On the Spatial Determinants of Educational Access*

Francesco Agostinelli[†] Margaux Luflade[‡] Paolo Martellini[§]

February 2024

Abstract

We study the spatial determinants of inequality in educational access, and assess the welfare effects of policies at scale trying to address this inequality. We develop a spatial equilibrium model of residential sorting and school choice, estimated using data from a large school district in the United States. The estimated model replicates the responses of neighborhood and school choice to quasi-experimental variation in peer school composition and school transportation provision. The model is used to evaluate the aggregate and distributional welfare effects of policies designed to enhance educational opportunities to low-income families: school-choice expansion (place-based) vs. housing vouchers (people-based). We estimate that both policies result in overall negative welfare effects, with only marginal improvements in low-income children's access to high-quality schools. Although eligible families generally benefit from these policies, some experience losses in equilibrium, in particular, those who invest in their children's education at baseline by residing in costly neighborhoods. We also find that geography and zoning regulations shape the spatial distribution of effects on higher-income families. Those attending high-quality schools located far away from low-income neighborhoods are shielded from the adverse effects of school-choice expansion policies. Those living in neighborhoods with more stringent zoning regulations are immune to the effects of housing voucher policies.

JEL Classification: I24, J13, R23, R31.

*We thank seminar participants at ASU, BFI China, Boston University, Cornell, Federal Reserve Bank of Minneapolis, NYU, Penn, Stanford, University of Arizona, UCLA, University of Minnesota, University of Oslo, Yale. Moreover, we are thankful to the attendees of workshops and conferences including the Chicago-Princeton Spatial Economics Conference 2023, Junior Spatial Conference 2022 at BFI, NBER SI 2023, Rohit and Harvanit Kumar Conference on Early Childhood Development at BFI, Saieh Family Fellows Workshop at BFI, and the Urban Economics Association Conference. We also thank the Office of Student Assignment of the Wake County Public School District, in particular the Student Information, Demographic and Spatial Analytics office, for sharing school boundary maps and helping us understand the WCPSS institutional setting. Ornella Darova provided outstanding research assistance. All errors are our own.

[†]Department of Economics, University of Pennsylvania. Email: fagostin@sas.upenn.edu

[‡]Department of Economics, University of Pennsylvania. Email: mluflade@sas.upenn.edu

[§]Department of Economics, University of Wisconsin-Madison. Email: martellini@wisc.edu

1 Introduction

Most cities in the United States are characterized by a significant degree of neighborhood income segregation, which translates into inequality in access to local amenities, such as high-quality public schools. This issue has attracted growing attention among policymakers, given the importance of early childhood education and peer effects on children's outcomes.

This paper asks two main questions. First, we quantify to extent to which institutional constraints to neighborhood and school choice rather than preference heterogeneity are responsible for the observed degree of sorting of families across neighborhoods and schools. Second, we investigate the aggregate and distributional implications of policies that grant low-income families access to high-quality schools. To answer these questions, we build a spatial equilibrium model of neighborhood and school choice. Families' choices are determined by preferences over school peer composition, commuting distance to school, house prices, as well as by local institutions like neighborhood-specific school choice sets and residential minimum lot size restrictions.

We estimate the model using data from the Wake County, North Carolina. We exploit the effect of longitudinal changes in school attendance boundaries on house price to identify families' willingness to pay for school quality. The model replicates these quasi-experimental effects, as well as the estimated effects of observed changes in school transportation on school enrollment. We find that low-income families disproportionally value proximity to school while high-income place a higher relative value on school peer composition, increasingly so the higher their children's skills.

We use the estimated model to evaluate the aggregate and distributional welfare impacts of two policies commonly discussed and implemented in the United States to mitigate inequality in access to high-quality schools: a place-based policy that expands school choice for families living in disadvantaged neighborhoods, and a people-based policy that offers housing vouchers to low-income families that locate to more affluent areas. Our results reveal that both policies generate a negative aggregate welfare effect while only marginally improving in access to high-quality schools for low-income children. This result is mainly driven by the detrimental effect of both policies on low-income inframarginal families who choose at baseline to pay a high house price to reside in neighborhoods offering access to high-quality schools. Finally, our findings show that the effectiveness of both policies varies based on the geography and housing regulations. High-quality schools located far away from low-income neighborhoods are shielded from the effects of school-choice expansion policies. High-income neighborhoods with more stringent zoning regulations are immune to the effects of housing voucher policies.

We develop a spatial equilibrium model of residential sorting and school choice that captures key institutional features common in numerous school districts across the United States: choice sets—or *portfolios*— of schools that vary by residential location, school transportation provision, and zoning restrictions on housing demand. Conditional on their residential location, families have access to a neighborhood-specific set of public schools they can apply to; they can also opt out the public school system and attend a private school. Families' school choices are determined by endogenous school quality—defined by the school peer composition—and disutility from commuting, which depends on distance from home and on whether transportation is provided. Seats in oversubscribed public schools are allocated to applicants via lottery. Residential choice is determined by the value of the school portfolio associated with each neighborhood, together with the cost of housing, the quality of exogenous neighborhood amenities, and neighborhood-specific zoning restrictions that impose a lower bound on housing consumption. Crucially for our policy counterfactuals, house prices, admission probabilities, and school (peer) quality are jointly determined in equilibrium.

Our empirical analysis focuses on Wake County, North Carolina, which serves as a natural setting for our analysis—it is covered by a single county-wide school district, the per-pupil expenditure is constant across schools in the district, and several institutional changes regarding the boundaries of school catchment area and school transportation have occurred in recent years. To map the model to the data, we build a new comprehensive dataset by combining several data sources. We use student-level administrative data from the North Carolina Education Research Data Center (NCERDC) that report, for each child attending public schools in NC, the school they attend, their test scores, and residential location. We merge these data with (i) yearly information from the Wake County Public School System (WCPSS) on school catchment areas and the portfolio of alternative public schools available to each address; (ii) information on school transportation provision to each address; and (iii) data on house prices and residential zoning regulation for the entire county.

We estimate the model via the method of simulated moments exploiting both longitudinal and cross-sectional variation for identification of model parameters. In particular, changes in the boundaries of school catchment areas over school years generate quasi-experimental variation in peer composition, which we leverage to identify families' willingness-to-pay (WTP) for school quality. Within-neighborhood sorting of children across schools by skill level helps identifying heterogeneity in families' valuation of school quality as a function of their child's skills. To identify commuting costs with and without available school transportation and by income, we use variation in distance traveled to school by families as a function of whether the attended school provides transportation from their neighborhood. We show that the identified commuting costs are able to further replicate the impact on school enrollment of longitudinal changes in the availability of school transportation to a given neighborhood. Parameters capturing the cost of private schools, which may vary by location and income group, are identified by the spatial variation in private school attendance, conditional on family income. Residual variation in neighborhoods sorting by income helps identifying (heterogeneity in) families' valuation of exogenous neighborhood amenities.

Parameter estimates reveal that economically disadvantaged (henceforth, *ED*) families have a lower WTP than higher-income (or *non-ED*) families for school peer quality. Conditional on income, families' valuation of school quality also increases with their child's skills. This gradient is particularly steep for higher-income families. For non-ED families with a child with skills at the 25th percentile of the skill distribution, thecompensated variation for a ten-percent decrease in peer quality in schools is $478, and it rises to $920 if the child's skills reach the 75th percentile of the skill distribution. The analoguous numbers for ED families are $85 to $150. The WTP to reduce commuting to school (of 1 mile) also varies across income groups, being twice as large for non-ED families, regardless of children's skills ($280 vs. $530). However, the relative WTP for school proximity compared to school peers, which represents a trade-off in school choice between distance and quality, is much higher for ED families, suggesting that they value proximity relatively more than school peer composition as compared to their higher-income counterparts.

We use the estimated model to conduct a welfare analysis of two policies commonly discussed and implemented in the United States to improve low-income families' access to high-quality schools: a place-based policy that expands school choice for families living in disadvantaged neighborhoods, and a people-based policy that offers housing vouchers to low-income families that locate to more affluent areas. In our place-based policy, we target low-income neighborhoods where over 60 percent of families are ED, expanding their school portfolios to include the top 40 schools in the district (which we deem *receiving* schools). In our people-based policy, we provide housing vouchers to ED families conditional on them locating within the catchment area of the same 40 receiving schools.

We find that both policies have negative aggregate welfare effects of similar magnitudes, amounting to an average $400 annual loss per family. While the beneficiaries of the policies obtain modest welfare gains—half of which are offset by higher commuting costs in the school choice expansion policy—the aggregate welfare loss is driven by the negative effects on families who reside in the catchment areas of the receiving schools in the baseline equilibrium. Among the latter, ED families are more exposed to—and are less able to insure against—the decreased school peer quality that results from the influx of new children into their school. In the school choice expansion policy, receiving schools that are located faraway from the targeted low-income neighborhoods are subject to high commuting costs (despite the provision of school transportation). Those schools, which tend to be located in more affluent neighborhoods, are shielded from the policy by their geographic location in the outskirts of the city. In the voucher policy, the role of distance is played by zoning regulations. While incoming children reside within the catchment areas of receiving schools, the presence of minimum size constraints on housing tilts their neighborhood choice toward less-regulated, lower income neighborhoods. Both policies highlight how preference heterogeneity and spatial constraints on the families' choices are central in evaluating the outcome of policies aimed at expanding access to schools for low-income families.

Related Literature. This paper is related to several strands of literature. First, this paper contributes to the literature on residential choice and school valuation, motivated by the well-documented fact that school quality is capitalized into house prices (Black, 1999). Epple and Sieg (1999) first developed a general empirical equilibrium framework in which households sort into neighborhoods based on their preferences for housing and local public goods, and local jurisdictions choose their property tax levels and provision of public goods.[1] Bayer et al. (2007) propose an empirical framework of neighborhood choice, where families have preferences over school quality and other neighborhood amenities. We depart from these studies by allowing each neighborhood to be associated with a portfolio of schools residents can choose from. We can then use our framework to analyze the equilibrium response of school composition and neighborhood sorting to counterfactual policies related to both school choice and housing. In contrast to these earlier studies, we also leverage longitudinal—instead of cross-sectional—variation in school catchment areas to disentangle households' willingness to pay for school quality from other neighborhood amenities.

[1] Sieg et al. (2004) further develop an empirical method to estimate willingness to pay for large changes in local public goods, in the presence of household spatial sorting, unobserved preference heterogeneity, and general equilibrium effects in the local housing market.

Our paper naturally relates to the empirical literature on school choice. We focus on the demand side of the education market and contribute to previous work interested in the determinants of school choice by endogenizing neighborhood choice—while recognizing the importance of the distance from home in families' choice of school for their children, this literature has largely treated home location as exogenously fixed (e.g., Hastings et al., 2009; Abdulkadiroğlu et al., 2017; Agarwal and Somaini, 2018; Kapor et al., 2020; Laverde, 2023). The role of school competition in shaping the outcomes of educational policies has been shown to be important (e.g., Singleton, 2019; Allende, 2020; Dinerstein and Smith, 2021; Campos and Kearns, 2023) but is beyond the scope of this paper.

Our approach to model neighborhood choice in the presence of school choice is in the spirit of Nechyba (2000), Epple and Romano (2003), and Avery and Pathak (2021). These papers propose models in which school quality and housing prices are determined in equilibrium and study the extent to which school choice policies increase low-income children's access to high-quality schools.[2] We also investigate these effects in equilibrium and show that geography plays an important role in determining the distributional effects such policies.[3] Park and Hahm (2023) examine the impact of the school assignment mechanism within a model of school and neighborhood choice. Our work differs from theirs by accounting for general equilibrium effects of spatial sorting on house prices and on school peer composition, which are central in assessing the response of families to—and the aggregate welfare impact of—the policies we analyze.

This paper also contributes to the growing urban literature that studies how the interaction between agglomeration and congestion forces shape city structure and individual outcomes (Ahlfeldt et al., 2015).[4] Recent contributions have focused on the effect of transportation infrastructure on commuting patterns (Heblich et al., 2020) and neighborhood sorting (Tsivanidis, 2019). Our paper also explores how transportation provision, along with local prices, affects access to desired locations within a city. However, we depart from much of the urban literature that deals with labor market access and explore heterogeneity across neighborhoods in their measure of educational access. Hsiao (2023) studies the provision of educational opportunities in a spatial equilibrium in which workers migrate toward labor markets with higher returns to education. Similar to us, Almagro and

[2] Nechyba (2000) focuses on school choice though private school vouchers; Epple and Romano (2003) and Avery and Pathak (2021) consider public school choice.

[3] Redding and Turner (2015) survey the theoretical and empirical literature on the impact of transportation costs on spatial economic activities.

[4] For a comprehensive review of recent advancements in quantitative models of economic geography, refer to Redding and Rossi-Hansberg (2017).

Domínguez-Iino (2022) delve into the role of endogenous local amenities in shaping the welfare distribution among residents within a city. They focus on leisure amenities, while our paper is centered around the quality of local schools.

Recent developments in macroeconomic research have shown a growing interest in understanding the impact of spatial segregation on human capital development. Early seminal papers examined the theoretical implications of neighborhood segregation on socio-economic inequality (e.g., Benabou, 1993; Durlauf, 1996; Fernandez and Rogerson, 1996). More recent studies have used quantitative dynamic models to examine how decisions related to residential choices affect the development of children's human capital through the transmission of skills across generations (Aliprantis and Carroll, 2018; Fogli and Guerrieri, 2018; Eckert and Kleineberg, 2021; Chyn and Daruich, 2022). Our research contributes to this field by delving into the micro-level institutional elements and regulatory factors that influence neighborhood sorting and educational choices, and thereby affecting aggregate welfare and its distribution.

2 Model

2.1 Overview

We build a static model of neighborhood and school choice by families with a child in elementary school age. Families are heterogeneous in terms of both their income and children's skills. Families choose their neighborhood of residence taking into account house prices, neighborhood-specific housing regulations, and exogenous and endogenous amenities. The endogenous amenity is given by the quality of the portfolio of schools associated with a specific neighborhood from which families must select the school they apply to. The set of schools in a neighborhood's portfolio is defined by school attendance boundaries, which we take directly from the data. Schools are characterized by some exogenous characteristics, namely their capacity and location in the city. The number and skill composition of children that apply to a particular school determine the equilibrium probability of admission to and the quality of that school, respectively. Home–school distance determines the disutility cost from commuting, which varies according to whether school transportation is provided.

The city is also composed of the rest of households, or other households, o, by which we mean households without a child in elementary school age. The rest of households—in contrast to families—are only heterogeneous with respect to their income and only choose their residential location. While families are the main focus of our analysis, incorporating other households into the model allows us to measure the families' valuation for school quality from changes in house prices, given that families represent only a fraction of the housing demand in the city.

2.2 Environment

Demographics. The city, which represents a single school district, is populated by a measure $m + 1$ of households. A measure 1 of households has one child in elementary school (families). A measure m of households does not have children in the relevant age group (the rest of households or 'others'). Families are of type (w, a), where w is the household income and a is the child's skills. The joint distribution over family types is exogenously given by $\phi(w, a)$. The rest of households have income w^o with distribution $\phi^o(w^o)$.

Geography. The city is two-dimensional and it is partitioned into N neighborhoods. With a slight abuse of notation, we denote by n both a neighborhood and the location of its centroid. Similarly, schools are denoted by s and each has a unique location in the city.

School Portfolios. In line with our empirical setting covering the Wake County Public School System (WCPSS) each neighborhood n is associated with a portfolio of schools, $\mathcal{L}_n$. The school portfolio of neighborhood n includes three mutually exclusive subsets of public schools and a subset of private schools. The first (singleton) subset consists of the base school, denoted by $\mathcal{B}_n$, which is the only school in the portfolio of the neighborhood that provides both guaranteed admission and bus transportation. For later use, we define a catchment area $\mathcal{C}_s$ as the set of neighborhoods n that share the same base school s, $\mathcal{C}_s = \{n : \mathcal{B}_n = \{s\}\}$. The second subset of public schools, denoted by $\mathcal{T}_n$, does not guarantee admission but provides bus transportation to students enrolled. The third subset of public schools does not guarantee admission nor they offer bus transportation from home to school ($\mathcal{NT}_n$). We refer to the latter two sets of schools as *option schools*. The probability of being admitted to an option school depends on a lottery system with admission probability denoted by p_s, described below. Note that each school might belong to a different set in different neighborhoods. For example, a school might be a base school for some neighborhoods, an option school with transportation for some other neighborhoods, an option school that does not provide transportation for yet another set of neighborhoods, and outside the choice set for the rest of the city.

Children within a neighborhood can also enroll in the set of available private schools,

$\mathcal{P}_n$. While our primary focus is on public school attendance and composition, we introduce private schools into the model to account for the possibility that families choose to opt out of the public school system. For simplicity, we limit this set to one private school per neighborhood–the closest to home–although multiple neighborhoods might share the same private school. We denote by $\mathcal{C}_s^{\mathcal{P}}$ the set of neighborhoods n that share the same private school s, $\mathcal{C}_s^{\mathcal{P}} = \{n : \mathcal{P}_n = \{s\}\}$. To summarize, the portfolio of schools available in each neighborhood is $\mathcal{L}_n = \{\mathcal{B}_n \cup \mathcal{T}_n \cup \mathcal{NT}_n \cup \mathcal{P}_n\}$.

Public School Capacity and Admission Probability. Base schools are mandated to admit students from their catchment area. Therefore, we assume that each school acting as base has sufficient capacity to guarantee admission to all children residing in their catchment area. This includes students who applied directly to their base schools and all children who did not obtain admission through the application lottery to their preferred option schools. Additionally, each school s provides a limited number of seats, denoted by q_s, to children for whom that school is an available option school given the neighborhood where they reside, $\{n : s \in \mathcal{T}_n \cup \mathcal{NT}_n\}$. If the number of applicants to school s exceeds its capacity q_s, applicants are rationed through a lottery that determines who is admitted. Hence, the admission probability to school s for a child in neighborhood n is equal to 1 if $s \in \mathcal{B}_n$, and to $p_s \leq 1$ if $s \in \{\mathcal{T}_n \cup \mathcal{NT}_n\}$. If children are not admitted to the option school they apply to, they can choose whether to enroll in their base or local private school. Families can always choose to enroll their children in a private school (and being admitted with probability one) without applying to any school in the public school system.

Preferences over Schools. We divide the choice of neighborhood and school into two sequential steps. This modeling decision is supported by the contextual information that parents must have verifiable proof of residency before they apply to schools within the neighborhood school portfolio. Conditional on living in a neighborhood n, families with children endowed with skill level a obtain the following utility from attending public school $s \in \mathcal{L}_n \setminus \{\mathcal{P}_n\}$

$$v_{a,w,s|n} = \gamma_{a,w} \ln \bar{a}_s - \kappa_{s,w} \tau_{ns} + \sigma_S \varepsilon_s \,. \tag{2.1}$$

Here, $\bar{a}_s$ represents the average skills of children enrolled in school s, which is endogenous to the school composition. Our assumption that demand for schools is a function of peers is motivated by two main reasons. First, Abdulkadiroğlu et al. (2020) analyze applicants' rank-ordered choices and find that parents value schools that enroll high-achieving peers. Interestingly, other measures of "school effectiveness" and "match effects" do not predict rank-ordered choices of families once school peer quality is controlled for.[5] We conduct a test in line with their approach utilizing our dataset. Although we do not directly observe rank-ordered choices of parents, our data includes house prices. Using the econometric model detailed in Section 4, we exploit longitudinal changes in school attendance boundaries to estimate the impact of changes in school peers on neighborhood house prices. When we augment our specification with a measure of school value-added estimated from test scores, we find the coefficient on such proxy for school effectiveness to be small and statistically indistinguishable from zero, and the coefficient on school peers to be unaffected. We interpret this result as evidence that parents prioritize the quality of peers when evaluating schools over other measures of effectiveness. Second, we think of our measure of school quality as the policy-relevant one, since other determinants of school quality (e.g., good teachers) tend to be disproportionally attracted to schools with a positively selected pool of children (e.g., Jackson, 2009).[6] Hence, in our counterfactual analysis, we interpret the endogenous changes in peer composition to account for both the direct impact of peers on demand for schools and their effects on other components of the educational production process.

The term $\kappa_{s,w}\tau_{ns}$ denotes the disutility associated with commuting. The commuting cost depends on the home-school road distance in miles τ_{ns}, family income w, and whether the school s provides bus transportation or not

$$\kappa_{s,w} = \begin{cases} \kappa_{T,w} & \text{if} \quad s \in \mathcal{T}_n \cup \mathcal{B}_n \\ \kappa_{NT,w} & \text{if} \quad s \in \mathcal{NT}_n \,. \end{cases}$$

Finally, $\sigma_S \varepsilon_s$ represents an idiosyncratic preference shock that follows a standard Extreme Value Type 1 distribution. The parameter σ_S governs the dispersion of these preference shocks.

The value of attending a private school is instead defined by

$$v_{a,w,\mathcal{P}_n|n} = \gamma_{a,w} \ln \bar{a}_{\mathcal{P}_n} - \kappa_{NT,w}\tau_{n\mathcal{P}_n} - \chi_{w,n} + \sigma_P \varepsilon_{\mathcal{P}_n}, \tag{2.2}$$

[5]The authors examine the New York City centralized high school assignment mechanism. Their empirical model defines a potential outcome equation for the test score of a child attending a given school. They use this model to conduct a decomposition exercise to analyze the observed variation in outcomes across schools using three key terms: peer quality, school effectiveness (defined by school indicators), and the match effect (defined by the interaction of school indicators with individual student characteristics).

[6]Our assumption that school quality is given by its childrens' skill composition is shared with ample previous work including Bayer et al. (2007), Epple and Romano (2003) and Avery and Pathak (2021), among others.

which is analogous to the value of attending a public school, with the exception of the additional term $\chi_{w,n}$ accounting for the distinct valuation of private schools over public schools based on family income. Throughout the paper, we will refer to this term as the private school cost for families, which incorporates factors such as the monetary expense of enrollment. In Wake County, the decision of private schools to offer transportation is decentralized compared to the public school system, and we lack detailed geographical data on individual private schools' transportation provision. Nevertheless, given that the vast majority of private schools do not offer transportation, we make the assumption that none do.

