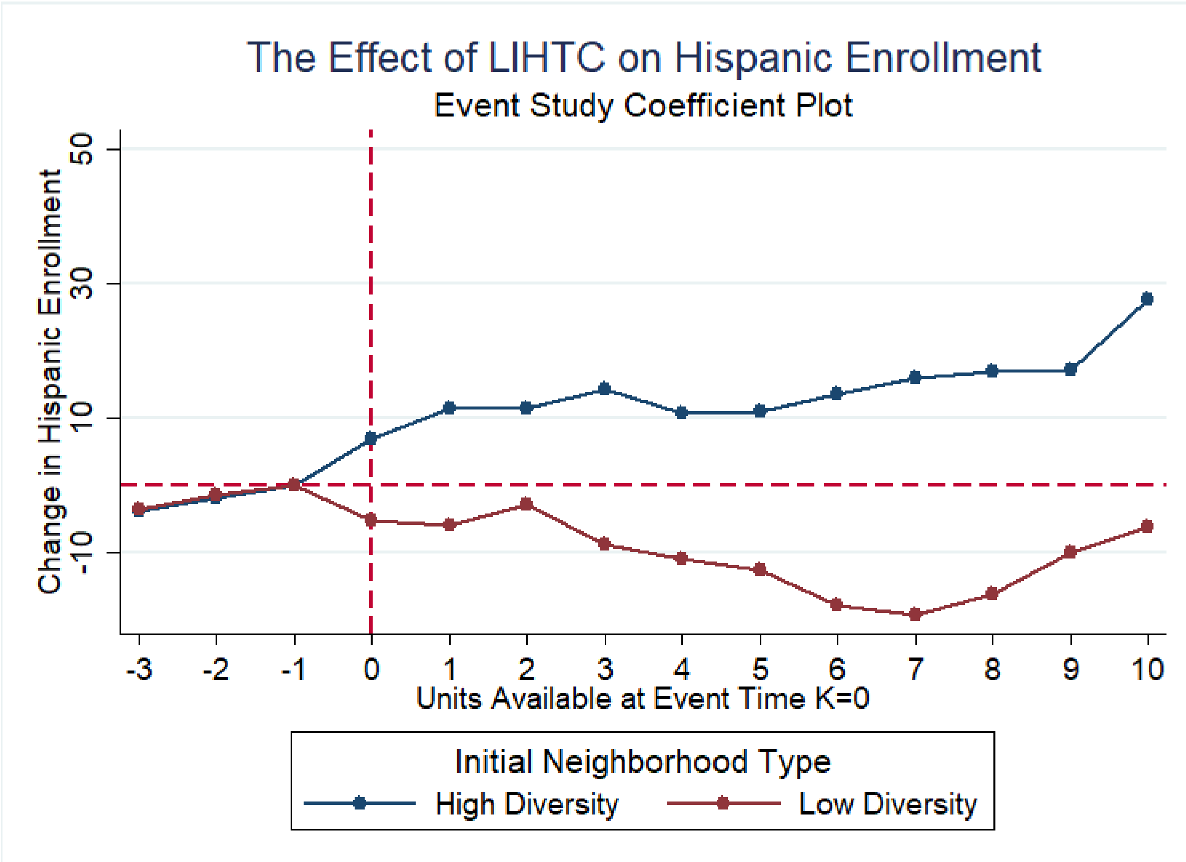


Figure 8: This diagram is a plot of the event study coefficients π_{gk} from regression equation 9 for g = 2, the high diversity group. The model predicts a decrease in diversity for these school types following LIHTC.