Each family can apply to only one option school, taking the equilibrium admission probabilities p_s as given.[7] Families applying to either their base or private school are guaranteed admission. For those applying to option schools, admission for their children is contingent upon winning the admission lottery. Families who are unsuccessful receive a value $v^{fallback}_{a,w|n} \equiv \max\{v_{a,w,\mathcal{B}_n|n}, v_{a,w,\mathcal{P}_n|n}\}$ from attending their preferred fallback school. It follows that the expected value of the school portfolio accessible to families in neighborhood n can be succinctly expressed as:

$$\bar{v}_{a,w}(\mathcal{L}_n) = \mathbb{E}_{\{\varepsilon_s\}}\left[\max_{s\in\mathcal{L}_n}\{p_s v_{a,w,s|n} + (1-p_s)v^{fallback}_{a,w|n}\}\right], \tag{2.3}$$

where the expectation is taken with respect to the realization of idiosyncratic preferences, unobserved at the time of choosing the neighborhood.[8] We deem the expression in Equation 2.3 our measure of *educational access*, which is reminiscent of the commuter market access in Ahlfeldt et al. (2015). Intuitively, a neighborhood has higher educational access when there are many schools, with good peers, that are located close to it.

Housing Supply and Zoning Restrictions. Each neighborhood is characterized by inelastic total housing supply H_n, as well as by neighborhood-specific housing regulation

[7]Option schools at the elementary level in the WCPSS include calendar options and magnet programs (see Appendix A.1 for details). Like in our model, families can apply to only one calendar option. As for magnet programs, for half of our sample (and in all previous years until then) families submit an ordered ranking of at most three options and are assigned via the Boston mechanism. As the Boston mechanism gives higher priority to the school ranked first, and since for most schools in the WCPSS the number of first-choice applicants exceeds the number of seats (Dur et al., 2018), most applicants are only considered for admission to their first choice (or the base school as their outside option), rendering all choices ranked below the first one irrelevant. Assignment during the last two years of our sample switched to Deferred Acceptance (DA). Since families can only apply to five out of the 15-20 options in the average choice set, the new mechanism retains strategic incentives, as in our model. For further discussion, see Appendix A.1.

[8]To avoid introducing additional notation, Equation (2.3) includes the trivial lottery in which a family simply chooses their base or private school.

(zoning).[9] In neighborhoods where regulation is binding, houses must adhere to the minimum housing size stipulated by neighborhood-specific zoning restrictions, denoted by h_n^{reg}. Conversely, in neighborhoods without housing regulations, the size of a house can be as small as h_0, which we interpret as the minimum physical space necessary for habitation.

Preferences over Neighborhoods. Families' utility over neighborhoods is represented by the following utility function:

$$v_{a,w,n} = u_{w,n} + \alpha_{0n} + \alpha_{1n} \log(w) + \bar{v}_{a,w}(\mathcal{L}_n) + \sigma_N \varepsilon_n \tag{2.4}$$

where

$$u_{w,n} = \max_{c,h} \left(\frac{1-\beta}{\beta}\right) \ln\left(\frac{c}{1-\beta}\right) + \ln\left(\frac{h}{\beta}\right) \tag{2.5}$$

represents utility from consumption of a numeraire good, c, and housing services, h. The second and third terms in Equation 2.4 are the value of neighborhood amenities that are common across households, α_{0n}, and amenities that vary by family income, $\alpha_{1n} \log(w)$, respectively. The term $\bar{v}_{a,w}(\mathcal{L}_n)$ is the expected value of the portfolio of schools associated with neighborhood n, described in Equation (2.3). Finally, the term $\sigma_N \varepsilon_n$ represents a neighborhood-specific taste shock, which is independent and identically distributed across households and neighborhoods according to the Extreme Value Type 1 distribution. The parameter σ_N determines the dispersion of these shocks, which families observe when choosing their neighborhood.

Families maximize utility subject to the following budget and housing constraints

$$\begin{aligned} \hat{w} = (1+\tau)w &\geq c + r_n h \\ h &\geq \underline{h}_n = \max\{h_0, h_n^{\text{reg}}\}. \end{aligned} \tag{2.6}$$

where the price of c is normalized to one, while each unit of housing costs r_n. Families divide their income between consumption of the numeraire good and housing. The endogenous transfer rate τ, which families take as given, originates from the land share of total housing expenditures. Specifically, we assume that a fraction μ of housing expenditure constitutes land rent, distributed proportionally to households' non-housing income

[9] For simplicity, we abstract from neighborhood-specific housing supply elasticity. Incorporating how the stock of housing responds to shifts in demand would alter the long-term housing market dynamics. We conjecture that such addition would not alter our main conclusions given the moderate response of house prices in our policy counterfactuals.

w, while the complementary fraction represents the production cost of the structure.

Our neighborhood amenities likely represent a combination of both truly exogenous amenities (e.g., parks, proximity to bodies of water) and potentially endogenous amenities that may evolve as the income composition of the neighborhood changes, independently of the neighborhood's school portfolio (Almagro and Domínguez-Iino, 2022). Disentangling the relative importance of exogenous and endogenous amenities is challenging due to the need for variation in neighborhood composition that is independent from the changes in school portfolios we exploit in our empirical analysis. If the desirability of neighborhood amenities is positively linked to neighborhood income, as it is commonly found in the literature (Diamond, 2016; Tsivanidis, 2019; Redding and Sturm, 2024), our approach provides a conservative evaluation of the equilibrium effects resulting from the endogenous sorting of households in response to the policies we analyze.

Preferences of the Rest of Households. The rest of households have the same preferences and are subject to the same constraints as families, except for two differences. First, we allow exogenous amenities and preference shocks for a given neighborhood to be different between families and the rest of households. This modeling choice aims to capture the idea that these households may place different values on specific local amenities—such as restaurants and gyms—compared to families with children in elementary school. Second, the quality of the neighborhood schools does not directly affect the choice of the rest of households. However, it is important to note that this does not imply that school quality has no bearing on the equilibrium residential sorting of the rest of households. Given that housing represents a rival service between families and the rest of households, the endogenous school quality does impact equilibrium prices, thereby affecting the resulting residential sorting of both groups. The resulting utility is then given by

$$v^o_{w,n} = u_{w,n} + \alpha^o_{0n} + \alpha^o_{1n} \log(w) + \sigma^o_N \varepsilon_n. \tag{2.7}$$

where $u_{w,n}$ is as in Equation 2.5 and the superscript o denotes the rest of the households, or others.

2.3 Equilibrium

Our spatial equilibrium follows naturally from aggregation of individual choices, which in turn are functions of aggregate outcomes like house prices r_n, school peers $\bar{a}_s$, admission probabilities p_s, and the transfer rate from land income τ.

Conditional on living in a given neighborhood, households choose how to allocate their income between the numeraire good c and housing h. Due to minimum housing constraints, the housing demand is nonlinear in income, and it is given by $h_{w,n} = \max\{\beta\hat{w}/r_n, \underline{h}_n\}$. The indirect utility from consumption is then equal to

$$u_{w,n} = \begin{cases} \frac{1}{\beta}\ln\hat{w} - \ln r_n & \text{if } \beta\hat{w} \geq r_n\underline{h}_n \\ \left(\frac{1-\beta}{\beta}\right)\ln\left(\frac{\hat{w}-r_n\underline{h}_n}{1-\beta}\right) + \ln\left(\frac{\underline{h}_n}{\beta}\right) & \text{if } \hat{w} > r_n\underline{h}_n > \beta\hat{w} \\ -\infty & \text{otherwise.} \end{cases} \tag{2.8}$$

For given house prices, more stringent zoning restrictions result in a higher fraction of families that are either priced out of the neighborhood or forced to spend a higher share of income on housing than they would like. The presence of zoning amplifies the contribution of housing demand to neighborhood income sorting and it will play an important role in our voucher counterfactual policy in Section 5.

Recall that the indirect utility over neighborhood n for families of type (a, w) is given by

$$v_{a,w,n} = u_{w,n} + \alpha_{0n} + \alpha_{1n}\log(w) + \bar{v}_{a,w}(\mathcal{L}_n).$$

Thanks to Extreme Value Type 1 idiosyncratic preferences, the measure of families living in neighborhood n satisfies the canonical logit formulation

$$\pi_{n|a,w} = \frac{\exp(v_{a,w,n})}{\sum_{\tilde{n}}\exp(v_{a,w,\tilde{n}})}.$$

Given their neighborhood of choice, families apply to schools according to the conditional probability

$$\pi_{s|n,a,w} = \Pr\left[v_{a,w,s|n} \geq v_{a,w,\tilde{s}|n}\ \forall\tilde{s} \in \mathcal{L}_n\right]$$

and the probability is induced by the idiosyncratic preference shocks over schools.[10] It is then straightforward to compute the measure of families living in neighborhood n and applying to school s

$$\pi_{s,n|a,w} = \pi_{n|a,w}\pi_{s|a,w,n}.$$

[10]The absence of a closed-form expression for choice probabilities when agents choose the lottery that maximizes their expected utility is a well-known feature of the empirical school choice literature. It escalates the computational burden of solving and estimating the model, as the choice probability (here, $\pi_{s|n,a,w}$) needs to be recovered by simulations. See, for instance, Agarwal and Somaini (2018); Calsamiglia et al. (2020); Luflade (2018); also see Agarwal and Somaini (2020) for a review of this literature.

The indirect utility and choice probabilities for the rest of households are only defined with respect to neighborhoods—not schools—and they are given by

$$\begin{aligned} v^o_{w,n} &= u_{w,n} + \alpha^o_{0n} + \alpha^o_{1n}\log(w) \\ \pi^o_{n|w} &= \frac{\exp(v^o_{w,n})}{\sum_{\tilde{n}}\exp(v^o_{w,\tilde{n}})}. \end{aligned}$$

Building toward the definition of our aggregate equilibrium variables, we define the following auxiliary variables. Let

$$\begin{aligned} \pi_{w,n} &= \Sigma_a \pi_{n|a,w}\phi(a,w) + m\pi^o_{n|w}\phi^o(w) \\ \pi_s &= \Sigma_{a,w}\Sigma_{n\notin \mathcal{C}_s\cup\mathcal{C}^{\mathcal{P}}_s}\pi_{s,n|a,w}\phi(a,w) \\ \pi^{attend}_{a,s} &= \begin{cases} \Sigma_{w,n} p_s \pi_{s,n|a,w}\phi(a,w) & \text{if} \quad s\in\mathcal{T}_n\cup\mathcal{NT}_n \\ \Sigma_{w,n}\left(\pi_{s,n|a,w}+\pi^{fallback}_{s,n|a,w}\right)\phi(a,w) & \text{if} \quad s\in\mathcal{B}_n\cup\mathcal{P}_n \end{cases} \end{aligned}$$

be the measure of households of type w who live in neighborhood n, the measure of applicants to s as an option school, and the measure of children of type a who *attend* school s, respectively. The first two variables are a straightforward aggregation of individual choices. The third variable is a combination of individual choices and school lottery outcomes. Specifically, for base and private schools, the measure of attendees must account for the measure of children who turn to school s as a fallback school after losing the admission lottery to their preferred option,

$$\pi^{fallback}_{s,n|a,w} = \sum_{\tilde{s}\in\{T_n\cup NT_n\}} \pi_{\tilde{s},n|a,w}(1-p_{\tilde{s}})\mathbb{I}\{v^{fallback}_{a,w|n} = v_{a,w,s|n}\}, \quad s\in\mathcal{B}_n\cup\mathcal{P}_n.$$

The market clearing condition for housing in neighborhood n reads

$$H_n = \sum_w \pi_{w,n}h_{w,n}. \tag{2.9}$$

and the transfer rate τ, proportional to housing expenditure, is then equal to

$$\tau = \mu\frac{\sum_n r_n H_n}{I}. \tag{2.10}$$

where I is total income in the city (summing both families' and the rest of households').

Admission probabilities are either equal to 1, if the school has enough seats to accommodate all applicants, or some value less than 1, if the school is oversubscribed. Formally,

$$p_s = \min\left\{\frac{q_s}{\pi_s}, 1\right\}. \tag{2.11}$$

Recall that such probability only applies to option schools for a given neighborhood. The same school would provide guaranteed admission to children who reside in neighborhoods for which that school acts as base. Similarly, private school admission is guaranteed to all those who apply—either as their first choice or after losing the lottery to access an option.

Last, the quality of peers is given by the average skills of children that attend a certain school,

$$\bar{a}_s = \mathbb{E}_s[a], \tag{2.12}$$

where the expectation $\mathbb{E}_s$ is taken with respect to the (conditional) distribution of children of skills a who attend school s, $\pi_{a,s}^{attend}$.

We are now ready to define an equilibrium for this economy.

Definition 1. *An equilibrium for this economy is a set of choice probabilities for families, $\{\pi_{s,n|a,w}\}$, and the rest of households, $\{\pi^o_{n|w}\}$, house prices $\{r_n\}$, school peers $\{\bar{a}_s\}$, admission probabilities $\{p_s\}$, and transfer rate $\{\tau\}$, such that:*

- *The choice probabilities $\{\pi_{s,n|a,w}\}$ and $\{\pi^o_{n|w}\}$ are induced by (i) families' solutions to the school choice problem (2.3), and choice of consumption, housing, and neighborhood to maximize the objective function (2.4), subject to budget and minimum housing constraints (2.6); and (ii) the rest of households' maximization of the objective function (2.7), subject to budget and minimum housing constraints (2.6);*

- *The housing markets clear (2.9), the admission probability is consistent with school capacity and applications (2.11), and school composition is consistent with individual choices and admission probabilities, (2.12);*

- *The transfer rate is consistent with the land share of housing expenditure as in (2.10).*

3 Mapping the Model to the Data

Our empirical analysis focuses on Wake County, North Carolina, which serves as a natural setting to investigate our questions of interest for several reasons. First, it is covered by a single county-wide school district, the Wake County Public School System (WCPSS), spanning approximately 850 square miles, which makes considerations about geographical access and transportation relevant. Second, the per-pupil expenditure is constant across schools in the district. Third, several institutional changes regarding the boundaries of school catchment area and school transportation have occurred over the past two decades, providing valuable variation for identification of the model parameters.

In this section, we describe our data sources, explain how the primitives of the model are mapped to the data, and provide descriptive statistics for our sample. Further details are reported in Section A.2.

3.1 Data Sources and Measurement

Our data come from several sources. Student-level data, which include school attended, end-of-grade test scores in Reading and Mathematics from grade three to eight, economically disadvantaged status (henceforth ED, measured by eligibility for free or reduced-price lunch), and residential address (until the 2016-17 school year), were obtained from the North Carolina Education Research Data Center (NCERDC). Starting in 2013-14, these data also document scores obtained at literacy assessments administered at the beginning of the kindergarten year. We focus on children entering kindergarten between Fall 2013 and Fall 2016, the four cohorts for which both residential addresses and kindergarten test scores are observed. To construct our measure of child skills (a), we aggregate kindergarten test scores into a unique index using Bartlett factor scores and standardize it by cohort. We discretize the resulting index of skills into deciles to define children's types, $a \in \{a_1, \ldots, a_{10}\}$. For each public school s, the number of seats reserved for option (as opposed to base) students in year t (q_{st}) is measured as follows. If s is oversubscribed in year t, then q_{st} is set to the count of kindergarten students attending s as an option in that year. If it is not, then q_{st} is set to the average number of kindergarten students attending s as an option in the years it is oversubscribed.[11]

The Wake County Public School System (WCPSS) provided yearly data showing the assignment of residential addresses to base schools and menus of options, as well as the availability of school transportation between each address and each option school. We exploit such institutional design of the public school system to define neighborhoods in our city. Specifically, a neighborhood n is the set of all contiguous addresses that share, in all

[11]While we do not observe individual applications, school-level information provided by the WCPSS indicates whether option schools are oversubscribed.

four school years in our sample, the same sets of base, option schools, and transportation availability–that is, the same portfolio of public schools. Figure A-4 shows the partition of Wake County into neighborhoods obtained with this definition, as well as 2010 Census tracts for comparison.

We combine several publicly available datasets to build neighborhood characteristics.[12] House prices are measured from the Wake County Real Estate Records, which contain all real estate transactions in the county since 1956. We use the average price per square foot observed in neighborhood n in year t as our measure of neighborhood-level house prices (r_{nt}). Information on the minimum lot size (MLS) in each neighborhood is obtained by collecting and standardizing zoning regulation data from the county and the twelve distinct municipalities and unincorporated areas in charge of regulating different land areas within Wake County. In the model, we assume households choose *house* size, rather than *lot* size, so we map our constructed neighborhood-level measure of MLS restrictions ($\underline{\text{mls}}_n$) into a minimum available house size (h_n^{reg}). Regressing observed house sizes (in square feet) on $\underline{\text{mls}}_n$ (in acres) yields the mapping: $\underline{h}_n = 641 + 892 \times \underline{\text{mls}}_n$. From this mapping, we derive the minimum house size in the absence of any housing regulation, $h_0 = 641$, and $h_n^{\text{reg}} = \underline{h}_n \mathbb{I}\{\underline{\text{mls}}_n > 0\}$.

Because the NCERDC data provides information about children's ED status, rather than family income, we use the American Community Survey (ACS) five-year estimates (2013-17) to measure income for families and other households. We aggregate the sixteen income brackets provided by the ACS to define ten discrete household income types, $w \in \{w_1, \ldots, w_{10}\}$.[13] Mapping our neighborhoods to 2010 Census tracts, we use tract-level information from the ACS to obtain the distribution of the other households' types ($\phi^o(w)$) as well as their sorting across neighborhoods ($\pi^o_{n|w}$). We also use this mapping of neighborhoods to census tracts to combine information from the ACS and the NCERDC and construct the joint distribution of family types ($\phi(w, a)$) and their sorting across neighborhoods ($\pi_{n|a,w}$). Namely, we derive the distribution of family income conditional on residential neighborhood and ED status from the ACS, which we then combine with the distribution of children skills conditional on residential neighborhood and ED status available in the NCERDC. See Appendix A.2.3 for the exact details of the construction of all neighborhood-level characteristics.

[12]See Appendix A.2.3 for the links to these datasets.

[13]See Appendix A.2.3 for details. To map the ten income types to the ED measure available in the NCERDC, we assign the lower four types to be ED (that is, with gross family income in the past 12 months below $50,000 in 2017 dollars). For reference, the eligibility thresholds for the school year 2017-2018 for reduced-price lunch were $45,510 annual income for a household of four.

Finally, we use the bi-annual National Center for Education Statistics (NCES) Private School Survey for years 2013, 2015, and 2017 for information about the location of private schools in Wake County. For each neighborhood n, we let the private school associated to the neighborhood ($\mathcal{P}_n$) be the nearest private school to the neighborhood centroid. We use Public Use Microdata Area (PUMA)-level information from the the ACS (2013–17) to construct for each of the eight PUMAs constituting Wake County, the share of kindergartners attending private school by ED status in our sample period.

3.2 Final sample

Our final sample consists of 34,428 students across four kindergarten cohorts.[14] Detailed descriptive statistics about the samples of schools, students, and neighborhoods are reported in Table A-3. Thirty-five percent of students in the sample qualify as ED. The average skill gap between ED and non-ED students is equal to 86 percent of a standard deviation of the log-skills distribution (average log-skills of -0.56 for ED versus 0.30 for non-ED students). About 75 percent of students attend their base school.

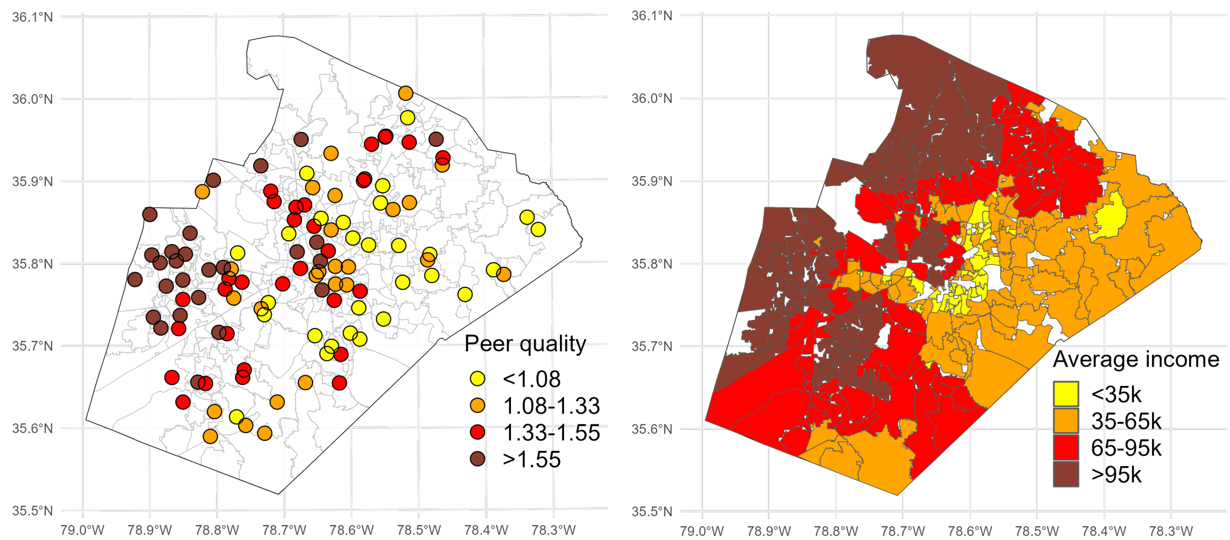
There are 111 public elementary schools in our sample. While all 111 schools are base schools, 76 (85 percent) are also option schools for some neighborhoods. Most option schools are oversubscribed. The share of ED students ranges from zero to ninety percent across schools, with a mean of 0.39 and a standard deviation of 0.22. In addition, between 2013–14 and 2016–17, 49 private schools offered kindergarten, enrolling on average 1,431 kindergartners per year.[15]

The final partition of Wake County into our 312 neighborhoods, with the 2010 Census tracts boundaries for comparison, is shown in Figure A-4. Figure 1 shows the geography of neighborhood income and public school peers quality. The two geographies align notably well—higher-income neighborhoods and schools with higher-skilled peers are located in the west and in some of the center of the county, while lower-income neighborhoods and schools with lower-quality peers can be found downtown and in the eastern part of county. Average house prices follow a very similar pattern (see Figure A-5(a)). Fi-

[14] We drop student observations for which school attended, address, kindergarten literacy assessments, or ED status is missing. We also exclude from the sample students attending a school outside the menu of schools attached to their residential address. Students may attend schools outside of the choice set attached to their residential address for multiple reasons. For instance, they may attend the same school as an older sibling (which was a part of the choice set in the past) or they may attend a school at which one of their parents is employed.

[15] Source: https://ncadmin.nc.gov/public/private-school-information/state-north-carolina-private-grade-k-12-school-statistics for private schools. Traditional public schools enrolled an average of 11,998 kindergartners a year over that period (https://www.wcpss.net/domain/100).

Figure 1: Geography of School Quality and Neighborhood Income



(a) Public school peers quality

(b) Neighborhood income

Each map in this figure shows Wake County, NC. The left map shows peer quality for each school in the sample. Our measure of peer quality is average (standardized) kindergarten literacy assessment score (source: NCERDC, Spring 2014 to Spring 2017). The peer quality thresholds used to determine the colors of the dots represent the quartiles of the distribution of school peer quality in the sample. The right map shows average household income by neighborhood (source: ACS five-year estimates 2013–17). The income thresholds used to determine the colors of the neighborhoods represent the quartiles of the distribution of average neighborhood income in the sample. Areas left uncolored are out of our sample and essentially non-residential.

nally, zoning regulations are mapped in Figure A-5(b). There is significant heterogeneity in MLS regulations across neighborhoods and higher-density neighborhoods are concentrated in the urban center of the county, that is, the city of Raleigh. Accordingly, there is significant variation in average house size across neighborhoods (mean of 882 square feet, standard deviation of 252 square feet).

On average, each neighborhood is associated with a choice set of 17 option schools in addition to its base. Option schools tend to be much farther away from the neighborhood than the base is (11 versus 3.7 miles on average). Twenty-three percent of neighborhoods experienced a change in base school over the sample period; all of them experienced a change in eligibility and/or transportation to some option school. As explained in the next section, these types of institutional changes are useful for the identification of families' valuation of school attributes. It has been well-established in the literature that, regarding changes in catchment areas in Wake County, "the selection of any given [neighborhood] for reassignment was, conditional on observable traits of the [neighborhood], essentially random and not manipulable or anticipated by [neighborhood] residents" (Hill et al., 2023, p. 7). While this quote is suggestive of the quasi-randomness in the re-design of catchment areas, we perform a formal statistical test and fail to reject the null hypothesis that changes in catchment areas are uncorrelated with pre-trends in house prices at the neighborhood level. That is, the redrawing of school boundaries does not occur in a systematic way within neighborhoods experiencing either notably higher or lower price trends (see Appendix A.1.2). The reasons underlying changes in transportation provision are not as well documented in the literature. To alleviate concerns about the endogeneity of these changes, we perform a similar statistical test as we do for changes in school catchment areas. Also in this case, we fail to reject the null hypothesis that the changes in transportation provision are uncorrelated with pre-existing trends in within-neighborhood school enrollment shares.

4 Model Identification and Estimation

We estimate the model using the Simulated Method of Moments (SMM). We recover the following parameters: preferences for school peers $\gamma_{a,w}$, commuting costs $(\kappa_{T,w}, \kappa_{NT,w})$, neighborhood amenities for families and other households, $(\alpha_{0n},\alpha_{1n})$ and $(\alpha^o_{0n}, \alpha^o_{1n})$ respectively, cost of attending a private school $\chi_{w,n}$, and the variance of preference shocks $(\sigma_N, \sigma_S, \sigma_P)$. We additionally obtain a model-implied measure of housing supply, denoted as H_n. While we emphasize that there is no one-to-one mapping between parameters and moments, we offer an intuitive argument of how each parameter is identified based on the selected moments. Detailed information regarding the moments computation is available in Appendix C. We report asymptotic standard errors for the structural estimates computed via the delta method.

4.1 Identification

Preferences for School Quality. We let preferences over school peers vary with skill type a and ED status, hence we estimate two vectors of parameters, $\{\gamma_{a,\text{ED}}\}$ and $\{\gamma_{a,\text{non-ED}}\}$, with $a \in \{a_1, \ldots, a_{10}\}$. These parameters are identified by two main types of variation. First, we estimate how house prices vary with changes in peer quality induced by the redrawing of school catchment areas. Unlike previous literature that mostly utilizes cross-sectional variation across school boundaries (see for example Black, 1999; Bayer et al., 2007), we instead leverage longitudinal variation resulting from policy-induced changes in boundaries within the same neighborhood. If parents value peer quality, an increase in the quality of their base school should trigger an increase in the demand for local housing, hence an increase in house prices. This type of variation is more robust to potential biases arising from sorting into neighborhoods, since cross-sectional variation might confound differences in school service provision between the two neighborhoods along the catchment area boundary, with unobserved heterogeneity in preferences among individuals residing on each side of the boundary.[16]

Second, the average peer composition across schools attended by families of given skill-income type is informative of the heterogeneity in preferences for school peers. Heterogeneous γ's can potentially increase or decrease the extent of school sorting with respect to the level induced by neighborhood sorting and home-school distance only–both of which we discuss below. Intuitively, house prices allow us to identify the absolute valuation of school peer quality (vs other neighborhood amenities, for instance), while school composition identifies the relative valuation of school peer quality across family types.

Starting from the identifying moment, we estimate the impact of peer quality on house prices, both in the data and in the model, using the following regression:

$$\Delta_n \ln r_{n,t} = \beta_1 \Delta_n \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t}},t} + \delta_t + \Delta_n \epsilon_{n,t}\,, \tag{4.1}$$

where $\Delta_n \ln r_{n,t}$ represents the within-neighborhood log-change in average house prices and $\Delta_n \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t}},t}$ is the policy-induced change in the quality of the catchment area associated to school $\mathcal{B}_n$ for two consecutive cohorts of kindergarten children in $t-1$ and t. Recall the definition of the catchment area of school s, $\mathcal{C}_{s,t} \equiv \{n | \mathcal{B}_n = s \text{ in school year } t\}$. Formally:

$$\Delta_n \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t}},t} = \ln \widehat{\bar{a}}_{\mathcal{C}_{\mathcal{B}_{n,t}},t-1} - \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t-1}},t-1}\,. \tag{4.2}$$

The variable $\widehat{\bar{a}}_{\mathcal{C}_{\mathcal{B}_{n,t}},t-1}$ is the mean test score of children who were living in period $t-1$ in the catchment area of neighborhood n's base school at time t. That is, $\widehat{\bar{a}}_{\mathcal{C}_{\mathcal{B}_{n,t}},t-1}$ captures the hypothetical composition of $\mathcal{C}_{\mathcal{B}_{n,t},t}$ if families' residential locations was the same as at time $t-1$. The variable $\Delta_n \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t}},t}$ then measures the change in the composition of

[16] Kuminoff and Pope (2014) highlight how exploiting longitudinal variation in school boundaries over extended periods requires assuming that the estimated hedonic price model is time-invariant. We believe this concern to be less relevant in our case as our panel data covers a relatively short period, spanning four school years. We replicate the same regression in our structural model, in which we also assume that individual preferences remain fixed throughout the period of analysis.

a base school potential attendees, from one cohort of kindergarten children to the next, induced only by changes in the boundaries of the catchment area, holding everything else (in particular residential choices) constant.

The parameter of interest is β_1, which we interpret as the average responsiveness of house prices to changes in catchment area quality. House prices may be affected by a change in catchment area through two channels. They may respond to the direct policy-induced changes in the pool of potential school peers within the catchment area. They may also adjust due to endogenous changes in the skill composition of the neighborhood, as families change their residential choices in response to the policy. The variable $\Delta_n \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t}},t}$ allows us to isolate the former channel. Our preferred estimate of β_1 is 0.092 (see Column (2) in appendix Table B-1), implying that an increase of 10 percent in school quality translates into a one-percent increase in house prices.[17]

The other 19 moments we use for the estimation of $\{\gamma_{a,\text{ED}}\}$ and $\{\gamma_{a,\text{non-ED}}\}$ are, for each skill level a, the difference between the average peer composition of schools attended by type-(a, ED) (resp. $(a, \text{non-ED})$) children and the average peer composition of schools attended by type-(a_1, ED) children. We adopt this approach as the overall skill distribution in the district is fixed and therefore only the relative school peer composition is informative about the extent of preference heterogeneity. The *level* of preferences for school peers is then pinned down by matching the estimate of β_1 described above.

Commuting Cost. We allow the disutility from commuting to vary by family ED status and school transportation provision. That is, we estimate the per-mile utility loss from commuting to the base school or to an option school with busing, $(\kappa_{T,\text{ED}}, \kappa_{T,\text{non-ED}})$, and the utility loss from commuting to an option school without busing or to a private school, $(\kappa_{NT,\text{ED}}, \kappa_{NT,\text{non-ED}})$. For each ED-status group, we target the average distance traveled to school conditional on attending an option school that either provides or does not provide transportation. Lower values of the parameter κ imply a smaller role for distance in

[17] As outlined in Section 2.2, when we expand the regression presented in Equation (4.1) to incorporate variations in school quality not due to peers by adding a measure of school value added, our findings indicate that, while the estimates for peer quality β_1 remain unchanged, variation in the value added resulting from the assignment of the neighborhood to a different base school do not affect house prices. School value added A_s is computed via a standard school value added model where A_s is given by the school fixed effect in a regression of children's 3rd-grade test scores on the child's kindergarten test score and average kindergarten peer test scores. The augmented model is:

$$\Delta_n \ln r_{n,t} = \beta_1 \Delta_n \ln \bar{a}_{\mathcal{C}_{\mathcal{B}_{n,t}},t} + \beta_2 \Delta_n A_{\mathcal{B}_{n,t}} + \delta_t + \Delta_n \epsilon_{n,t} \,,$$

where $\Delta_n A_{\mathcal{B}_{n,t}}$ represents the changes in school value added induced by changes in the assigned base school for different cohorts in neighborhood n. The coefficient β_2 of school value added in this augmented model is equal to 0.016 with a standard error of 0.025.

determining school choice and–all else equal–longer commutes.

As one of our counterfactual policies involves expanding the set of option schools that provide transportation, we exploit similar longitudinal variation in the data to investigate the extent to which the model replicates the observed changes in school attendance. Specifically, our focus is on examining the within-neighborhood longitudinal changes in transportation availability and on how that impacts attendance at option schools. Table B-2 and Table B-3 in Appendix B.2 show estimates obtained from the data and from the model, respectively. The model matches the data very well. In particular, the model is able to quantitatively replicate the increase in attendance to schools that introduce transportation provision, with a decreasing effect the further the school is from home.

Idiosyncratic School Preferences. The parameter that governs the dispersion of idiosyncratic preferences for schools, denoted by σ_S, is identified by the residual dispersion in public school choice. Our target moment is the standard deviation of within-neighborhood and school year enrollment rates across public schools, averaged across neighborhoods and school years. Intuitively, the higher the value of σ_S the more equal the shares, as observable differences between schools (i.e. peer composition and distance) play a smaller role in school choice.

Private School Valuations and Idiosyncratic Preferences. Our parsimonious treatment of private schools is dictated by the lack of individual-level data on children attending them. Shares (by ED status) of children attending private school are only observed at the PUMA level. We map neighborhoods to the PUMA they belong to, let χ vary at the PUMA level, rather than at neighborhood level, and choose the values of $\chi_{\text{ED,PUMA}}$ and $\chi_{\text{non-ED,PUMA}}$ that allow the model to replicate those attendance shares. Using data from the WCPSS in the school year 2015-2016, Dur et al. (2022) find that applicants who lose the option school lottery are 10 percentage points less likely to enroll in public schools than those winning the lottery. We set σ_P to match the same moment in our model. For given share of children enrolled in private schools, higher values of σ_P increase the likelihood that families choose private schools upfront rather than after losing a public option lottery.

Neighborhoods: Amenities, Housing Supply, Idiosyncratic Preferences. The neighborhood amenities are set to replicate the share and average income of families and the rest of households that live in each neighborhood. The parameters α_{0n} and α^o_{0n} capture the common valuation for neighborhood n, while α_{1n} and α^o_{1n} let neighborhood preferences vary systematically with household income–after accounting for observable neighborhood attributes like school quality, house prices, and zoning restrictions.

To identify the dispersion in idiosyncratic neighborhood preferences (σ_N), we use empirical findings from the Moving to Opportunity (MTO) experiment in Galiani et al. (2015). We replicate the MTO experiment within our model, where we proxy poverty rates with ED status available in our school administrative data. Poor families living in neighborhoods with more than 40 percent of poor households at baseline are offered a housing voucher equal to the 40th percentile of the rent distribution in Wake County, which they can use if they choose to live in a neighborhood where the share of ED household is below 40 percent. The value of σ_N is then chosen to replicate the take-up rate of 63 percent reported by Galiani et al. (2015).[18] Interestingly, we observe that families utilizing the voucher opt to relocate to neighborhoods with an average ED share of 25 percent, a figure closely resembling the 19 percent reported by Galiani et al. (2015).[19]

We further investigate the ability of the model to account for the heterogeneous residential responses of families to changes in neighborhood compositions following the redesign of catchment areas. We do so by estimating a version of Equation 4.1, wherein the change in prices on the left-hand side is replaced by the change in the share of ED families residing in the neighborhood. It is important to note that the right-hand side variable reflects the change in the composition of the catchment area of a neighborhood's base school induced solely by the redesign of its boundaries, excluding the subsequent changes in residential sorting, which is the focal point of the present analysis. The estimated impact in the data is equal to -0.13 and is significant at the 5% level, showing that non-ED families exhibit greater elasticity in residential responses to improvements in the skill composition of their catchment areas compared to ED families. Interestingly, the empirical estimate is closely replicated by the model, with a value of -0.15.

Finally, we assume that housing supply takes the form $\ln H_{n,t} = \ln H_n + \ln H_t$. The parameter H_n is informed by the average price in each neighborhood across years, while H_t determines the average price across neighborhoods in each given year. A key property of this specification is that housing supply does not vary systematically across neighborhoods over our years of analysis. Hence, this specification is consistent with our reduced-form regression 4.1, which exploits the *differential* effect on house prices of changes in

[18]Since we adopt a parsimonious specification for amenities and we model rich income heterogeneity, dispersion in idiosyncratic neighborhood preferences could in principle be identified from cross-sectional variation in choice shares, along the lines of the identification of σ_S. We believe that our current strategy, based on experimental evidence, does not rely excessively on our parametric assumptions and the granularity of the income distribution we observe.

[19]We set $\sigma_N^o = \sigma_N$ for simplicity. We have experimented with alternative values of σ_N^o and found negligible differences in outcomes once $(\alpha_{0n}^o, \alpha_{1n}^o)$ are re-estimated as well.

school boundaries to identify preference for school quality. We believe this assumption to be reasonable in our setting, considering the short time span between the re-design of catchment areas and the start of the school year. Furthermore, we expect that violations of this assumption would likely result in an increase in housing supply in neighborhoods benefiting from an improvement in the composition of their catchment area, making our estimates of the valuation for peers a conservative lower bound.

4.2 Estimation Results

Estimates for the parameters governing the valuation of school peer composition, commuting costs, and the scale of idiosyncratic preferences are shown in Table 1. Estimates for neighborhood amenities and private school valuations are shown in Figures C-1 and C-2.

We find the valuation for school peers to be increasing in the children's own skills. Such increase is much steeper for high-income than for low-income families. One interpretation for this result is that children's skills in kindergarten reflect children's innate ability. Moreover, parents–particularly those with higher incomes–may place greater importance on the quality of their child's school peers according to their child's inherent abilities, resulting in the estimated positive correlation between preferences and skills. An alternative explanation posits that the heterogeneity in skills arises from unequal parental investments during the early stages of children's lives. Parents who prioritize the quality of kindergarten school peers might be those who invested in their children's skills early on, thus generating the observed positive correlation between γ's and our measure of skills (Heckman and Mosso, 2014). It is easy to imagine that both initial conditions and endogenous investments in early life contribute to the realized heterogeneity in skill endowments upon entering kindergarten. However, assessing the relevance of each, while certainly valuable, is beyond the scope of this paper.

Turning to commuting costs, our estimates indicates that families value the provision of school transportation. The marginal cost of commuting an extra mile to schools without access to school transportation is more than double the marginal cost with transportation. We emphasize that attending schools located far away from home is costly even when transportation is provided. Families may have various reasons for disliking distance, including the necessity to wake up their children earlier in the morning and the increased inconveniences associated with accidents or sickness episodes while at school. Finally, the marginal cost of commuting varies by family income, with lower-income families

Table 1: Estimated Parameters

Peer Quality Valuation	γ_{a_1}	γ_{a_2}	γ_{a_3}	γ_{a_4}	γ_{a_5}	γ_{a_6}	γ_{a_7}	γ_{a_8}	γ_{a_9}	$\gamma_{a_{10}}$
ED Families	0.57	0.51	0.51	0.61	0.72	0.82	0.85	0.79	1.00	1.36
	(0.22)	(0.21)	(0.21)	(0.22)	(0.22)	(0.23)	(0.23)	(0.24)	(0.26)	(0.30)
Non-ED Families	0.43	0.68	0.77	0.96	1.06	1.19	1.30	1.57	1.94	2.82
	(0.18)	(0.21)	(0.22)	(0.24)	(0.25)	(0.27)	(0.29)	(0.32)	(0.35)	(0.43)
Commuting Cost	κ_{NT}	κ_T								
ED Families	0.12	0.05								
	(0.006)	(0.003)								
Non-ED Families	0.08	0.03								
	(0.005)	(0.002)								
Dispersion Parameter of Shocks	σ_N	σ_S	σ_P							
All Families	1.08	0.23	0.61							
	(0.12)	(0.02)	(0.12)							

The table shows estimates for preferences over peer quality, commuting cost, and dispersion parameters for utility shocks over neighborhoods, public schools, and private schools. Standard errors are reported in parentheses.

experiencing a higher cost of commuting per mile traveled to school.

Preferences for neighborhoods are determined by income-specific amenities and idiosyncratic shocks. Figure C-1 displays the geographic distribution of estimated amenities for the average ED and non-ED family. We highlight two features of our estimates. First, regardless of ED status, there exists a considerable dispersion in preferences for neighborhood amenities. Second, non-ED families exhibit a pronounced preference for neighborhoods located in the western part of Wake County, where high-quality schools are also located. Conversely, preferences among ED families appear to be less spatially concentrated.

As shown in Table 1, we also find that that the variance for idiosyncratic preferences for neighborhoods is larger than for schools (1.08, as compared to 0.23 and 0.61 for public and private schools, respectively). This is consistent with the many reasons for which the attractiveness of neighborhoods may be vary idiosyncratically across households, like proximity to the parents' workplace or to the children's grandparents.

Finally, Figure C-2 in Appendix shows the estimates for the utility from attending a private school, χ. Unsurprisingly, all our estimates are negative, indicating that all else equal families would prefer to attend a public school. Although heterogeneous by ED status, the estimated PUMA-level private-school disutility parameters closely match the average private-school tuition in each PUMA in Figure C-3. In addition, we estimate the variance of idiosyncratic preferences for private schools to be larger than for public

Table 2: Compensated Variation

	25th percentile child skills		75th percentile child skills	
	ED families	Non-ED families	ED families	Non-ED families
2017$ EQUIVALENT				
Peer quality increases by 10% of std	85	478	150	920
Dist. to schools decreases by 1 mile	263	545	291	519
Transportation to all options	260	511	238	326
Cost of Private School down by 10%	35	255	47	561

The table shows the compensating variation in house prices ($) required to keep families indifferent following the change outlined in the leftmost column, while holding everything else constant. The four columns of the table report the compensated variation as a function of family income (qualifying as ED or not) and the child's skills (25th and 75 percentiles of the skill distribution).

schools (0.61 vs. 0.23 in Table 1), which is expected since our model does not consider potentially important factors influencing private school choice, such as religious affiliation.

Interpreting the Estimates. To better assess the magnitude of the estimated parameters, Table 2 shows the dollar amounts different types of families would be willing to pay for changes in peer quality, distance to school, transportation provision, or the cost of attending private schools. We show our results for ED- and non-ED families and for the 25th and 75th percentile of the skill distribution within ED status, which correspond to the 2nd (4th) and 6th (9th) decile of the overall skill distribution for ED (Non-ED) families.

The first row reflects the gradient in families' school valuation with respect to their child's skills found in Table 1. High-income families with a high-skilled child are willing to pay about twice as much for an increase in school quality by 10 percent of a standard deviation as high-income families with a child with lower skills ($920 vs $478). The gradient is still present for low-income families albeit less steep and at lower scale, as their willingness to pay for the same change ranges from $85 to $150 as children skills increase from the 25th to 75th percentile of distribution.

Comparing the compensated variation results in Table 2 for distance (second row) and bus transportation (third row) with the compensated variation for peers helps understand the relative value families place on shorter commutes and transportation provision with respect to school quality. Non-ED families with children of lower skill levels assign a similar value to reducing their school commute by one mile or obtaining school transportation for option schools as they do to having a better peer composition by 10 percent of a standard deviation. However, the importance of peers increases significantly, doubling or even tripling when compared with distance and transportation provision, for families with higher-skilled children. Economically-disadvantaged families consistently prioritize commuting distance and school transportation provision over school peer quality. Their willingness to pay for reducing distance to school is two to three times greater than their willingness to pay for peer quality.

Families also exhibit significant heterogeneity in their willingness to pay for neighborhood amenities and private schools. Using our estimates we convert the dispersion in valuation for amenities displayed in Figure C-1 into their dollar value. We find that moving from the 25th to the 75th percentile of the distribution of neighborhood amenity is valued approximately $\$9,700$ and $\$43,800$ for ED and non-ED families, respectively. In both cases, these estimated values represent about forty percent of their income. Concerning private schools, the willingness to pay ranges from \$35 to \$561, reflecting heterogeneous preferences over private education, denoted by χ. While this parameter captures both the monetary and non-monetary costs of attending private schools, our estimate aligns closely with the observed level of private school tuitions. The primary source of this heterogeneity in willingness to pay for private schools appears to stem from income differences rather than heterogeneity in children's skills.