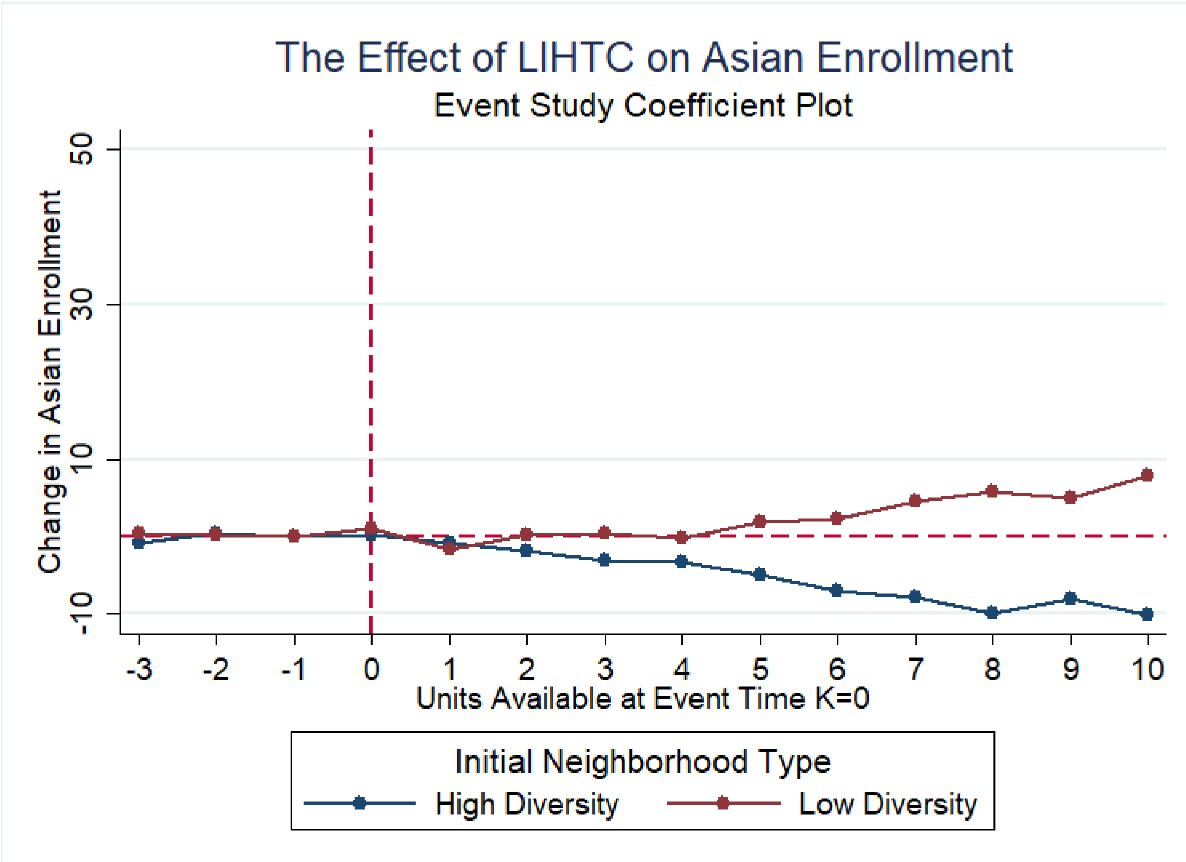


Figure 9: This diagram is a plot of the event study coefficients π_{gk} from regression equation 9 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