5 Policy Counterfactuals

Our model estimates shed light on key determinants of residential and school choice. Heterogeneous parental preferences for exogenous and endogenous (i.e., educational access) neighborhood amenities, coupled with zoning regulations, generate the observed neighborhood income segregation. The existence of neighborhood-specific school catchment areas and the unequal provision of school transportation across neighborhoods translate residential segregation into school segregation.

In this section, we investigate the aggregate and distributional impact of two of the most debated policies aimed at addressing inequality of educational access in the United States. We first consider a place-based policy that expands school choice for lower-income neighborhoods by including in their choice set schools with higher-skilled peers, which are mostly located in high-income neighborhoods. We then turn to a people-based policy that offers housing vouchers to low-income families willing to live in the catchment area of those same high quality schools, that is, in neighborhoods for which those schools are base schools.

Before proceeding further with the description of the policies and the presentation of our results, we briefly introduce the metric we adopt in order to evaluate the impact of these policies on aggregate welfare and on each of its determinants. We quantify the effect of the policies on family welfare as the equivalent variation in units of annual income. Specifically, let ΔV_i be the change in value for agent i and λ_i be family i's marginal utility of income. The welfare impact of the policy is then equal to

$$\Delta W = \sum_i \frac{\Delta V_i}{\lambda_i}. \tag{5.1}$$

Computing the welfare impact of the policy is not only of obvious interest in its own right, but it provides a natural way to measure the impact of the policies on various individual outcomes in a common unit.

In addition, we break down ΔV_i into its various determinants: peers, commuting, unobserved preferences for public schools, unobserved preferences for private schools, exogenous neighborhood amenities, and consumption (both numeraire and housing). We emphasize that due to the lottery-based admission process for option schools, the contribution to welfare from school-specific variables reflects the probability distribution over the school ultimately attended by the child in family i. In Appendix D.1, we present the mathematical formulation for the breakdown of ΔV_i into its determinants.

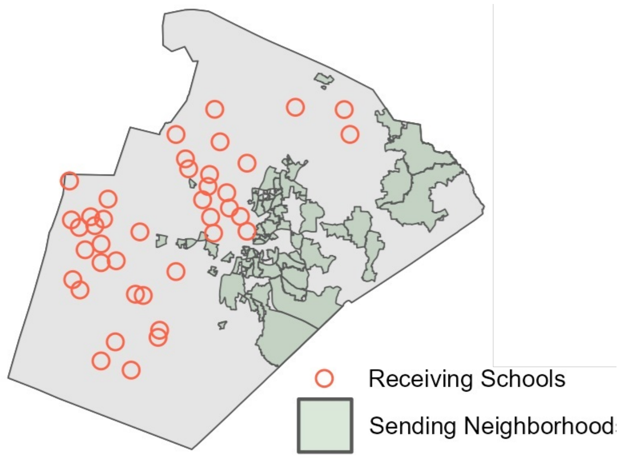
We apply our counterfactual policies to the institutional environment of the 2015-2016 school year, for which we have direct estimates of the probability of choosing a private school after losing the lottery to option schools. Implementing the same policies in any of the other school years in our sample delivers very similar results. Last, we report our welfare analysis only for families. The rest of households are affected by the policy only via its impact on house prices, delivering an average welfare change that is two orders of magnitude smaller than the one experienced by families.

5.1 Expanding School Choice

Over the past few decades, there has been extensive discussion surrounding school-choice policies which aim to uncouple residential location from school assignments (Epple and Romano, 2003; Cullen et al., 2006; Hastings and Weinstein, 2008; Deming et al., 2014). In our analysis, we expand the school choice set for the lowest-income neighborhoods in the city and include, as options with school transportation, the schools with the highest skill composition. We investigate the aggregate and distributional consequences of this policy, with a particular emphasis on how such policy affects the composition of peers for families with different income and residential location in the city.

Policy Design. We identify the 40 schools with the best peer composition at baseline. We

Figure 2: School Choice Expansion Policy



The figure shows the locations of the sending neighborhoods (green areas) and receiving schools (red dots) involved in the school choice expansion policy.

refer to these schools—and to the neighborhoods in their catchment areas—as *receiving*. We add these receiving schools to the school portfolio of neighborhoods with the highest shares of ED families in the county—above 60 percent. We refer to these as *sending* neighborhoods.[20] To accommodate the increase in eligible option students, each receiving school is endowed with a number of additional option seats equal to ten percent of its baseline capacity. In cases where a school receives more applications than it has available option seats, these seats are allocated via a lottery, as described in the model. So as to keep the aggregate number of option seats constant in the county, we decrease the number of option seats in each non-receiving school by 25 percent.

The spatial distribution of sending neighborhoods and receiving schools is shown in Figure 2. Sending neighborhoods are located in the eastern and center parts of the county while receiving schools tend to be in the western side. This spatial division between West and East maps closely to the income distribution shown in Wake County, as shown in Figure 1.

Results. We obtain two main results. First, while the initial goal of the policy is to allow lower-income children to attend schools with better peers, ED families experience virtu-

[20]A few lower-income neighborhoods in center city have one of the receiving school as base at baseline. These neighborhoods are not included in the *sending* group despite their large share of ED families because their choice set of schools is not changed by the policy. A few receiving schools are already in the choice set of some sending neighborhoods at baseline as options; in this case, transportation is added if it is not available at baseline.

Table 3: Heterogeneity among ED Families and Baseline Neighborhood Choices

	Neighborhood chosen at baseline		
	Receiving	Sending	Other
Average share of ED families	17	73	33
Average house price for ED fam. (per sqft)	8.85	6.42	7.86
Average child skills cond. on ED	1.03	0.89	0.98

The table shows the average share of ED families among residents of receiving, sending, and other neighborhoods at baseline. It also reports the average children skills among ED families conditional on their baseline neighborhood choice, as well as the average house price they pay (per square foot) at baseline.

ally no change in peer composition on average. Second, the policy has heterogeneous welfare effects not only across income groups, but also within the set of low-income families, according to the family's location within the city.

Figure 3 shows the impact of the policy on welfare (black bars) and on its determinants (colored bars), as detailed in Equation (D-1), for different sets of families. Figure 3(a) shows the average welfare gains across all families and by ED status. The policy induces an average welfare loss of \$450 per family (left panel). Losses are particularly large for non-ED families ($-$\$721, right panel) while ED families on average benefit by a small amount (+\$73, center panel). Remarkably, changes in peer quality account for virtually none of the welfare gains experienced by ED families (middle red bar).

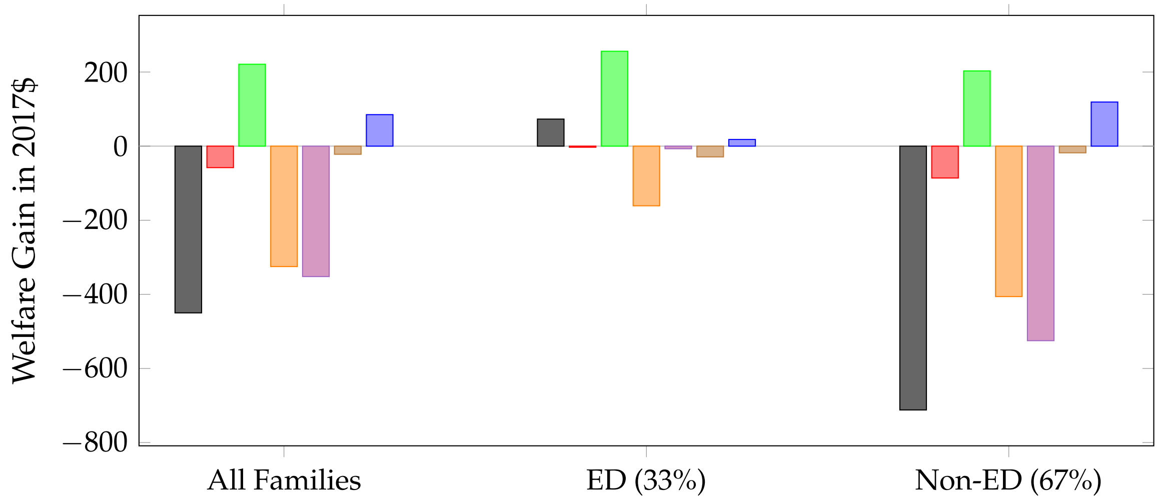
The average outcome masks significant heterogeneity among ED families, in particular as a function of their location choice in the baseline equilibrium. By design of the policy, ED families represent the majority of residents in sending neighborhoods. Nonetheless, ED families are also present in receiving neighborhoods at baseline, comprising 17 percent of the receiving population. As detailed in Table 3, these families spend more in housing to access higher-quality base schools for their children and are positively selected in terms of their children's skills, especially compared to ED families choosing sending neighborhoods at baseline.

Figure 3(b) shows how such heterogeneity in sorting across neighborhoods maps into unequal welfare changes for three distinct subsets of ED families—ED families residing in receiving neighborhoods at baseline ("ED receivers"); ED families residing in sending neighborhoods at baseline and who apply to a receiving school after the school choice expansion ("ED compliers"); and all other ED families.

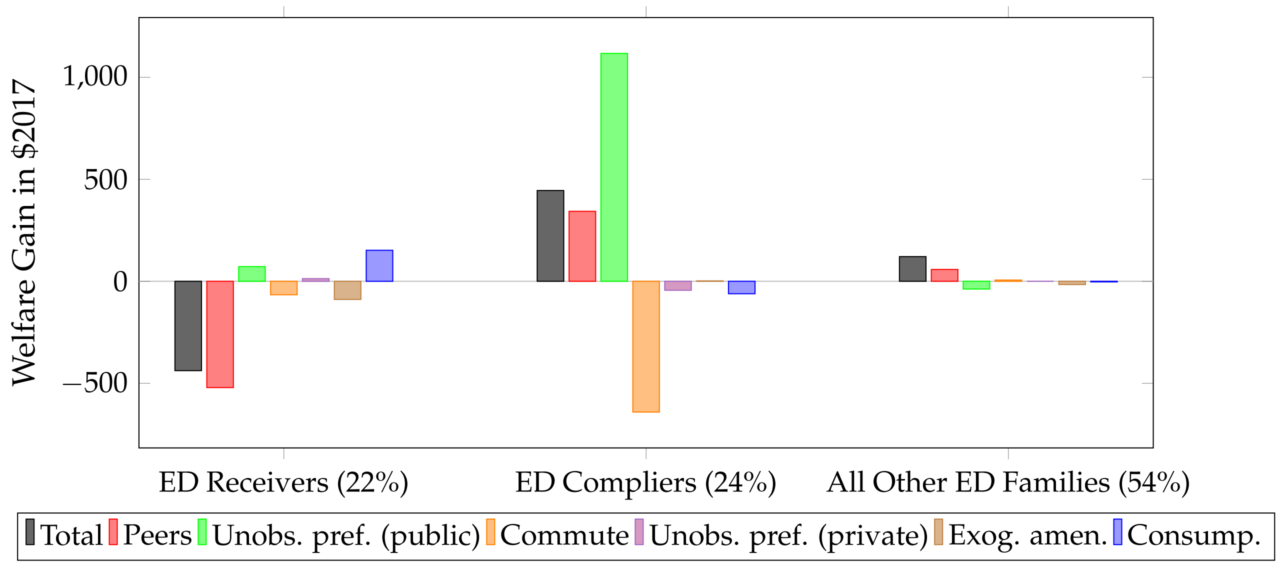
On the one hand, as one would expect, ED compliers experience a sizeable increase in

Figure 3: Decomposition of Welfare Effects for the School Choice Policy

(a) Decomposition for all families, and by ED status



(b) Decomposition for a partition of ED families



The top chart illustrates the welfare changes induced by the school choice expansion policy for all families (left), ED families (center), and non-ED families (right). The bottom chart illustrates the welfare changes induced by the school choice expansion policy for a partition of ED families—those who live in receiving neighborhoods at baseline (or *Receivers* for short, left), those who live in sending neighborhoods at baseline and are induced by the policy to apply to receiving schools (*Compliers*, center), and all other ED families (right). For each set of families, the first bar on the left (black) shows the total average change in welfare resulting from the policy. The second to seventh (colored) bars show the decomposition of this average welfare change into the changes induced by the policy in the following determinants of family utility (from left to right): school peers ($\bar{a}$), idiosyncratic preference received from attending a public school if doing so ($\sigma_S \varepsilon_s$), commuting costs (τ), net value from attending a private school if doing so ($\chi_{w,n} + \sigma_P \varepsilon_{\mathcal{P}_n}$), exogenous neighborhood amenities ($\alpha_{0n} + \alpha_{1n} \log(w) + \sigma_N \varepsilon_n$), and value from consumption ($u_{w,n}$). The formal decomposition of welfare into its components is shown in Equation D-1.

Table 4: Outcomes of the School Choice Policy for Receiving Families

	Non-ED Receivers	ED Receivers
Change in average school peers (% std)	-11	-38
Change in share attending private school (pp)	5.0	1.1
Share locating outside receiving nbhds (%)	5.7	7.7

The tables shows how the school choice expansion policy affects attended school quality, private school attendance, and residential choice for families living in receiving neighborhoods at baseline, by ED status.

welfare, primarily coming from access to public schools with better peers and for which they have a stronger idiosyncratic preference (due the increase in the size of their choice set). However, the costs associated with the longer commute to receiving schools (averaging +3.14 miles) eliminates almost half of the welfare gains experienced by compliers.[21] On the other hand, ED families who live in receiving neighborhoods at baseline experience a net welfare loss that originates from the significant decline in their average school peer composition after the inflow of new children. In that respect, ED receivers fare more similarly to the average non-ED family than to target ED families. In fact, ED receivers experience an even larger decline in peer quality than non-ED receivers, as shown in Table 4. This result is due to different exposure to the policy and the adoption of different behavioral responses to offset its negative consequences.

The difference in *exposure* between ED and non-ED receivers lies in the geography of neighborhood income and it is illustrated in Figure 4. Receiving schools that are farther away from sending neighborhoods systematically enroll fewer and higher-skilled ED compliers—symmetrically, the easternmost neighborhoods located farther away from the receiving schools experience smaller take-up rates and a stronger selection on skills. These results are intuitive, as commuting costs—even in the presence of transportation—represent a screening mechanism for families with lower-skilled children. The westernmost part of the city is the precisely where the most affluent families live, while ED receivers are primarily located in the center and center-north of the county, closer to lower income-income neighborhoods (see Figure 1 in Section 3).

High- and low-income families in receiving neighborhoods also differ in terms of *responses* to the policy. To mitigate their welfare loss from decreased peer quality, non-ED

[21] Increased commuting by school transportation also translate into increased expenditures for the school district. The policy results in a 21.8 percent increase in overall commuting distance made via school transportation, which corresponds to an additional school transportation expenditure of $678.30 per child. Further details on this calculation are provided in Appendix D.3.

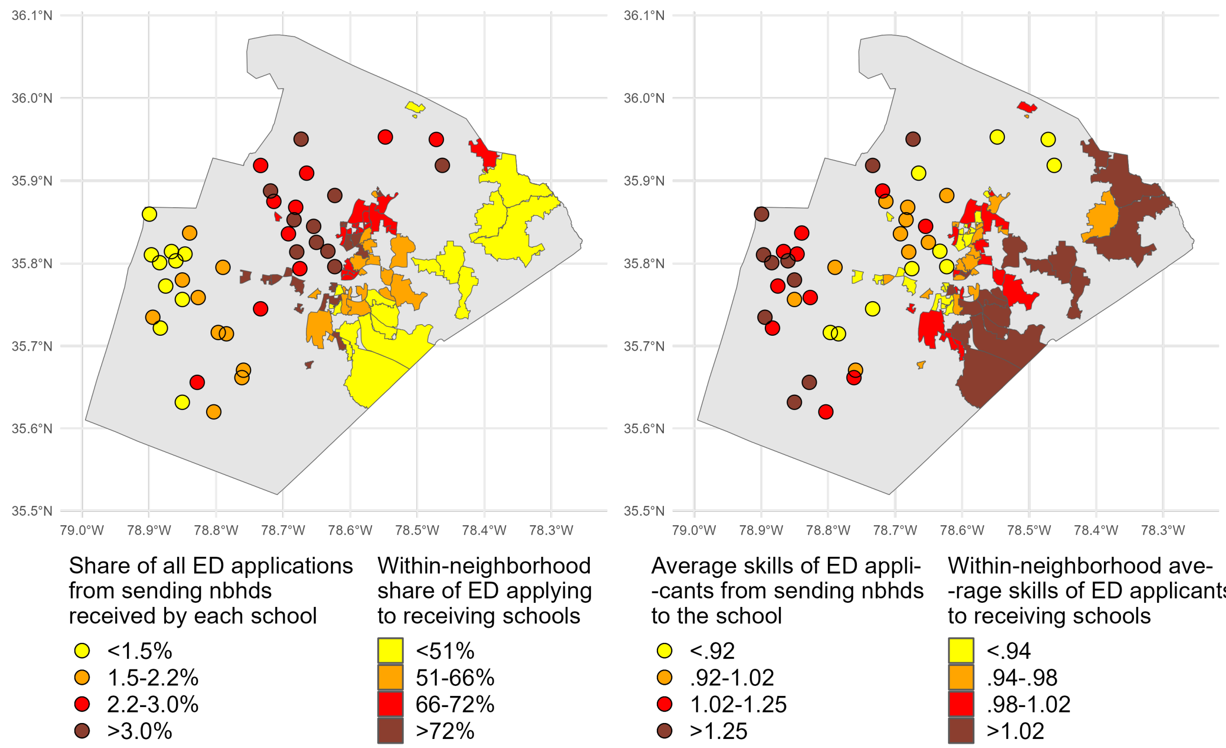
receivers turn to private schools (+5 pp, see Table 4). To do so, they pay the cost of attending these schools and, as private schools tend to be located further away than public schools and not to offer transportation, they also suffer extra commuting costs. The losses non-ED families are willing to incur from attending and traveling to private schools as an insurance against a decrease in school quality are symptomatic of their estimated strong preference for peer quality. In contrast, private school attendance among ED families only increases by one percentage point. Instead ED families who live in receiving neighborhoods at baseline turn to other public school options that offer a better idiosyncratic match rather than access to higher-quality peers. As an effort to ease access to these alternative public schools, 7.7 percent of them choose to locate outside of the receiving neighborhoods. The alternate public options are typically closer than private schools and provide transportation, yielding lower additional commuting costs.

We acknowledge that the policy we consider in this section represents just one possible way of expanding school choice. In Appendix D-1, we further illustrate the welfare effects derived from two alternative policies, each distinct from the one detailed here with respect to the targeted sending neighborhoods. In particular, we alternatively offer receiving schools as options with transportation to either of two sets of neighborhoods according to their share of ED families (below 31 percent or between 31 and and 60 percent).[22] We find that while welfare changes are overall small, less progressive policies translate into higher welfare gains since they target higher ability children, entail lower commuting cost (the sending neighborhoods being closer to receiving schools), and trigger a smaller flight of receiving families into private schools.[23] Last, we check the robustness of our conclusions to adopting an alternate welfare function, namely a utilitarian one. We formally show that utilitarian welfare gains are the sum of changes in our money-metric measure, that captures the efficiency gains from the policy, and changes in an additional term that captures the benefit from redistribution. In all three versions of school choice expansion, the sign of the aggregate utilitarian welfare gains is the same as our money-metric measure.

[22] These cutoffs are chosen so that the measure of families in sending neighborhoods is constant across all three versions of the policy.

[23] Given the geographic concentration of compliers in receiving schools located closer to the city center (Figure 4) one might wonder whether the negative welfare impact on receiving ED families could be reduced by restricting receiving schools to be those overwhelmingly attended by non-ED families, on the westernmost side of the city. However, the increased home-school distance would result in a much lower take-up rate and negligible overall welfare gains for ED families.

Figure 4: Spatial Heterogeneity in Exposure and Take-up for the School Choice Policy



The figure shows spatial patterns in the take-up of the school choice expansion policy. In the left map, receiving schools (dots) are colored as a function of the share of all applications from ED families in sending neighborhoods they receive; and sending neighborhoods are colored as a function of the share of their ED population who applies to receiving schools. In the right map, receiving schools (dots) are colored as a function of the average skill level of the ED applicants they receive from sending neighborhoods they receive; and sending neighborhoods are colored as a function of the average skill level of their ED applicants to receiving schools.

5.2 Housing Vouchers

As our second policy counterfactual, we analyze housing vouchers, a people-based policies that received considerable attention among both academics and policy-makers (Kling et al., 2007; Chetty et al., 2016; Aliprantis and Richter, 2020; Bergman et al., Forthcoming). Our policy counterfactual provides housing vouchers to low-income families enabling them to reside within the catchment area of high-quality schools. Moreover, housing vouchers address one of the shortcomings highlighted in the previous section regarding the school choice expansion policy. In that analysis, we found that commuting costs significantly reduce the welfare gains for those taking advantage of the policy, and concentrate the inflow into schools that are in the city center where relatively more ED families live. Housing vouchers offset this spatial barrier by allowing low-income families to reside in neighborhoods near high-quality schools, including potentially in the more affluent areas at the outskirts of the city.

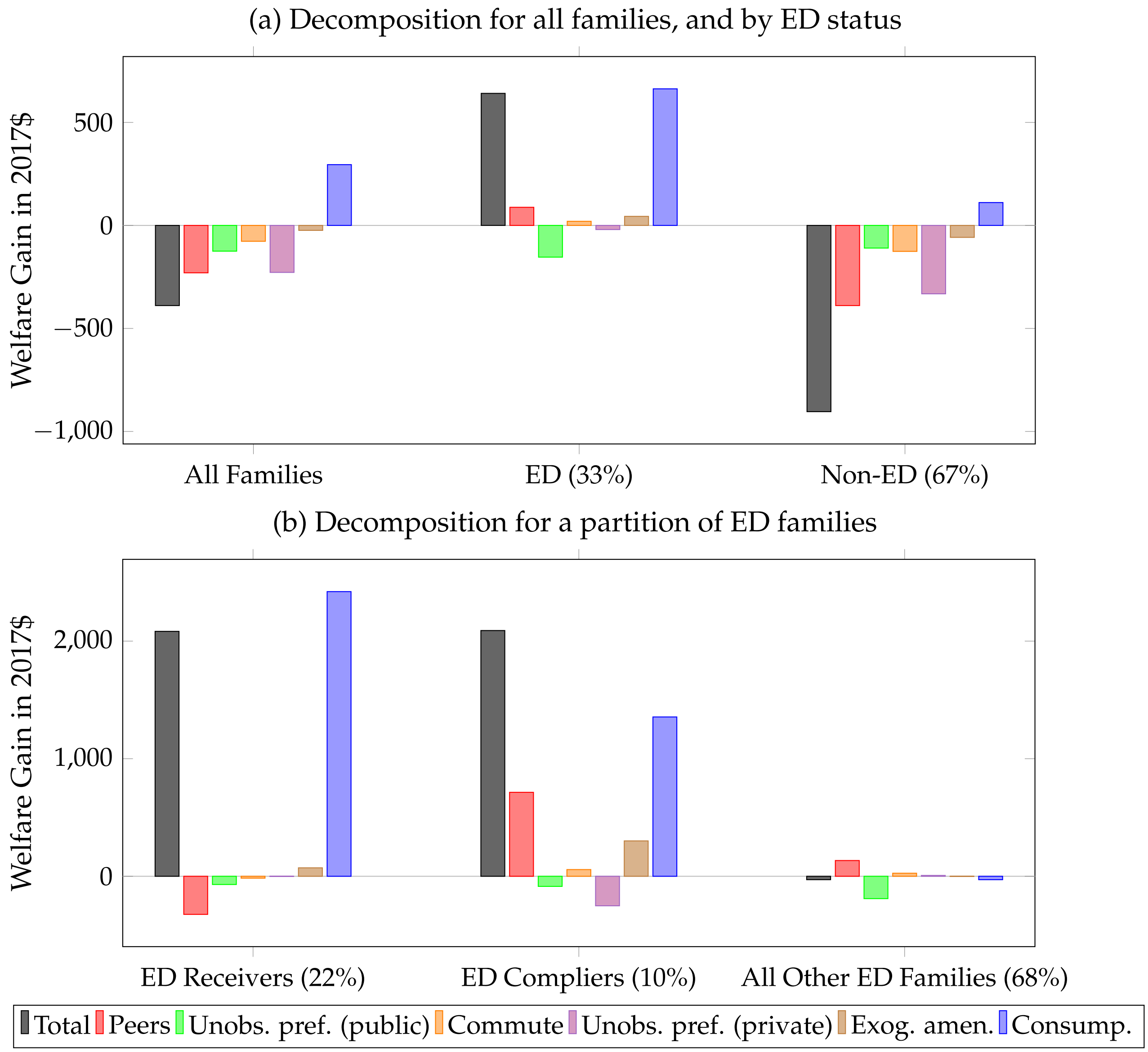
Policy Design. To make the two counterfactual policies comparable, we offer housing vouchers to ED families under the condition that they choose to live within the catchment area of one of the 40 high-quality receiving schools identified in the school choice expansion policy. We design the housing voucher to ensure that every eligible family can afford housing of a typical size for families in their income bracket–that is, the average house size in the baseline equilibrium–while spending only 25 percent of their income towards housing. The fact that the dollar value of the voucher scales with income dampens the otherwise negative selection on income, conditional on being an ED family, induced by alternative flat payments. In addition, the dollar value of the voucher is highest in expensive neighborhoods, making those more accessible to potential takers.

Results. We present our results in a similar way as we did for the school choice expansion counterfactual, and focus our comments on the comparison between the two policies. Average welfare effects and their decomposition are shown in Figure 5. Just like in Figure 3, the top chart (Figure 5(a)) shows the decomposition of welfare effects for all families, as well as for ED and non-ED families separately. The bottom chart (Figure 5(b)) shows the decomposition of welfare changes for a partition of ED families: those families who live in receiving neighborhoods at baseline (again, for short, *ED receivers*); those ED families who do not live in receiving neighborhoods at baseline but do under the policy (*compliers*); and all other ED families. The voucher policy exhibits an average welfare loss of $389, as depicted by the leftmost black bar in Figure 5(a). This outcome is quantitatively very close to the impact induced by the school choice expansion policy.[24]

On the distributional front, the main difference between the two policies lies in the much larger gains experienced by ED families and the much larger losses experienced by non-ED families under the voucher policy relative to the school choice expansion policy (black bars). Focusing on ED families first, the larger gains for compliers are mostly driven by the 'mechanical' effects of the voucher on consumption of housing and the numeraire good. Thanks to the voucher, compliers also locate in neighborhoods with higher amenities. Differently from the school choice expansion policy, idiosyncratic preferences over public schools play a minimal role here, as it was a feature associated with broadening their school choice set. In addition, the absence of significant welfare loss from com-

[24]The (unconditional) average monetary cost of the voucher amounts to $460 per family, or about two-third the cost of expanding school choice.

Figure 5: Decomposition of Welfare Effects for the Voucher Policy



The top chart illustrates the welfare changes induced by the housing voucher policy for all families (left), ED families (center), and non-ED families (right). The bottom chart illustrates the welfare changes induced by the housing voucher for a partition of ED families—those who live in receiving neighborhoods at baseline (or *Receivers* for short, left), those who do not live in receiving neighborhoods at baseline but do under the policy (*Compliers*, center), and all other ED families (right). For each set of families, the first bar on the left (black) shows the total average change in welfare resulting from the policy. The second to seventh (colored) bars show the decomposition of this average welfare change into the changes induced by the policy in the following determinants of family utility (from left to right): school peers ($\gamma\bar{a}$), idiosyncratic preference received from attending a public school if doing so ($\sigma_S\varepsilon_s$), commuting costs ($\kappa\tau$), net value from attending a private school if doing so ($\chi_{w,n} + \sigma_P\varepsilon_{\mathcal{P}_n}$), exogenous neighborhood amenities ($\alpha_{0n} + \alpha_{1n}\log(w) + \sigma_N\varepsilon_n$), and value from consumption ($u_{w,n}$). The formal decomposition of welfare into its components is shown in Equation D-1.

Table 5: Outcomes of the Voucher Policy for Receiving Families

	Non-ED Receivers	ED Receivers
Change in average school peers (% std)	-16	-26
Change in share attending private school (pp)	3.9	0.4
Share locating outside receiving nbhds (%)	4.2	0.5

The tables shows how the housing vouchers policy affects attended school quality, private school attendance, and residential choice for families living in receiving neighborhoods at baseline, by ED status.

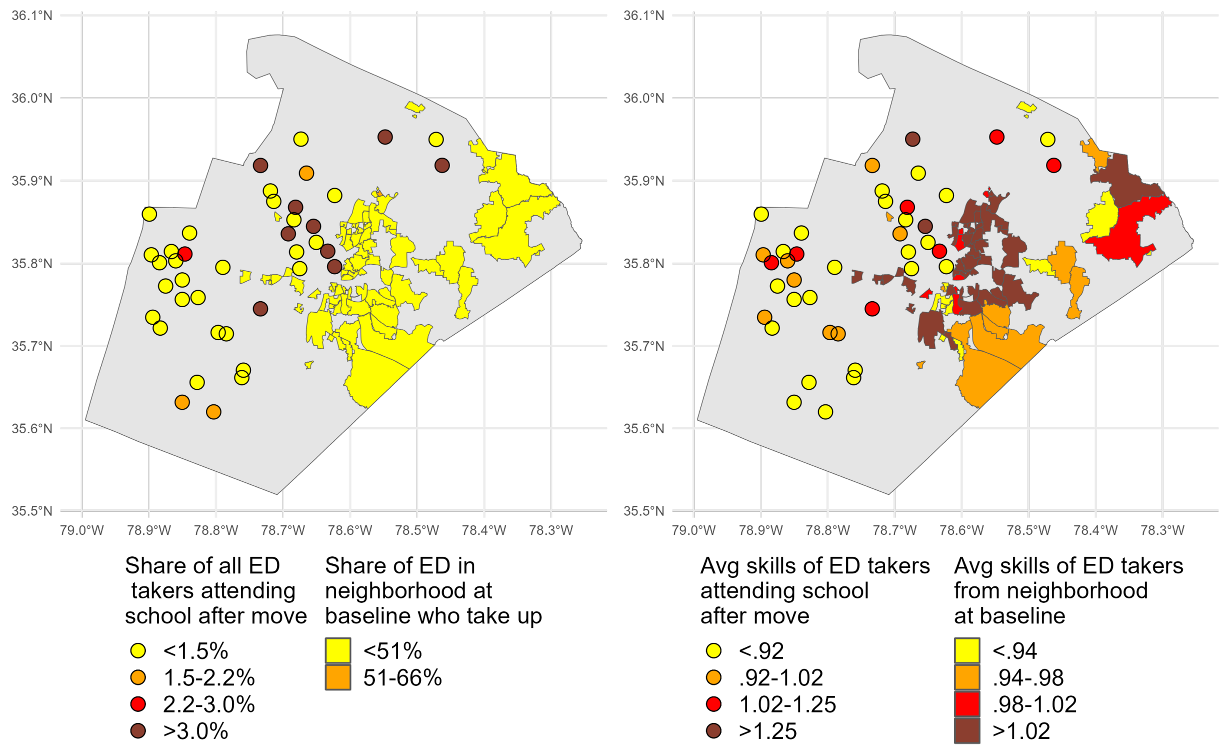
muting costs stems from the fact that compliers now reside in receiving neighborhoods rather than commuting from distant sending neighborhoods.

ED families residing in receiving neighborhoods at baseline experience significant welfare gains under the voucher policy, in contrast to the losses from the school choice expansion policy. Under both policies, they experience lower-skilled school peers. However, under the voucher policy, they qualify for the housing voucher if they choose to remain in receiving neighborhoods. Only half a percent of them opt to live outside of receiving neighborhoods under the voucher policy, compared to 7.7 percent under the school choice expansion (compare Table 5 with Table 4). The increase in consumption value more than compensates for the loss from peers, thereby explaining their overall welfare gains.

The gap in peer-quality losses between ED and non-ED receivers is significantly lower under the voucher policy compared to the school choice policy. Under the latter, the peer-quality losses were -38 and -11 percent of a standard deviation for ED and non-ED receivers, respectively, while they are equal to -26 and -16 percent under the voucher policy. To understand this result, Figure 6 illustrates the average quality of ED compliers attending each receiving school under the voucher policy (right panel) and the distribution of ED compliers across different receiving schools (right panel). Comparing Figure 6 to Figure 4 reveals that the magnitude and quality of the influx of ED compliers are more evenly distributed across receiving schools under the voucher policy than under the school choice expansion policy. As a consequence, non-ED families experience a substantial decline in the quality of their school peers, hence a larger welfare loss under the voucher policy.

The east-to-west gradient observed in both panels of Figure 4 is virtually absent in Figure 6. This is because, under the voucher policy, compliers live in receiving neighborhoods instead of commuting from sending neighborhoods. As a consequence, west-

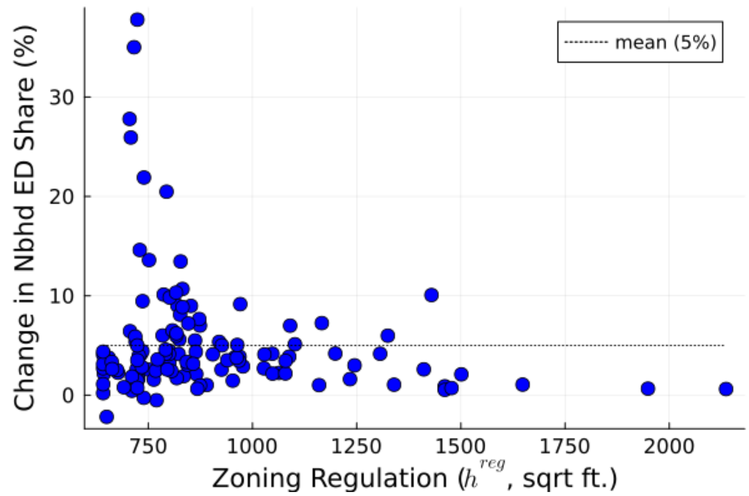
Figure 6: Spatial Heterogeneity in Exposure and Take-up for the Voucher Policy



The figure shows spatial patterns in the take-up of the housing voucher policy. In the left map, receiving schools (dots) are colored as a function of the share of all applications from ED families in sending neighborhoods they receive; and sending neighborhoods are colored as a function of the share of their ED population who applies to receiving schools. In the right map, receiving schools (dots) are colored as a function of the average skill level of the ED applicants they receive from sending neighborhoods they receive; and sending neighborhoods are colored as a function of the average skill level of their ED applicants to receiving schools.

ernmost receiving schools, located in neighborhoods with lower baseline shares of ED families, receive fewer and more positively selected compliers (in terms of skills) compared to other receiving neighborhoods under the school choice expansion policy. The absence of commuting costs then reduces the spatial heterogeneity in exposure to compliers across receiving neighborhoods. Yet, a gap persists in policy exposure between ED and non-ED receiving families. Figure 7 sheds light on one of its determinants by revealing housing regulation as a barrier to access of ED compliers into receiving neighborhoods. Families using vouchers tend to relocate toward areas with lower housing regulation. Conversely, neighborhoods characterized by stringent zoning regulations, typically the domains of higher-income households, witness negligible influx of voucher-holding families who would have to complement the voucher with their own resources in order

Figure 7: Zoning and Change in Neighborhood Composition under the Voucher Policy



The figure shows how receiving neighborhoods' shares of ED families change as the voucher policy is implemented as a function of their level of zoning regulation, h_n^{reg} (blue dots). Zoning regulation is defined as the minimum housing size (in squared feet) allowed in a neighborhood. The dashed line represents the average change in shares of ED families in the receiving neighborhoods.

to satisfy the minimum size requirement. This result highlights the role of zoning regulations as a shield for higher-income families in receiving neighborhoods against the effects of the voucher policy.

6 Conclusion

We build and estimate a spatial equilibrium model of neighborhood and school choice that accounts for the key institutional determinants of educational access in the United States: school catchment areas, school transportation provision, and zoning regulations. The model is identified by the observed empirical pattern of residential sorting and school choice, both in the cross-section and using quasi-experimental evidence on how demand for neighborhoods and schools respond to changes in the composition of school catchment areas and in transportation provision.

The model sheds light on the trade-offs involved in the design of large-scale policy interventions that aim to reduce inequality of educational opportunities. On the one hand, we find that common instances of such policies like school-choice expansion and housing vouchers are moderately effective at improving their beneficiaries' outcomes. On the other hand, families who have already made significant investments in their children's education by residing in affluent neighborhoods experience a decline in welfare. Low-income families willing to pay high housing costs in order to access a better composition of peers are particularly hurt by a school choice expansion as their geographic proximity to compliers overwhelmingly exposes them to an additional inflow into their schools. Relocation to other neighborhoods or increased attendance of private schools are imperfect and costly insurance mechanism that families with high valuation for school quality implement in order to dampen the negative impact of the policy. When evaluated in dollar-equivalent terms, both policy counterfactuals deliver a substantial welfare loss.

Our analysis can be expanded along several interesting dimensions. First, we abstracted from dynamic considerations in families' choice of schools and neighborhoods, which might affect the estimation of families' preferences whenever school portfolios are expected to change over time and moving cost vary with distance. Second, our model could be enriched with other possible responses of parents to policies aimed at integrating heterogeneous peer groups within educational settings, such as seeking special classroom assignment for their children (Crema, 2023) or adopting parenting styles at home that control their children's social interactions (Agostinelli et al., 2023). In both instances, these mechanisms would impact the effective exposure of children to the change in school peer effects induced by the proposed policies. Third, we see the design of optimal local institutions, such as school boundaries and zoning restrictions, as natural—albeit challenging—improvements on our counterfactual analysis. We see all these aspects as fruitful avenues for future research.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag Pathak, "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match," *American Economic Review*, 2017, *107* (12), 3635–3689.

Abdulkadiroğlu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters, "Do Parents Value School Effectiveness?," *American Economic Review*, May 2020, *110* (5), 1502–39.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Annual Review of Economics*, 2018, *86* (2), 391–444.

— **and** — , "Revealed Preference Analysis of School Choice Models," *Annual Review of Economics*, 2020, *12* (1), 471–501.

Agostinelli, Francesco, Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti, "It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects," 2023. NBER Working Paper 27050.

Ahlfeldt, Gabriel, Stephen Redding, Daniel Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Aliprantis, Dionissi and Daniel R. Carroll, "Neighborhood dynamics and the distribution of opportunity," *Quantitative Economics*, 2018, *9* (1), 247–303.

— **and Francisca G.-C. Richter**, "Evidence of Neighborhood Effects from Moving to Opportunity: LATEs of Neighborhood Quality," *The Review of Economics and Statistics*, 10 2020, *102* (4), 633–647.

Allende, Claudia, "Competition Under Social Interactions and the Design of Education Policies," 2020. Working Paper.

Almagro, Milena and Tomás Domínguez-Iino, "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," 2022. Unpublished Manuscript.

Avery, Christopher and Parag Pathak, "The Distributional Consequences of Public School Choice," *American Economic Review*, 2021, *111* (1), 129–152.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

Benabou, Roland, "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics*, 1993, *108* (3), 619–652.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer, "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice," *American Economic Review*, accepted Forthcoming. Econofact Podcast on MTO and CMTO Interview on CMTO research The animated stick-figure summary of CMTO Vox 2023 Coverage Vox 2019 Coverage Also available as NBER Working Paper No. 26164.

Black, Sandra E., "Do Better Schools Matter? Parental Valuation of Elementary Education," *The Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

Calsamiglia, Caterina, Chao Fu, and Maya Güell, "Structural Estimation of a Model of School Choices: The Boston Mechanism versus Its Alternatives," *Journal of Political Economy*, 2020, *128* (2), 642–680.

Campos, Christopher and Caitlin Kearns, "The Impact of Public School Choice: Evidence from Los Angeles's Zones of Choice," *The Quarterly Journal of Economics*, 2023.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Chyn, Eric and Diego Daruich, "An Equilibrium Analysis of the Effects of Neighborhood-based Interventions on Children," 2022. Unpublished Manuscript.

Crema, Angela, "The Effects of Charter Schools on Classroom Segregation," 2023. Working Paper.

Cullen, Julie Berry, Brian A Jacob, and Steven Levitt, "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, 2006, *74* (5), 1191–1230.

Davila, Eduardo and Andreas Schaab, "Welfare Accounting," 2023. Unpublished Manuscript.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger, "School Choice, School Quality, and Postsecondary Attainment," *American Economic Review*, March 2014, *104* (3), 991–1013.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dinerstein, Michael and Troy Smith, "Quantifying the Supply Response of Private Schools to Public Policies," *American Economic Review*, 2021, *111* (10), 3376–3417.

Dur, Umut, Robert Hammond, and Thayer Morrill, "Identifying the Harm of Manipulable School-Choice Mechanisms," *American Economic Journal: Economic Policy*, 2018, *10* (1), 187–213.

_ , _ , Matthew Lenard, Melinda Morrill, Thayer Morrill, and Colleen Paeplow, "The Attraction of Magnet Schools: Evidence from Embedded Lotteries in School Assignment," 2022. Working Paper.

Durlauf, Steven N., "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1996, *1* (1), 75–93.

Eckert, Fabian and Tatjana Kleineberg, "Saving the American Dream? Education Policies in Spatial General Equilibrium," 2021. Unpublished Manuscript, Yale University.

Epple, Dennis and Holger Sieg, "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, 1999, *107* (4), 645–681.

_ and Richard Romano, "Neighborhood Schools, Choice, and the Distribution of Educational Benefits," in Caroline M. Hoxby, ed., *The Economics of School Choice*, University of Chicago Press, 2003, pp. 227–286.

Fernandez, Raquel and Richard Rogerson, "Income Distribution, Communities, and the Quality of Public Education," *The Quarterly Journal of Economics*, 1996, *111* (1), 135–164.

Fogli, Alessandra and Veronica Guerrieri, "The End of the American Dream? Inequality and Segregation in US Cities," 2018. Unpublished Manuscript, University of Chicago.

Galiani, Sebastian, Alvin Murphy, and Juan Pantano, "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment," *American Economic Review*, 2015, *105* (11), 3385–3415.

Hastings, Justine S. and Jeffrey M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments*," *The Quarterly Journal of Economics*, 11 2008, *123* (4), 1373–1414.

_ , Thomas J. Kane, and Douglas O. Staiger, "Heterogeneous Preferences and the Efficacy of Public School Choice," 2009. Unpublished Manuscript.

Heblich, Stephan, Stephen Redding, and Daniel Sturm, "The Making of the Modern Metropolis: Evidence from London," *The Quarterly Journal of Economics*, 2020, *135* (4), 2059–2133.

Heckman, James J. and Stefano Mosso, "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, 2014, *6* (1), 689–733.

Heckman, James, Rodrigo Pinto, and Peter Savelyev, "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 2013, *103* (6), 2055–86.

Hill, Darryl V., Rodney P. Hughes, Matthew A. Lenard, David D. Liebowitz, and Lindsay C. Page, "New schools and new classmates: The disruption and peer group effects of school reassignment," *Economics of Education Review*, 2023, *92*, 102316.

Hsiao, Allan, "Educational Investment in Spatial Equilibrium: Evidence from Indonesia," 2023. Unpublished Manuscript.

Jackson, Kirabo, "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation," *Journal of Labor Economics*, 2009, *27* (2), 213–256.

Kapor, Adam, Christopher Neilson, and Seth Zimmerman, "Heterogeneous Beliefs and School Choice Assignment Mechanisms," *American Economic Review*, 2020, *100* (5), 1274–1315.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kuminoff, Nicolai V. and Jaren C. Pope, "Do "capitalization effects" for public goods reveal the public's willingness to pay?," *International Economic Review*, 2014, *55* (4), 1227–1250.