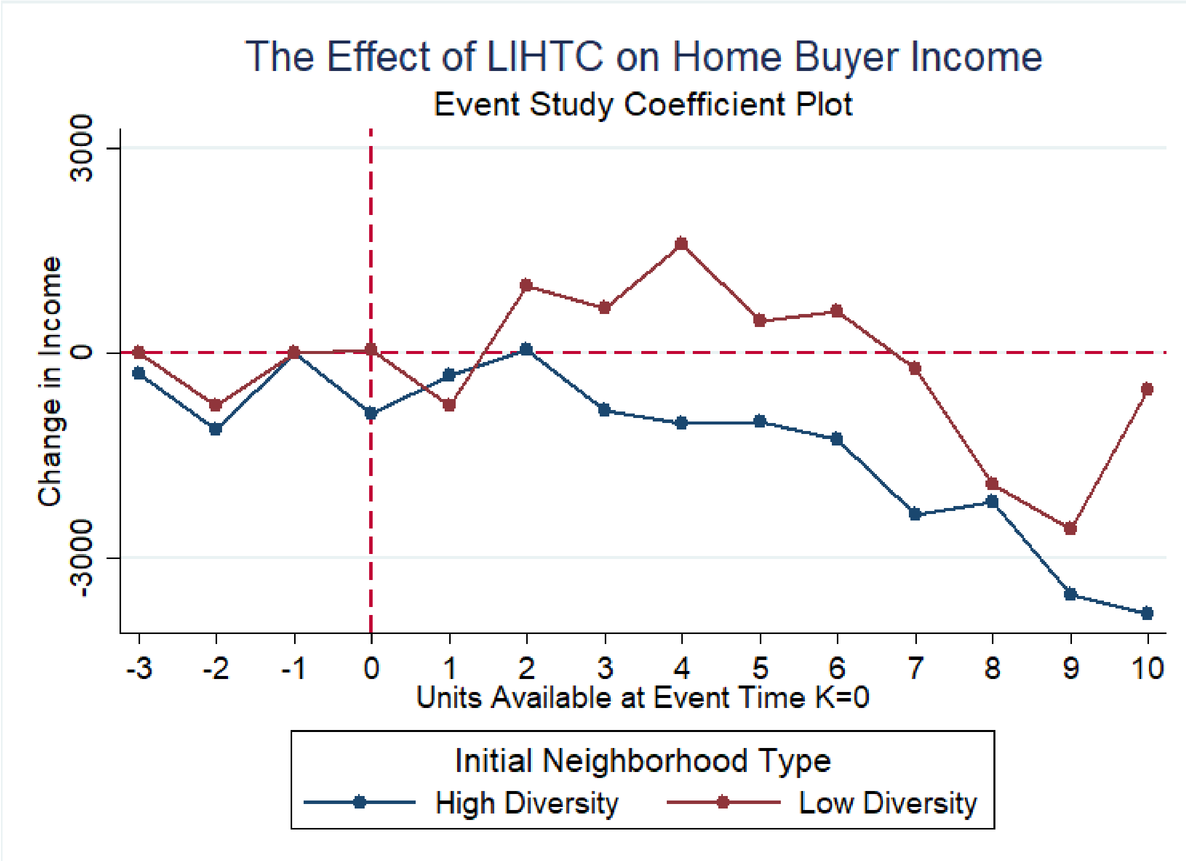


Figure 10: This is a plot of the event study coefficients from regressing log(applicant income) on the interacted event time dummies and the covariates included in first stage, for the initially low diversity school zones. For the low diversity school zones, the income of new home buyers did not change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

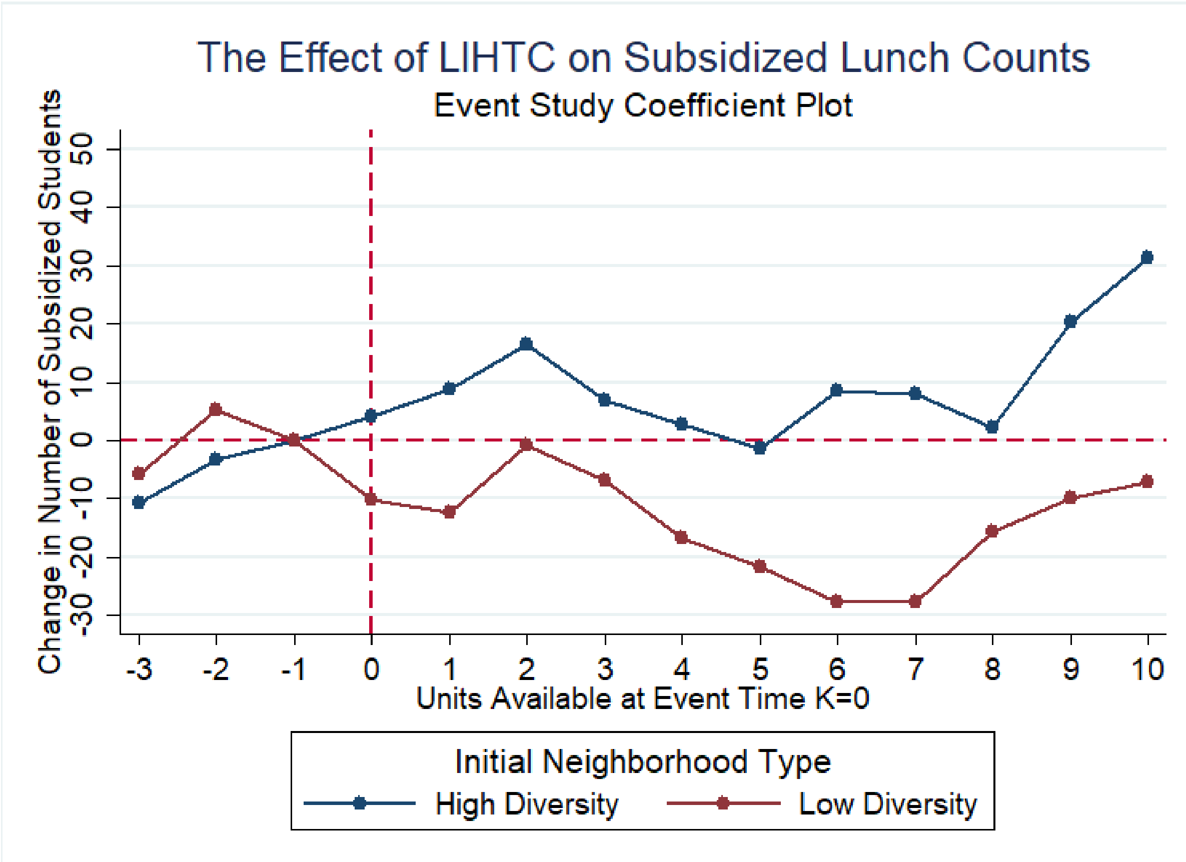


Figure 11: This is a plot of the event study coefficients from regressing log(applicant income) on the interacted event time dummies and the covariates included in first stage, for the initially high diversity school zones. For the high diversity school zones, the income of new home buyers did not change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