Laverde, Mariana, "Distance to Schools and Equal Access in School Choice Systems," 2023. Unpublished Manuscript.

Luflade, Margaux, "The Value of Information in Centralized School Choice Systems," 2018. Unpublished Manuscript.

Nechyba, Thomas, "Mobility, Targeting, and Private-School Vouchers," *American Economic Review*, 2000, *90* (1), 130–146.

Parcel, Toby L. and Andrew J. Taylor, *The End of Consensus: Diversity, Neighborhoods, and the Politics of Public School Assignments*, Chapel Hill, NC: University of North Carolina Press, 2015.

Park, Minseon and Dong Woo Hahm, "Location Choice, Commuting, and School Choice," 2023. Working Paper.

Redding, Stephen J and Daniel M Sturm, "Neighborhood Effects: Evidence from Wartime Destruction in London," *Working Paper*, 2024.

Redding, Stephen J. and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58.

_ **and Matthew A. Turner**, "Chapter 20 - Transportation Costs and the Spatial Organization of Economic Activity," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier, 2015, pp. 1339–1398.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh, "Estimating the general equilibrium benefits of large changes in spatially delineated public goods," *International Economic Review*, 2004, *45* (4), 1047–1077.

Singleton, John, "Incentives and the Supply of Effective Charter Schools," *American Economic Review*, 2019, *109* (7), 2568–2612.

Tsivanidis, Nick, "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's Transmilenio," 2019. Unpublished Manuscript.

A Data appendix

A.1 Additional institutional details

The Wake County Public School System (WCPSS) is the county-wide school district covering Wake County, North Carolina, which is the county of the state capital, Raleigh. The WCPSS was, in 2019–20, the fourteenth largest school district in the United States, with more than 161,000 students. The geography of the county and the location of the public elementary schools open during our sample period (2013-14 to 2016-17) are shown in Figure 1(a).

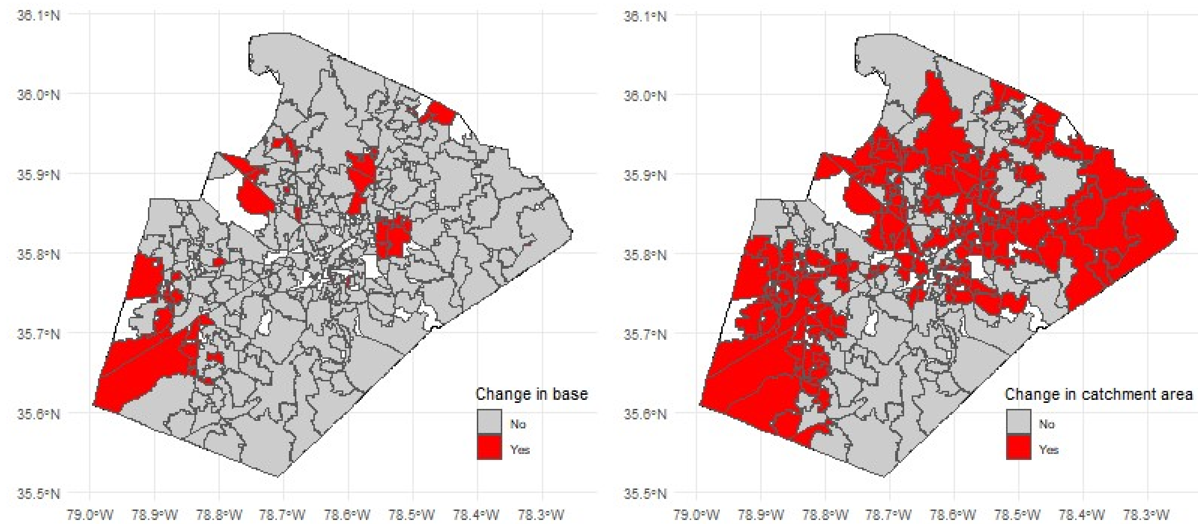
A.1.1 Public school choice in the WCPSS

Each address in Wake County is associated with a *base* school at which the child is guaranteed a seat and transportation. The school district offers two main ways for parents to have their child attend a public school other than their base: magnet programs and calendar transfers, each of which we describe below —however, when we bring our structural model to the data, the two types of options are not differentiated and are pooled under the umbrella category of *option* school.

Historically, from the creation of the district in 1976 until 2000, the student assignment policy was driven by the goal of promoting racial diversity in schools. Residential addresses were assigned to base schools so that each school would have 15–45 percent of Black students. Magnet schools were created as a second instrument to facilitate racial integration in schools: a number of urban schools were endowed with special curricula (e.g., arts, foreign languages) expected to attract white suburban students. Starting from the 2000–01 academic year and until 2011–12, the WCPSS moved from the goal of ensuring racial diversity in schools to that of ensuring socioeconomic balance. Assignments of addresses to base schools was then supposed to serve the goal that no school had more than 40 percent of students eligible for free or reduced-price lunch (that is, economically disadvantaged, or *ED*) nor more than 25 percent of students below the state's reading standards for their grade. While socioeconomic balance in schools was a target for the school board until the early 2010s, pressure to accommodate population growth across the county[25] has been the main driver of school reassignments, as illustrated by this quote from Parcel and Taylor (2015, p. 53) who said reassignment "from school to

[25] Over the 2000–10 decade, the public school population in the WCPSS increased from about 95,000 to more than 140,000.

Figure A-1: Changes in Base Schools and Catchment Areas

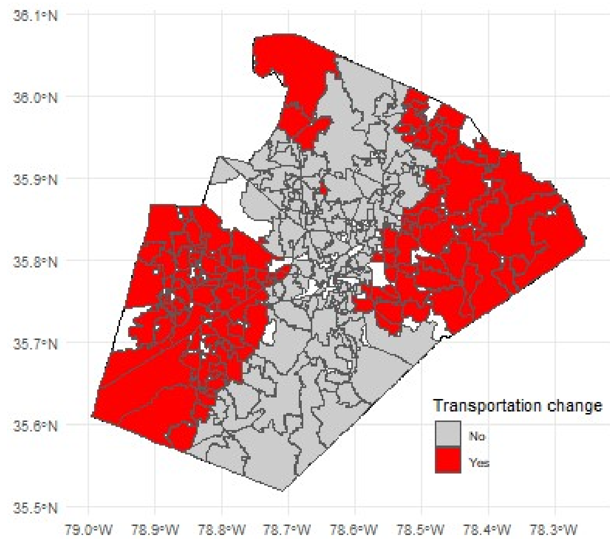


The map on the left shows in red all neighborhoods that experienced a change of base school (that is, were reassigned to another base school) between 2013-14 and 2016-17. When a neighborhood is reassigned to another base school, the catchment area of both its initial and its final base schools change—as the neighborhood is removed from the former and added to the latter. This means that the neighborhoods which remained assigned to the initial base school and those that were already assigned to the final base school experience a change in the their catchment area, although they do not change base schools. Taking this into account, the map on the right shows in red the neighborhoods experiencing changes in their base school *or* in their catchment area over the sample period. Areas left uncolored are out of our sample and essentially non-residential.

school [was] because of population growth, and that is what it was. The busing was not intended primarily for diversity but just to fill in . . . schools." Figure A-1 illustrates changes in catchment areas over the sample period. The map on the left shows in red all neighborhoods that experienced a change of base school (that is, were reassigned to another base school) between 2013-14 and 2016-17. When a neighborhood is reassigned to another base school, the catchment area of both its initial and its final base schools changes—as the neighborhood is removed from the former and added to the latter. This means that the neighborhoods which remained assigned to the initial base school and those that were already assigned to the final base school experienced a change in the their catchment area, although they did not change base schools. Taking this into account, the map on the right shows in red the neighborhoods experiencing changes in their base school or their catchment area over the sample period.

Magnet schools are the main instrument of public school choice in the WCPSS. In our period of interest, the WCPSS had 26 magnet schools at the elementary school level. Based on their residential address, parents can apply to a subset of these programs for their

Figure A-2: Changes in Transportation Provision to Option Schools



The figure shows the neighborhoods experiencing a change in transportation provision to at least one of their options over the sample period. That is, a neighborhood n is colored in red on the map if and only if there is at least one school s that is in neighborhood n's portfolio a option for n for two consecutive years in the sample period *and* transportation provision between n and s changes across these two years. Areas left uncolored are out of our sample and essentially non-residential.

child. Also based on their residential address, parents may or may not be offered school transportation to the magnet school.[26] Magnet choice set and transportation provision do not only change cross-sectionally, they also change over time during the period of interest, with several magnets expanding and/or changing their transportation provision. Figure A-2 provides an illustration of transportation changes over time. In red, it shows the neighborhoods experiencing a change in transportation provision for a school in their portfolio within the sample period. That is, a neighborhood n is colored in red on the map if and only if there is at least one school s that is in neighborhood n's portfolio as an option for n for two consecutive years and transportation provision between n and s changes across these two years—in other words: $s \in \mathcal{NT}_{n,t}$ and $s \in \mathcal{T}_{n,t+1}$, or $s \in \mathcal{T}_{n,t}$ and $s \in \mathcal{NT}_{n,t+1}$. In addition to these changes in *transportation provision*, all neighborhoods experienced *eligibility* changes, in the sense that new options were added to all choice sets over the sample period. Twenty-four magnet schools saw a change in the set of neighborhoods eligible to apply and/or in their transportation provision over the sample period (see bottom panel of Table A-3).

[26]Figure 1 in Dur et al. (2018) shows a screenshot of the online platform parents can use to apply; the fourth column in the table illustrates the variation of transportation provision across schools and residential addresses.

Calendar transfers allow students to attend a school running on a different calendar than their base school. Schools in the WCPSS operate following one of two calendars—the traditional September-to-June academic calendar or a year-round calendar designed as a response to the rapid population growth to allow schools to accommodate more students.[27] Each base school is paired with one school that operates on the other calendar to which families in the catchment areas can apply and to which transportation is provided.

Assignment to magnet and calendar options is centralized. It proceeds in two steps (first families can apply to magnets, then they may apply to calendar options), and can be summarized as follows:

- Magnet seats —"[Ninety] percent of magnet seats are assigned via the Boston Mechanism [up until the 2014-15 school year, and via the Deferred Acceptance algorithm since 2015-16.] For elementary schools, priority points at school s depend on whether the student's sibling will attend school s next year (highest priority), whether the student lives in a high-performing [area] based on historical test score data (second highest), and whether the student's base school is overcrowded (third highest). [Ties within coarse priority groups are broken using random lottery numbers.] . . . Finally, 10 percent of magnet seats are assigned through a pure lottery; specifically, a lottery that is independent of a student's priority points. The district introduced the 10 percent lottery to encourage more students to participate in the magnet application process." (Dur et al., 2018, p. 192).
- Calendar seats —Throughout our sample period, families may apply to the one alternate-calendar school associated to their base. For kindergarten-entry applicants, priority points at school s depend on whether the student's sibling will attend school s next year (highest priority), whether the student applies to attend a school with a calendar matching a next sibling's school calendar (second highest), and whether the student's base school is overcrowded (third highest). Ties within coarse priority groups are broken using random lottery numbers.

In the model, we make no distinction between calendar and magnet applications (all "options"). We assume that students can apply to at most one option school, and that assignment is made by pure lottery. A key feature of the Boston Mechanism is that students who rank a school first get higher priority for that school than applicants who rake the school

[27] In year-round schools, students are placed on four different tracks, each of them alternating year-round between nine weeks of class and three weeks of break. At any point in time, one of the four tracks is on break while the other three are in attendance, allowing the school to serve a larger number of students.

lower on their application list.[28] If, for each school, the number of first-choice applicants exceeds the number of seats, each applicant will only be considered for admission to their first choice, rendering all choices ranked below the first one irrelevant. In that regard, our single-application assumption is a reasonable approximation of the Boston Mechanism. While the Deferred Acceptance mechanism used starting 2015-16 does not prioritize first-rank applicants over others, it retains the incentive to strategize as the number of choices students can submit is capped (three to five choices allowed, out of a choice set of, on average, about 20 schools).

A.1.2 On the exogeneity of institutional changes

The identification of preferences for school quality relies on within-neighborhood variation over time, and requires the changes in school quality induced by changes in catchment areas to be unanticipated by households. Here, we provide evidence of the exogeneity of these institutional changes.

Changes in base schools' catchment areas. While the school board targeted socioeconomic balance in schools until the early 2010s, pressure to accommodate unequal population growth across the county has been the main driver of base school reassignments as illustrated by this quote from Parcel and Taylor (2015, p. 53): reassignment "from school to school [was] because of population growth, and that is what it was. The busing was not intended primarily for diversity but just to fill in . . . schools." In addition, while the fact that changes in catchment areas were likely was well-known to families over the period of interest (Parcel and Taylor, 2015), Hill et al. (2023, p. 7) argue that "the selection of any given geographic node for reassignment was, conditional on observable traits of the node, essentially random and not manipulable or anticipated by [neighborhood] residents. . . . As a result of the reassignment plan, geographically proximal and observationally similar [neighborhoods] were treated differently. Students from the same geographic area but different assignment nodes, who had been assigned to attend the same school in one year, would be assigned to attend different schools the following year." To further show that changes were not anticipated by households, we test whether policy-induced changes in base school quality (defined in Equation (4.2)) correlate with pre-trends in house prices. We estimate a regression similar to Equation (4.1), but using house-price changes in pre-policy years as the left-hand-side variable—namely, changes in average

[28] Dur et al. (2018, p. 192) note that the "WCPSS used the Boston Mechanism for the reason that Boston and many other districts used it: it is intuitive, easy to explain, and maximizes the number of students assigned to their reported first choice."

within-neighborhood house prices from $(t-2)$ to $(t-1)$, from $(t-3)$ to $(t-1)$, and from $(t-4)$ to $(t-1)$. Results are shown in Table A-1, and indicate that policy-induced changes in base school quality do not predict pre-policy changes in house prices.

Table A-1: Policy-induced changes in base school quality and pre-trends in house prices

	Change in (log) house prices		
	$(t-2)$ to $(t-1)$	$(t-3)$ to $(t-1)$	$(t-4)$ to $(t-1)$
Change in (log) base peer quality, $(t-1)$ to t	-0.008	-0.018	0.038
	(0.63)	(0.063)	(0.054)
Obs.	811	803	795
Sale-year F.E.	Yes	Yes	Yes

The table shows test results for whether policy-induced changes in base-school quality (defined by Equation (4.2)) correlate with changes in house prices in pre-policy years for different time horizons. We refer to policy-induced changes as the reassignments of school catchment areas taking place in the WCPSS between 2013-14 and 2016-17 (as illustrated in Figure A-1). Standard errors are clustered at the neighborhood level and reported in parentheses. *, **, and *** indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Changes in option school transportation provision. The reasons underlying changes in transportation provision are not as well documented in the literature (nor in the minutes of school board meetings) as those underlying changes in base schools' catchment areas. The argument for the exogeneity of policy changes can therefore not be made in the same way as it was for changes in base schools' catchment areas. Instead, to assess whether neighborhoods and schools chosen for the changes could be predicted based on their observable characteristics, we test whether transportation provision in year t correlate with pre-trends in within-neighborhood school enrollment shares. We estimate a regression similar to Equation (B-1), but using as the left-hand-side variable changes in the share (both overall and conditional on ED status) of students from neighborhood n attending school s—namely, changes in average within-neighborhood enrollment shares from $(t-1)$ to t, and from $(t-2)$ to $(t-1)$. Results are shown in Table A-2, and indicate that transportation provision does not predict past changes in enrollment shares.

A.2 Data Sources, Variables Construction, and Descriptive Statistics

Here, we provide details about data sources; construction of the student sample and key variables; as well as descriptive statistics.

Table A-2: Transportation provision and pre-trends in enrollment shares

	Change in within-neighborhood enrollment share			
	Overall		Conditional on ED	
	$(t-1)$ to t	$(t-2)$ to $(t-1)$	$(t-1)$ to t	$(t-2)$ to $(t-1)$
Transportation provided at t	-0.0001	-0.005^{*}	0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.004)
Obs.	15,947	10,061	12,651	7,901
School $\times$ Year F.E.	Yes	Yes	Yes	Yes
Neighborhood F.E.	Yes	Yes	Yes	Yes

This table shows test results for whether transportation provision for a given neighborhood to a given option school correlates with changes in within-neighborhood enrollment shares to that school in previous years. *Overall* enrollment share is used as outcome in the first two columns; share *conditional on being ED* is used as outcome in the last two columns. Standard errors are clustered at the neighborhood level and reported in parentheses. *, **, and *** indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

A.2.1 Students

Data source. Student-level data were obtained from the North Carolina Education Research Data Center[29] (NCERDC). The data show, for each year and each student enrolled in a North Carolina public school, the school the child is enrolled in and a set of demographic variables (gender, economically disadvantaged status, or *ED*). Starting in 2006, the data also show the student's residential census block group and (a noisy version of) residential coordinates. For all years, the data show end-of-grade test scores in math and reading for grades three to eight. For the school years 2013-14 through 2016-17, literacy test scores are also available for the beginning and the end of the academic year for kindergarten.

Final sample construction. The final estimation sample is obtained after three successive sample restrictions:

1. We restrict the sample to students enrolled in kindergarten in a Wake County public school in school years 2013–14 to 2016–17 and with the following information not missing for the kindergarten year: residential address, school attended, ED status, literacy assessment score.
2. After matching students to their neighborhood, we count the number of students assigned to each neighborhood, and exclude neighborhoods with fewer than ten

[29] https://childandfamilypolicy.duke.edu/research/nc-education-data-center/, accessed August 2021.

students over the four years of data. Our final sample consists of 312 neighborhoods.

3. Given their residential neighborhood and detailed administrative information about catchment areas, we are able to determine whether each student attends a school that was indeed in that student's choice set for his kindergarten school year. More precisely, we observe three sets of students: (1) children attending a school assigned to their neighborhood as a base or option school when they entered kindergarten; (2) children attending a school that is not in the choice set attached to their neighborhood in their kindergarten year, but assigned to their neighborhood as a base or option school within five years prior; and (3) children attending a school that was never part of the choice set attached to their address. We assume students of type (2) were grandfathered into attending the same school as an older sibling. For students of type (3), we check whether the mismatch between school attended and choice set can be explained by measurement error in residential address coordinates (introduced by the NCERDC to preserve privacy). If there is a neighborhood within one mile of their observed address that includes the attended school in its choice set for the student's kindergarten year, we infer that measurement error was the source of the mismatch, and we reassign the student to that neighborhood. If there is no such neighborhood, we exclude the student from the sample since his choice of school cannot be explained given the choice set.

Construction of the skills measure. We use beginning-of-the-year kindergarten literacy assessment test scores as a measure of a student's baseline skills. We use DIBELS (Dynamic Indicatord of Basic Early Literacy) and TRC (Text Reading and Comprehension) assessments from the mCLASS assessment system (`https://amplify.com/northcarolina/`) available in the NCERDC data, in particular: First Sound Fluency score, Letter Naming Fluency score, TRC score, as well as the Composite score. We combine the multiple assessment scores into one skill measure and address measurement error using Bartlett factor scores (Heckman et al., 2013). The obtained baseline skill measure is standardized by cohort. We define skill types, $a \in \{a_1, \ldots, a_{10}\}$, as the deciles of the continuous standardized baseline skill distribution.

A.2.2 Schools

Public schools: location, peer quality, capacity. Figure 1(a) in the main text shows the location of public elementary schools in Wake County, as well as the distribution of school peer quality across schools. School peer quality for school s and year t is measured as the average standardized skills for kindergarten students enrolled in school s in year t. All kindergarten students with non-missing kindergarten literacy assessment scores (and school attended) are used to compute school peer quality. For each school s, the number of seats reserved for option (as opposed to base) students in year t (q_{st}) is measured as follows. If s is oversubscribed in year t, then q_{st} is set to the count of kindergarten students attending s as an option in that year. If it is not, then q_{st} is set to the average number of kindergarten students attending s as an option in the years it is oversubscribed. While we do not observe individual applications, school-level information provided by the WCPSS indicates whether option schools are oversubscribed.

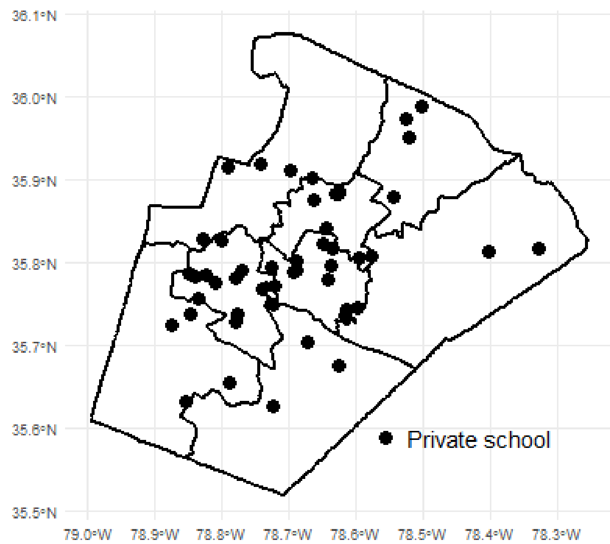
Private schools: data sources, location, and attendance. We use the bi-annual National Center for Education Statistics (NCES, https://nces.ed.gov/surveys/pss/pssdata.asp, accessed August 2021) Private School Survey for years 2013, 2015, and 2017 for information about the location of private schools offering Kindergarten in Wake County. We use data at the PUMA level from the ACS (5-year estimates, 2013-17) to construct the share of kindergartners attending private school and average income conditional on private versus public school attendance by PUMA. In particular, we use the variables "School attendance", "Grade level attending (detailed version)", "Public or Private school", "Total family income", and "PUMA" (downloaded from https://usa.ipums.org/usa/index.shtml, accessed August 2021). Figure A-3 shows the location of private schools offering kindergarten in Wake County between 2013-14 and 2016-17, as well as the boundaries of the eight PUMA partitioning Wake County. We set $\tau_{n\mathcal{P}_n}$ to the distance between the centroid of neighborhood n and its closest private school, and we construct, for each of the eight PUMAs constituting Wake County, the share of kindergartners attending private school by ED status.