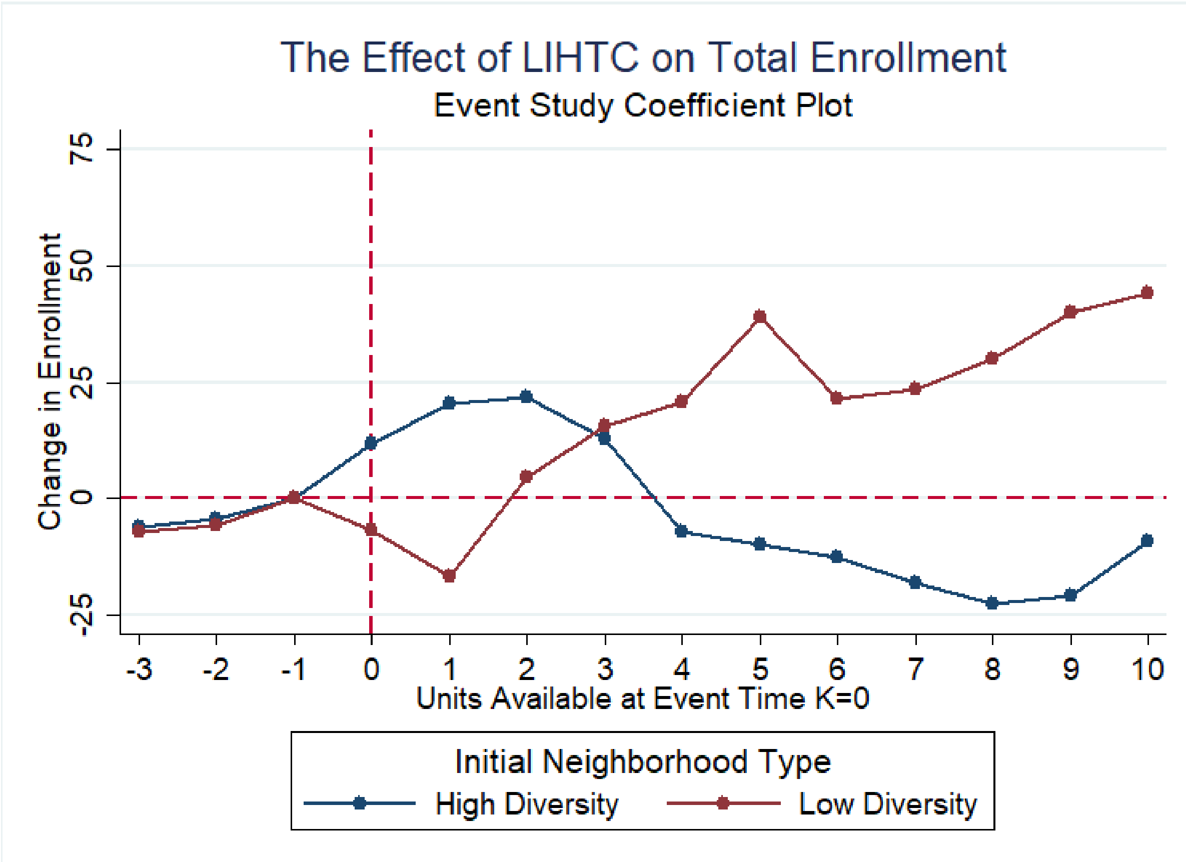


Figure 12: This diagram is a plot of the event study coefficients π_{gk} from regression equation 9 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