A.2.3 Neighborhoods

Data sources. The WCPSS provided yearly data (maps) showing the assignment of residential addresses to base schools and menus of (public) options, as well as the availability of school transportation between each address and each option school.

Construction of neighborhoods. Each neighborhood n is characterized by a sequence of public school portfolio from school year 2013–14 to school year 2016–17: $\{ (\mathcal{B}_{n,t}, \mathcal{T}_{n,t}, \mathcal{NT}_{n,t}) \mid t = 2013, \ldots, 2016\}$, where $\mathcal{B}_{n,t}$ is the base school associated with n in year t, $\mathcal{T}_{n,t}$ is the set of option schools providing transportation to neighborhood n in year t, $\mathcal{NT}_{n,t}$ is the

Figure A-3: Private Schools Location and PUMA boundaries



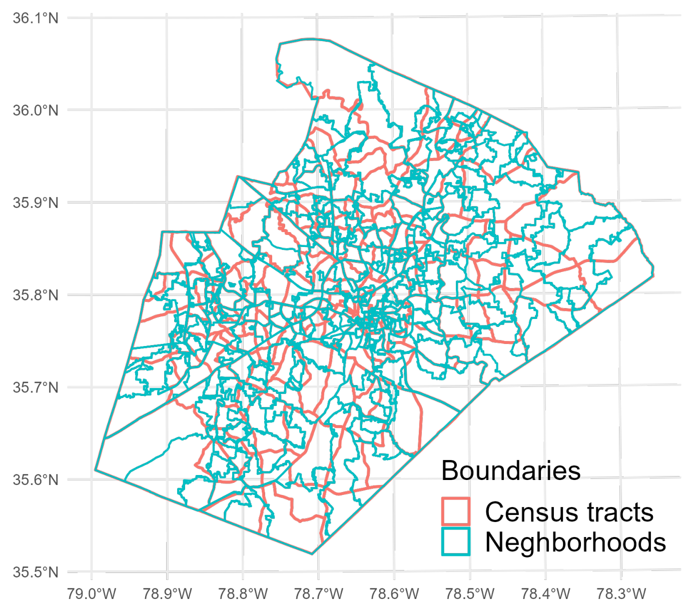
The map shows PUMA boundaries and the location of private schools offering kindergarten between Fall 2013 and Spring 2017 in Wake County.

set of option schools in the choice set of neighborhood t but not providing transportation. Neighborhood n is the union of all *contiguous* points with public school choice menu $\{ (\mathcal{B}_{n,t}, \mathcal{T}_{n,t}, \mathcal{NT}_{n,t}) \mid t = 2013, \ldots, 2016\}$. Formally, let us denote each residential address by its coordinates (x, y). Then, $(x, y) \in n$ only if the following three points are satisfied:

1. (x, y) has base school $\mathcal{B}_{n,t}$ in school year t, for each t.
2. $\mathcal{T}_{n,t}$ is the set of all public schools (except for $\mathcal{B}_{n,t}$) providing transportation to (x, y) in school year t.
3. $\mathcal{NT}_{n,t}$ is the set of all public schools open for application to (x, y) but not providing transportation to (x, y) in school year t.

In addition, we require neighborhoods to consist of fully contiguous points so if two regions share the same public school portfolio $\{ (\mathcal{B}_{n,t}, \mathcal{T}_{n,t}, \mathcal{NT}_{n,t}) \mid t = 2013, \ldots, 2016\}$ but are not touching, they make up distinct neighborhoods. Our definition of neighborhoods implies that at any point in our sample period, two addresses in the same neighborhood share the same portfolio of public schools—base and options with and without transportation. Conversely, two addresses can be in distinct neighborhoods for two reasons. Either their respective portfolios of public schools differ at some point in the sample period or, if they share the same portfolio of public schools, they are part of two geographic regions with no common border. Figure A-4 shows the obtained partition of Wake County into neighborhoods, with 2010 Census tract boundaries for comparison.

Figure A-4: Neighborhood Boundaries and 2010 Census tracts for comparison



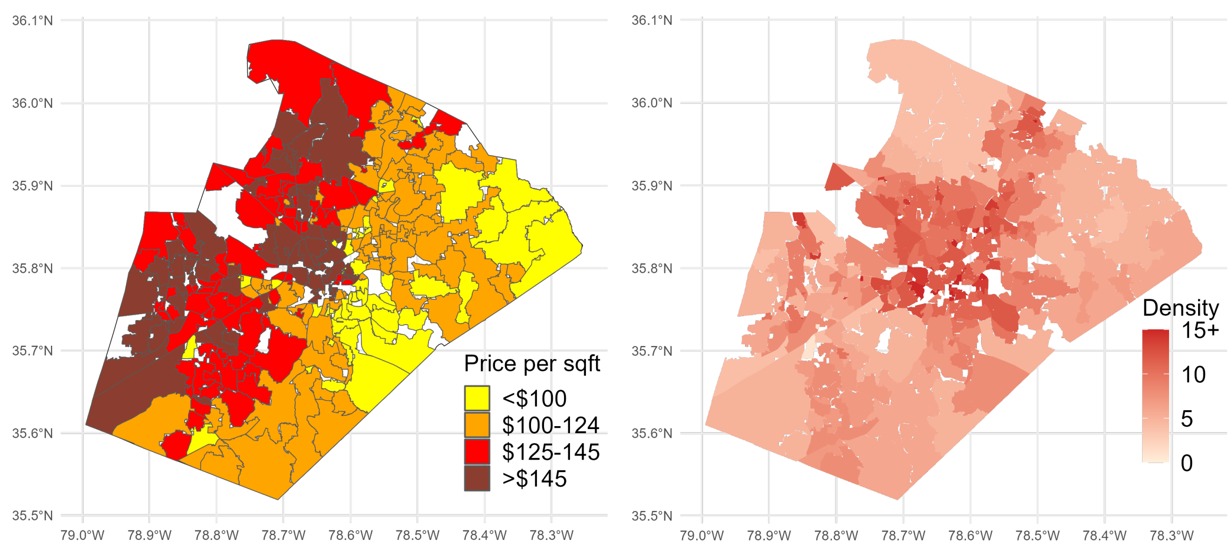
The map shows the boundaries of the constructed neighborhoods (blue), as well as the boundaries of the 2010 Census tracts (red).

Distance between schools and neighborhoods τ_{ns}. As the distance between neighborhood n and school s, we use the road distance between the centroid of n and s. The road distance between any two points is computed using the OSRM package, which is an interface between R and the OSRM API. OSRM is a routing service based on OpenStreetMap data.[30]

Zoning data and minimum house size. Multiple entities are in charge of zoning regulations in the county. While part of county land is regulated by the county itself, the zoning in other areas is done by a number of different local municipalities and/or unincorporated areas—namely: Raleigh, Apex, Cary, Fuquay-Varina, Garner, Holly Springs, Knightdale, Morrisville, Rolesville, Wake Forest, Wendell, and Zebulon. Geographic data on the zoning regulations for each entity is publicly available at: https://data-wake.opendata.arcgis.com/ (accessed August 2021). Each entity uses its own zoning categories and labels. By harmonizing regulation categories and labels across entities, we create a geographical dataset that gives, for any (residential) point in the county, the associated minimum lot size (MLS) regulation. Figure A-5(b) represents MLS regulations over (res-

[30] https://www.openstreetmap.org/, accessed August 2021.

Figure A-5: House Prices (Left) and Minimum Lot Size regulations (Right)



(a) House prices

(b) Zoning

The map on the left shows average house prices per square foot by neighborhoods (expressed in 2017 dollars, source: Wake County Real Estate Transaction data 2013–2017). The dollar thresholds used to determine the colors of the neighborhoods represent the quartiles of the distribution of within-neighborhood average price per square foot in sample. The map on the right shows density regulations (in dwelling units, du, per acre) throughout Wake County. Areas left uncolored are out of our sample and essentially non-residential.

idential land in) Wake County. Density regulations are typically expressed in dwelling units (du) per acre —the stronger the regulation, the lower the density allowed. Lighter areas in Figure A-5(b) are zoned for lower density, meaning that fewer dwelling units are allowed to be built on one acre of land. The inverse of density gives the more intuitive measure for MLS, which is expressed in acre per lot. There is a relatively wide range of MLS regulations throughout Wake County—from more than 25 du/acre in the urban center of the county, to less that 1 du/acre in the western periphery.

To each neighborhood, we attach a MLS. For neighborhoods that overlap multiple zoning areas with distinct MLS restrictions, the neighborhood-level MLS restriction is constructed as the least constraining MLS in the neighborhood. Formally, $\underline{\text{mls}}_n = \min\{\underline{\text{mls}}(x,y) \mid (x,y) \in n\}$, where (x,y) simply denotes the coordinate of any point in Wake County zoned for residential use, and $\underline{\text{mls}}(x,y)$ is the MLS restriction in place at that point. In the model though, we assume households choose and are constrained in their choice of *house* size, rather than *lot* size. We map neighborhood restrictions on minimum lot size

($\underline{\text{mls}}_n$) into a minimum house size available (h_n^{reg}). Regressing observed house sizes (in square feet) on our measure of minimum lot size ($\underline{\text{mls}}_n$, in acres) yields the mapping: $\underline{h}_n = 641 + 892 \times \underline{\text{mls}}_n$. From this mapping, we compute $h_n^{\text{reg}} = E[\underline{h}_n | \underline{\text{mls}}_n]$ for each neighborhood n, as well as the essential minimum housing $h_0 = 641$ (minimum house size in the absence of regulation, $\underline{\text{mls}}_n = 0$).

Real estate data and house prices. Publicly available records from Wake County show details about all real estate transactions in Wake County starting from 1956.[31] For each property sold, these data show the sale price and date, exact address of the property, characteristics of the lot and of the buildings/units, if any. In particular, we use the following characteristics in the analysis: sale date, sale price, acreage of the lot, year the building was built, whether the building is for residential use, and its type (single-family house, apartment, etc.), and heated area. We use heated area as our measure of house size.

To construct a measure of average house price by neighborhood and year, we proceed in three steps. First, we convert all prices into 2017 dollars to be consistent with household income provided in 2017 dollars in the ACS. Second, for each neighborhood and year, we compute an average sale price per square foot. We trim outlier cases (heated area smaller than 550 square feet or larger than 8,000 square feet with a sale price above $10 millions). For the structural estimation, prices are imputed for (neighborhood, year)-pairs with no observed transaction. If prices are observed for at least two years (out of four) for the neighborhood, prices for the missing year(s) are imputed using a linear neighborhood-specific trend. If prices are observed for fewer than two years for the neighborhood, prices for the missing years are imputed using average log-skills of children, share of ED families, and year. Average house prices (in dollar per square foot) are depicted in Figure A-5(a).

Household income w and average neighborhood income $\bar{w}_n$. We use the following (tract- and county-level) variables from the ACS five-year estimates (2013–17): "Family Type by Presence of Own Children Under 18 Years by Family Income in the Past 12 Months (in 2017 Inflation-Adjusted Dollars)" (NHGIS Code AIJA) and "Own Children Under 18 Years by Family Type and Age" (NHGIS Code AHZU). Data were downloaded from https://www.nhgis.org/ (accessed August 2021). The ACS five-year estimates (2013–17) table "Family Type by Presence of Own Children Under 18 Years by Family Income in the Past 12 Months (in 2017 Inflation-Adjusted Dollars)" (NHGIS

[31] https://www.wakegov.com/departments-government/tax-administration/real-estate, accessed August 2021

Code AIJA) gives household counts by census tracts for families with and without children and for 16 brackets of household income. We use variables AIJAE004–AIJAE019, AIJAE040–AIJAE055, and AIJAE075–AIJAE090 to characterize the income distribution of our "families," and AIJAE021–AIJAE036, AIJAE057–AIJAE072, and AIJAE092–AIJAE107 for other households. To construct our discrete household income types, we aggregate the 16 ACS brackets into ten: family income in the past 12 months below \$15,000; within \$15–25,000; \$25–35,000; \$35–50,000; \$50–75,000; \$75–100,000; \$100–125,000; \$125–150,000; \$150–200,000; and above \$200,000. Net household income w for households of each type ($w \in \{w_1, \ldots, w_{10}\}$) is constructed in three steps. First, gross income is assumed to be the middle point of the bracket (and \$250,000 for the top bracket "above \$200,000"). Next, net income is obtained from gross income using the NBER TAXSIM program,[32] assuming the following household characteristics: married couple, spending 28 percent of their income on a mortgage, and with one dependent younger than 13. We use these household characteristics for all households in the model, that is families and other households. To aggregate the ten discrete household income levels into the ED and non-ED categories available at the student level in the NCERDC data, we assign the lower four brackets (that is, with family income in the past 12 months below \$50,000 in 2017 inflation-adjusted dollars) to ED, and the six higher brackets to non-ED. ED status in the NCERDC is determined by eligibility for free or reduced-price lunch. Income levels for eligibility to the programs are determined annually by the USDA.[33] For reference, in 2017, the eligibility thresholds for reduced-price lunch (below 185 percent of the federal poverty line) were \$45,510 annual income for a household of four.[34]

Average income in neighborhood n is obtained as: $\bar{w}_n = \sum_{w \in \{w_1,\ldots,w_{10}\}} w \times Pr(w \mid n)$, where w is the level of household income for each discrete type (constructed as indicated above) and $Pr(w \mid n)$ is the share of households with income type w in neighborhood n. The distribution of average neighborhood income in the county is shown in Figure 1(b) in the main text.

Joint distribution of parental income and child skills, and families' sorting across neighborhoods. The joint distribution of family income and child skills is not directly observed. On the one hand, the NCERDC data, which contain individual information about children skills, only report ED and non-ED as measures of socioeconomic status.

[32] https://users.nber.org/~taxsim/, accessed August 2021.

[33] https://www.fns.usda.gov/cn/income-eligibility-guidelines, accessed August 2021.

[34] https://www.federalregister.gov/documents/2017/04/10/2017-07043/child-nutrition-programs-income-eligibility-guidelines, accessed August 2021.

On the other hand, the ACS, which shows household counts by income bracket, does not contain any information about children skills. We construct the joint distribution $\phi(w, a)$ as follows. We assume that conditional on skill type a, income is distributed as a log-normal, and we set the mean and variance of this distribution to match:

1. the share of ED students conditional on skill type a observed in the NCERDC data
2. average income conditional on skill type a, which we derive as follows:
   (a) Assuming that the four lower income types (obtained by aggregating the sixteen ACS brackets into ten) map to ED, the distribution of family income types in each neighborhood n conditional on ED status is given by the ACS: $Pr(w \mid n, \text{ED}) = \frac{Pr(w|n)}{\sum_{w' \in \{w_1,...,w_4\}} Pr(w'|n)}$ if $w \in \{w_1, \ldots, w_4\}$ and $Pr(w \mid n, \text{non-ED}) = \frac{Pr(w|n)}{\sum_{w' \in \{w_5,...,w_{10}\}} Pr(w'|n)}$ if $w \in \{w_5, \ldots, w_{10}\}$, while $Pr(w \mid n, \text{ED}) = 0$ if $w \in \{w_5, \ldots, w_{10}\}$ and $Pr(w \mid n, \text{non-ED}) = 0$ if $w \in \{w_1, \ldots, w_4\}$.

   (b) Combining with the NCERDC data, total counts of families of each income and skill types are given by: $N_{a,w} = \sum_n N_{a,\text{ED},n} \times Pr(w \mid n, \text{ED})$ if $w \in \{w_1, \ldots, w_4\}$, and $N_{aw} = \sum_n N_{a,\text{non-ED},n} \times Pr(w \mid n, \text{non-ED})$ if $w \in \{w_5, \ldots, w_{10}\}$, where $N_{a,\text{ED},n}$ (resp. $N_{a,\text{non-ED},n}$) is the count of ED (non-ED) students with skill type a living in neighborhood n observed in the NCERDC.

   (c) Average income conditional on skill type a is then given by: $\frac{\sum_{w \in \{w_1,...,w_{10}\}} N_{a,w} \times w}{\sum_{w' \in \{w_1,...,w_{10}\}} N_{a,w'}}$

The distribution of family types across neighborhoods is constructed as $\pi_{n|a,w} = N_{a,\text{non-ED},n} \times Pr(w \mid n, \text{non-ED})$ if $w \in \{w_1, \ldots, w_4\}$, and $N_{n|a,w} = N_{a,\text{non-ED},n} \times Pr(w \mid n, \text{non-ED})$ if $w \in \{w_5, \ldots, w_{10}\}$.

A.2.4 Descriptive statistics

Table A-3 shows descriptive statistics on our final samples of neighborhoods, students, and public schools.

Table A-3: Descriptive Statistics

	Mean	Stdev	Min	Max
PANEL A: NEIGHBORHOOD SAMPLE				
# of transactions obs. per year	51.67	56.45	0	322
Avg sale price by sqft	127.43	35.45	13.05	323.2
Avg MLS regulation (acre)	0.15	0.24	0	0.92
Avg minimum house size (in sqft)	885.81	252.1733	641.74	2,236.23
Avg # of school options (excl. base)	17.75	0.63	14	19
Avg # of school options w/ transp. (excl. base)	4.06	0.70	1	6
Distance to base sch. (miles)	3.71	3.24	.08	16.82
Avg. distance to option sch.	11.00	6.18	0.44	34.59
Has base change during period	0.23	0.45	0	1
Has catchment area change during period	0.63	0.48	0	1
Has change in option set during period	1	0	1	1
Avg # of student obs. per year	27.59	28.38	3	197
Share of econ. disadv. (ED) students	0.42	0.31	0	1
# of neighborhoods in sample	*312*			
# of neighborhood-year obs.	*1,248*			
PANEL B: STUDENT SAMPLE				
Is economically disadvantaged (ED)	0.35	0.48	0	1
Attends base, cond. on being ED	0.76	0.43	0	1
Attends base, cond. on being non-ED	0.73	0.44	0	1
Attends option w/ transp., cond. on ED	0.16	0.37	0	1
Attends option w/ transp., cond. on non-ED	0.21	0.41	0	1
Attends option w/o transp., cond. on ED	0.08	0.27	0	1
Attends option w/o transp., cond. on non-ED	0.06	0.24	0	1
Log-skills (standardized w/in cohort) cond. on ED	−0.56	0.83	−1.44	3.45
Log-skills (standardized w/in cohort) cond. on non-ED	0.30	0.95	−1.44	4.23
# of student-yr obs.	*34,428*			
PANEL C: PUBLIC SCHOOL SAMPLE				
Avg peer quality	1.34	0.38	0.51	2.56
Share econ. disadv. (ED) students	0.39	0.22	0.00	0.90
# of student obs. in sample (per year)	80.63	29.91	4	195
Is option school for some address	0.85	0.36	0	1
Has catchm. area change during period (base)	0.55	0.50	0	1
Has elig./transp. change during period (options only)	0.91	0.28	0	1
# of schools in sample	*111*			
# of school-year obs.	*428*			

The table shows descriptive statistics on our final samples of neighborhoods, students, and public schools. In Panel A (respectively B and C), the mean, standard deviation, minimum, and maximum are taken over the sample of neighborhood-year (respectively student-year, school-year) observations.

B Evidence from policy variation

B.1 House-price capitalization of school quality

Table B-1 shows parameter estimates for equation (4.1).

Table B-1: Changes in School Quality and Effects on House Prices

	House Price Psf (Log)			
Changes In School Quality (Log)	0.090*	0.092**	0.094**	0.096**
	(0.047)	(0.046)	(0.046)	(0.045)
Observations	812	812	812	812
Sale-Year F.E.	No	Yes	Yes	Yes
Control: Heated Area	No	No	Yes	Yes
Controls: Year House Built and Deeded Acreage	No	No	No	Yes

The table shows the effect of school quality on house prices. School quality is measured as the average test score of children attending the base school associated with the neighborhood in which the house is located. Details on the construction of the variable for changes in school quality are provided in Section 3.1 and Appendix A.2. Standard errors are clustered at the neighborhood level and reported in parentheses. *, **, and *** indicate statistical significance at the 10, 5, and 1-percent levels, respectively.

B.2 Commuting cost and the role of transportation

The within-neighborhood longitudinal variation in the provision of school transportation can be used to assess the extent to which commuting distance and transportation matter for families' choice of school. We estimate the following regression:

$$\pi_{s,n,t} = \beta_1 \mathbb{1}\left\{\text{Bus}\right\}_{s,n,t} + \beta_2 \tau_{ns} + \beta_3 \mathbb{1}\left\{\text{Bus}\right\}_{s,n,t} \times \tau_{ns} + \delta_{s,t} + \delta_n + \epsilon_{s,n,t} \,, \tag{B-1}$$

where $\pi_{s,n,t}$ is the share of children from cohort t in neighborhood n attending option school s; $\mathbb{1}\left\{\text{Bus}\right\}_{s,n,t} \in \{0,1\}$ indicates whether transportation is provided to school s from neighborhood n in year t; τ_{ns} denotes the distance between the school and the neighborhood. The inclusion of the interaction term between distance and the availability of school transportation allows the benefit of transportation to vary as the distance between the school and the neighborhood increases. This spatial heterogeneity is in line with our structural model. We control for school-specific aggregate trends in enrollment ($\delta_{s,n}$) and include neighborhood fixed effects (δ_n) to exploit the within-neighborhood (across cohorts of kindergarten students) institutional variation in transportation provision.[35]

Table B-2 shows that the provision of school transportation increases school enrollment by 2.4 percentage points in the hypothetical scenario of a zero-mile home-school distance. As the distance to school increases, the impact of providing transportation decreases (-3.3 percentage points per 10 miles of distance), and it becomes null at a distance of approximately seven miles. The distance itself reduces the attractiveness of that school, with an extra mile reducing the share of children going to that school from a given neighborhood by 2.8 percentage points.