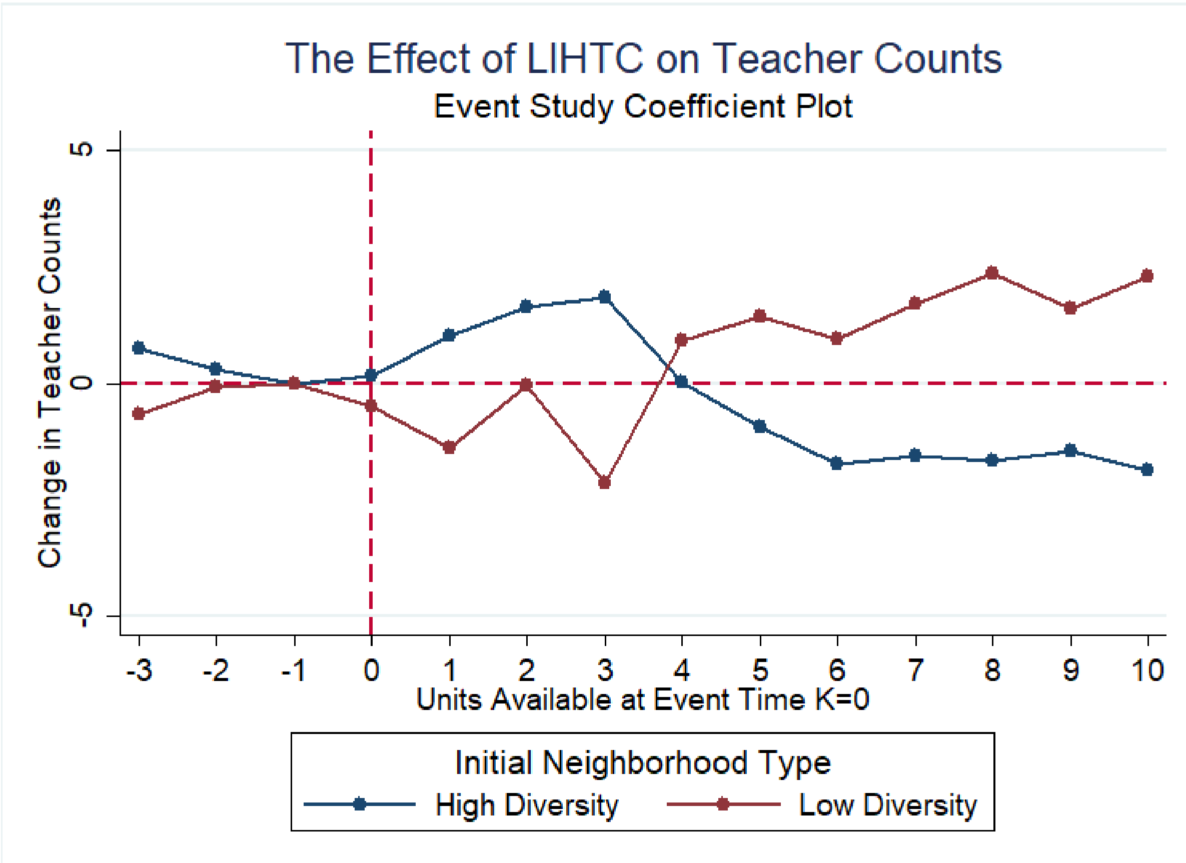


Figure 13: This diagram is a plot of the event study coefficients π_{gk} from regression equation 9 for g = 1, the low diversity group. The model predicts an increase in diversity for these school types following LIHTC.

All Regions

East North Central
East South Central
Middle Atlantic
Mountain
New England
Pacific
South Atlantic
West North Central
West South Central

East North Central

Illinois (IL)
Indiana (IN)
Michigan (MI)
Ohio (OH)
Wisconsin (WI)

East South Central

Alabama (AL)
Kentucky (KY)
Mississippi (MS)
Tennessee (TN)

Middle Atlantic

New Jersey (NJ)
New York (NY)
Pennsylvania (PA)

Mountain

Arizona (AZ)
Colorado (CO)
Idaho (ID)
Montana (MT)
New Mexico (NM)
Nevada (NV)
Utah (UT)
Wyoming (WY)

New England

Connecticut (CT)
Maine (ME)
Massachusetts (MA)
New Hampshire (NH)
Rhode Island (RI)
Vermont (VT)

Pacific

California (CA)
Oregon (OR)
Washington (WA)

South Atlantic

Delaware (DE)
Florida (FL)
Georgia (GA)
Maryland (MD)
North Carolina (NC)
South Carolina (SC)
Virginia (VA)
West Virginia (WV)

West North Central

Iowa (IA)
Kansas (KS)
Minnesota (MN)
Missouri (MO)
Nebraska (NE)
North Dakota (ND)
South Dakota (SD)

West South Central

Arkansas (AR)
Lousianna (LA)
Oklahoma (OK)
Texas (TX)

Figure 14: US Census Divisions. Source: National Center for Environmental Information, US Department of Commerce