Table B-2: Neighborhood School Transportation and School Enrollment

	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{Bus}\}_{s,n,t}$	0.008***	0.007***	0.024***	0.023***
	(0.002)	(0.002)	(0.005)	(0.005)
$\mathbb{1}\{\text{Bus}\}_{s,n,t} \times \tau_{ns}$			-0.033***	-0.033***
			(0.005)	(0.005)
τ_{ns}			-0.029***	-0.029***
			(0.003)	(0.003)
Observations	22,153	22,153	22,153	22,153
Neighborhood F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	No	Yes	No
School F.E.	Yes	No	Yes	No
School F.E. $\times$ Year F.E.	No	Yes	No	Yes

The table shows estimates of the effect of providing school transportation to/from an option school on enrollment. Standard errors are clustered at the neighborhood level and reported in parentheses. *, **, and *** indicate statistical significance at the 10, 5, and 1-percent levels, respectively.

Table B-3 shows parameters values obtained when estimating Equation (B-1) on data simulated from the estimated structural model.

[35] The specific structure of the network of schools and neighborhoods in Wake County, where one school serves multiple neighborhoods, allows us to distinguish between aggregate school trends in enrollment and the neighborhood-specific effects of transportation.

Table B-3: Neighborhood School Transportation and School Enrollment (Model)

	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{Bus}\}_{s,n,t}$	0.014	0.015	0.031	0.031
$\mathbb{1}\{\text{Bus}\}_{s,n,t} \times \tau_{ns}$			-0.030	-0.029
τ_{ns}			-0.019	-0.020
Neighborhood F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	No	Yes	No
School F.E.	Yes	No	Yes	No
School F.E. $\times$ Year F.E.	No	Yes	No	Yes

The table shows the model-predicted effect of providing school transportation to/from an option school on enrollment.

C Estimation appendix

Skill type is discretized into ten bins (a) corresponding to the deciles of the baseline skill distribution, and family income type is discretized into the ten income bins (w) from the ACS. Let T denote the number of years used in estimation ($T = 4$), N the number of neighborhoods ($N = 312$), A the number of children's skills bins ($A = 10$). We estimate model parameters using $A \times 2 + 1 + 4 + 1 + 16 + 1 + (N-1) \times 2 + (N-1) \times 2 + 1 + N + T = 1,604$ moments, which we define formally here. Below, we use $a(i)$, $s(i)$, and $n(i)$ to denote child i's skill level, the school they attend, and the neighborhood they live in, respectively. We use $t(i)$ to denote the cohort i is from (that is, the school year child i attends kindergarten). $\bar{a}_{s(i)}$ denotes the average peer quality in the school attended by child i for the year i attends the school (the subscript t is dropped to simplify notation). We write $\#\mathcal{A}$ to denote the number of elements in the set $\mathcal{A}$.

C.1 Data moments

1. Average peer quality in the school attended by a child with skill type a and ED status eds, for all $a \in \{a_1, \ldots, a_{10}\}$ and $eds = 0, 1$ [$A \times 2$ moments]

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\#\mathcal{A}_{a,t,eds}}\sum_{i\in\mathcal{A}_{a,t,eds}}\bar{a}_{s(i)} \quad \text{where } \mathcal{A}_{a,t,0} = \{i \mid a(i) = a, i \text{ is non-ED, and } t(i) = t\}$$
$$\text{and } \mathcal{A}_{a,t,1} = \{i \mid a(i) = a, i \text{ is ED, and } t(i) = t\}$$

2. Regression coefficient (β_1) of changes in house prices on changes in associated school quality; see Equation (4.1) and Table B-1.

3. Average (over years and neighborhoods) distance to school attended conditional on transportation being provided, separately for ED and non-ED students [2 moments]

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\#\mathcal{A}_{t,eds}}\sum_{i\in\mathcal{A}_{t,eds}}\tau_{n(i)s(i)} \quad \text{where } \mathcal{A}_{t,0} = \{i \mid i \text{ is non-ED}, s(i) \in \mathcal{T}_{n(i)t}, \text{ and } t(i) = t\}$$
$$\text{and } \mathcal{A}_{t,1} = \{i \mid i \text{ is ED}, s(i) \in \mathcal{T}_{n(i)t}, \text{ and } t(i) = t\}$$

4. Average (over years and neighborhoods) distance to school attended conditional on transportation not being provided, separately for ED and non-ED students [2 moments]

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\#\mathcal{A}_{t,eds}}\sum_{i\in\mathcal{A}_{t,eds}}\tau_{n(i)s(i)} \quad \text{where } \mathcal{A}_{t,0} = \{i \mid i \text{ is non-ED}, s(i) \in \mathcal{NT}_{n(i)t}, \text{ and } t(i) = t\}$$
$$\text{and } \mathcal{A}_{t,1} = \{i \mid i \text{ is ED}, s(i) \in \mathcal{NT}_{n(i)t}, \text{ and } t(i) = t\}$$

5. Average variance of neighborhood-level attendance share across schools [1 moment]

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\sum_n \#\mathcal{A}_{n,t}}\sum_{n} var_{n,t} \times \#\mathcal{A}_{n,t}$$

where

$$\begin{aligned}
\mathcal{A}_{n,t} &= \{i \mid n(i) = n \text{ and } t(i) = t\} \\
\mathcal{A}_{s,n,t} &= \{i \mid s(i) = s, n(i) = n \text{ and } t(i) = t\} \\
\Pr(s \mid n,t) &= \frac{\#\mathcal{A}_{s,n,t}}{\#\mathcal{A}_{n,t}} \\
\text{and } var_{n,t} &= \sum_{s \in \mathcal{L}_{nt} \setminus \mathcal{P}_n} \left[\left(\Pr(s \mid n,t) - \sum_{s \in \mathcal{L}_{nt} \setminus \mathcal{P}_n} \Pr(s \mid n,t) \right)^2 \right]
\end{aligned}$$

6. Private-school attendance shares at the PUMA level by ED status [8×2 moments] —obtained directly from the ACS (five-year estimates 2013-17).

7. Share attending private school among applicants to option school who lose the admission lottery —taken from Dur et al. (2022)

8. Empirical shares of families across neighborhoods [$(N-1)$ moments]

$$\frac{1}{T} \sum_{t=1}^{T} \frac{\#\{i \mid n(i) \text{ and } t(i) = t\}}{\#\{i \mid t(i) = t\}}, \text{ from the NCERDC}$$

9. Empirical shares of other households across neighborhoods [$(N-1)$ moments] —obtained by mapping the census tract-level households counts from the ACS to our neighborhoods.

10. Average income for families and the rest of/other households in each neighborhood [$N \times 2$ moments] —constructed as explained in A.2.3

11. Take-up rate from the MTO experiment reported in Galiani et al. (2015)

12. Average (over time) house prices in each neighborhood and average (over neighborhood) house prices in each year [$N + T$ moments]

$$\frac{1}{T} \sum_t \text{price}_{nt} \;\text{ for each } n \quad \text{ and } \quad \frac{1}{N} \sum_n \text{price}_{nt} \;\text{ for each } t$$

C.2 Model moments

Model-generated moments can be written as a function of the model parameters. Recall from Section 2.3 that:

$$\pi_{n|a,w} = \frac{\exp(v_{a,w,n})}{\sum_{\tilde{n}} \exp(v_{a,w,\tilde{n}})} \quad \text{with} \quad v_{a,w,n} = u_{w,n} + \alpha_{0n} + \alpha_{1n} \log(w) + \bar{v}_{a,w}(\mathcal{L}_n),$$

where $\bar{v}_{a,w}(\mathcal{L}_n) = \mathbb{E}_{\{\varepsilon_s\}} \left[\max_{s \in \mathcal{L}_n} \{ p_s v_{a,w,s|n} + (1-p_s) v^{\mathcal{BP}}_{a,w|n} \} \right]$, with $v^{\mathcal{BP}}_{a,w|n} = \max\{v_{a,w,\mathcal{B}_n|n}, v_{a,w,\mathcal{P}_n|n}\}$. The school year subscript t is dropped to simplify exposition. The probability of applying to school s conditional on neighborhood n and child skills k,

$$\pi_{s|n,a,w} = Pr \left[\hat{v}_{a,w,s|n} \geq \hat{v}_{a,w,\tilde{s}|n} \; \forall \tilde{s} \in \mathcal{L}_n \right],$$

does not have a closed-form solution and is estimated by simulation.
The probability that a family of type (a, w) chooses neighborhood n and applies to school $s \in \mathcal{L}_n$ is: $\pi_{n,s|a,w} = \pi_{s|n,a,w} \times \pi_{n|a,w}$ If p_s is the admission probability to school s conditional on applying, then the probability that a family of type (a, w) chooses neighborhood n and attends school $s \in \mathcal{L}_n$ is: $\pi^{\text{att}}_{n,s|a,w} = \pi_{n,s|a,w} \times p_s$. Similarly, $\pi^{\text{att}}_{s|n,a,w} = \pi_{s|n,a,w} \times p_s$. Then:

1. Average peer quality in the (public) school attended by a child with skills type a and ED status eds, for all $a \in \{a_1, \ldots, a_{10}\}$ and $eds = 0, 1$ [$K \times 2$ moments]

$$\frac{1}{T} \sum_t \left\{ \sum_w \left(\sum_n \sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_n} \pi^{\text{att}}_{n,s|a,w} \times \bar{a}_s \right) \times \phi(w \mid a, eds) \right\}$$

2. Regression coefficient (β_1) obtained when estimating Equation (4.1) on model-generated data

3. Average (over years and neighborhoods) distance to optiom school attended conditional on transportation being provided, separately for ED and non-ED students

$$\frac{1}{T} \sum_t \left\{ \sum_a \sum_w \left(\sum_n \sum_{s \in \mathcal{T}_{n,t}} \pi^{\text{att}}_{n,s|a,w} \times \tau_{n,s} \right) \times \phi(a, w \mid eds) \right\}, \text{ for } eds = 0, 1$$

4. Average (over years and neighborhoods) distance to option school attended condi-

tional on transportation not being provided, separately for ED and non-ED students

$$\frac{1}{T}\sum_{t}\left\{\sum_{a}\sum_{w}\left(\sum_{n}\sum_{s\in\mathcal{NT}_{n,t}}\pi^{\text{att}}_{n,s|a,w}\times\tau_{n,s}\right)\times\phi(a,w\mid eds)\right\},\ \text{for } eds=0,1$$

5. Average standard deviation of neighborhood-level attendance share across public schools [1 moment]

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{n}\left(\sum_{s\in\mathcal{L}_{nt}\setminus\mathcal{P}_n}\left[\left(\Pi_{s|n,t}-\sum_{s\in\mathcal{L}_{nt}\setminus\mathcal{P}_n}\Pi_{s|n,t}\right)^2\right]\right)\times\Pi_{n,t}$$

where, for each year t, $\Pi_{n,t}=\sum_{a,w}\pi_{n|a,w}\times\phi(a,w)$ and $\Pi_{s|n,t}=\sum_{a,w}\pi^{\text{att}}_{s|n,a,w}\times\phi(a,w)$.

6. Private-school attendance shares by ED status at the PUMA level [8×2 moments]

7. Share attending private school among applicants to option schools who lose the admission lottery in 2015 [1 moment]

$$\frac{\sum_{s,n,a,w}(1-p_s)\pi_{s,n|a,w}\phi(a,w)\times\mathbb{I}\{s\in\mathcal{T}_n\cup\mathcal{NT}_n\}\times\pi^{fallback}_{\mathcal{P}_n,n|a,w}}{\sum_{s,n,a,w}(1-p_s)\pi_{s,n|a,w}\phi(a,w)\times\mathbb{I}\{s\in\mathcal{T}_n\cup\mathcal{NT}_n\}}$$

8. Distribution of families across neighborhoods [$(N-1)$ moments]

$$\frac{1}{T}\sum_{t}\sum_{w}\sum_{s}\pi_{n|a,w}\times\phi(a,w)\text{ for each } n$$

9. Distribution of other households across neighborhoods [$(N-1)$ moments]

$$\frac{1}{T}\sum_{t}\sum_{w}\pi^{o}_{n|w}\times\phi^{o}(w),\quad\text{for each } n$$

10. Average neighborhood income for families and the rest of/other households in each neighborhood [$N\times 2$ moments]

$$\frac{1}{T}\sum_{t}\frac{\sum_w w\sum_a\pi_{n|a,w}\phi(a,w)}{\sum_w\sum_a\pi_{n|a,w}\phi(a,w)}\text{ and }\frac{1}{T}\sum_{t}\frac{\sum_k w\pi^{o}_{n|w}\phi^{o}(w)}{\sum_w\pi^{o}_{n|w}\phi^{o}(w)}\text{ for each } n$$

11. Take-up rate for the following policy: families living in neighborhoods with more than 40 percent of ED households (at baseline) are offered a housing voucher equal to the 40th percentile of the rent distribution in Wake County, which they can use if they choose to live in a neighborhood where the share of ED household is below 40 percent (at baseline). The moment we use is the fraction of eligible families who take up the housing voucher [1 moment]

12. Average (over time) equilibrium house prices in each neighborhood and average (over neighborhood) equilibrium house prices in each year [$N + T$ moments]

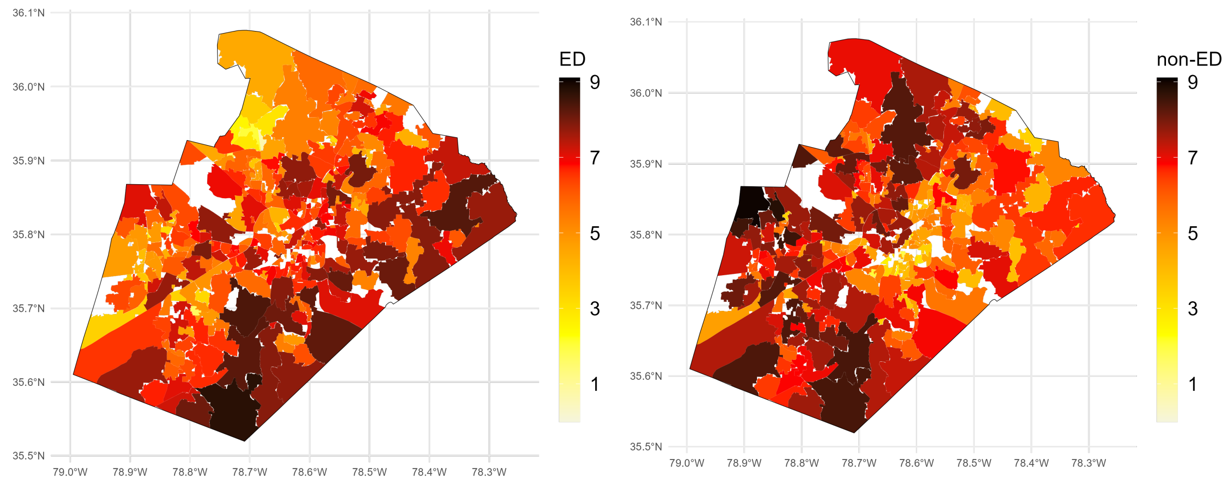
$$\frac{1}{T}\sum_t r_{nt} \text{ for each } n \quad \text{and} \quad \frac{1}{N}\sum_n r_{nt} \text{ for each } t$$

C.3 Additional estimates

Figure C-1 shows the value of neighborhood amenities for families with average income conditional on being ED (left panel) or non-ED (right panel) implied by our estimated for α_{0n} and α_{1n}. Consistently with Equation (2.4), the value of neighborhood n amenities for a family with income w is given by: $\alpha_{0n} + \alpha_{1n}\log(w)$. The income levels w used to produce Figure C-1 are \$24,210 (left panel) and \$104,395 (right panel), which correspond to the average income in Wake County conditional on the family being ED and non-ED, respectively.

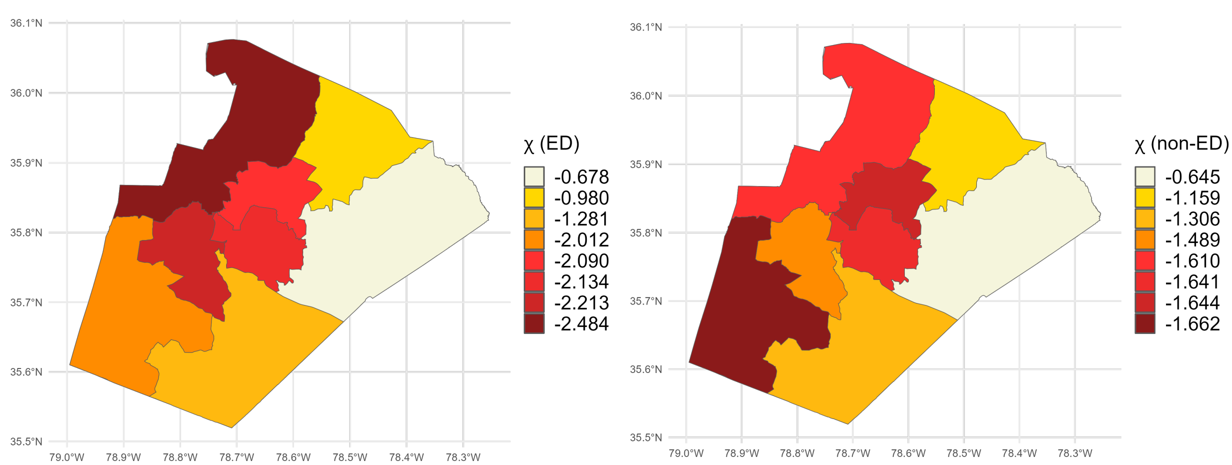
Figure C-2 shows estimates for the PUMA-level parameter χ for ED and non-ED families. The estimated PUMA-level private school costs χ generally align with average private-school tuition in each PUMA, which are shown as a reference in Figure C-3.

Figure C-1: Estimated value of neighborhood amenities for average-income ED and non-ED families



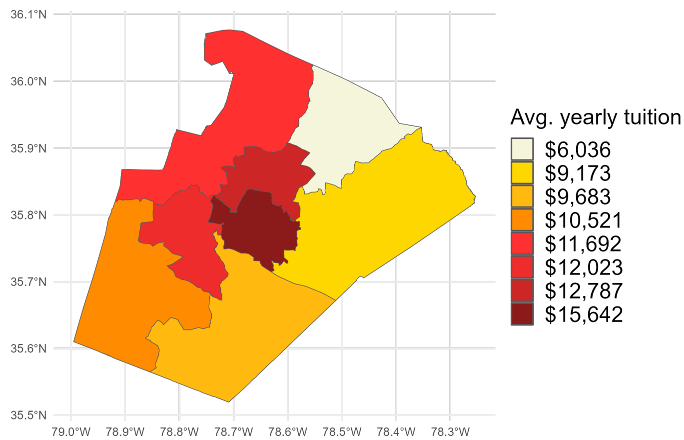
The map on the left (right) shows the neighborhood amenities valuation (in utils) for the average-income ED (non-ED) family implied by our estimates for $\alpha_{0,n}$ and $\alpha_{1,n}$. The value of neighborhood n amenities for a family with income w is given by: $\alpha_{0n} + \alpha_{1n}\log(w)$. The income levels w used to produce the figure are \$24,210 (left panel) and \$104,395 (right panel), which correspond to the average income in Wake County conditional on the family being ED and non-ED, respectively.

Figure C-2: Estimated values of χ for ED and non-ED families



The map on the left (right) shows estimated values for the PUMA-level private-school parameters $\chi_{w,n}$ for ED (non-ED) families.

Figure C-3: Average Private Schools Tuition by PUMA



The map shows average private-school tuition by PUMA in Wake County. Average tuition is constructed in line with the mapping of neighborhoods to private schools in the model. First we map each neighborhood to its closest private school, then we take a weighted average, over neighborhoods and within PUMAs, of the closest private school tuition using the share of families in each neighborhood as weights. Tuition levels were collected from each school's website for academic year 2022-23. If tuition information was unavailable for a neighborhood's closest private school, the second closest private school was used to produce the figure.

D Counterfactual appendix

D.1 Decomposition

For a given allocation of the economy, we construct the resulting welfare for agent i, V_i, as follows:

$$\begin{aligned}
\underbrace{V_i}_{\text{total}} = \underbrace{u_{w,n}}_{\text{consumption}} &+ \underbrace{\gamma_{a,w}[p_s \log(\bar{a}_s) + (1-p_s)\log(\bar{a}_{\mathcal{O}_n})]}_{\text{peers}} + \\
&\underbrace{p_s\left(-\kappa_{s,w}\tau_{s,n}\right) + (1-p_s)\left(-\kappa_{\mathcal{O}_n,w}\tau_{\mathcal{O}_n,n}\right)}_{\text{commute}} + \\
&\underbrace{p_s\left(\sigma_S \varepsilon_{si}\right)\mathbb{I}_{s\notin\mathcal{P}_n} + (1-p_s)\left(\sigma_S \varepsilon_{\mathcal{B}_n i}\right)\mathbb{I}_{\mathcal{O}_n=\mathcal{B}_n}}_{\text{unobs. pref. (public)}} + \\
&\underbrace{p_s\left(\chi_{w,n} + \sigma_P \varepsilon_{\mathcal{P}i}\right)\mathbb{I}_{s\in\mathcal{P}_n} + (1-p_s)\left(\chi_{w,n} + \sigma_P \varepsilon_{\mathcal{P}i}\right)\mathbb{I}_{\mathcal{O}_n=\mathcal{P}_n}}_{\text{unobs. pref. (private)}} + \\
&\underbrace{\alpha_{0n} + \alpha_{1n}\log(w) + \sigma_N \varepsilon_{ni}}_{\text{exog. amenities}}
\end{aligned} \tag{D-1}$$

where (a, w) denotes agents' i exogenous type, and n and s denote the neighborhood they live in and the school they apply to, respectively. Recall that children that are not admitted to their option school of choice enroll into their preferred fallback alternative—either their base and their local private school. Here, $\mathcal{O}_n$ denotes this preferred fallback school.

To compute the welfare gains of a given policy, we construct the value of agent i in the equilibrium allocation under that policy and then compute the difference relative to agent i's value in the baseline equilibrium: $\Delta V_i = V_i^{counterfactual} - V_i^{baseline}$. The change in each component of welfare is defined analogously.

D.2 Utilitarian Welfare and Alternative School Choice Policy Target

Here, we show how to compute the welfare gains of a counterfactual policy using a general welfare criterion. We follow Davila and Schaab (2023) and pose a social welfare function,

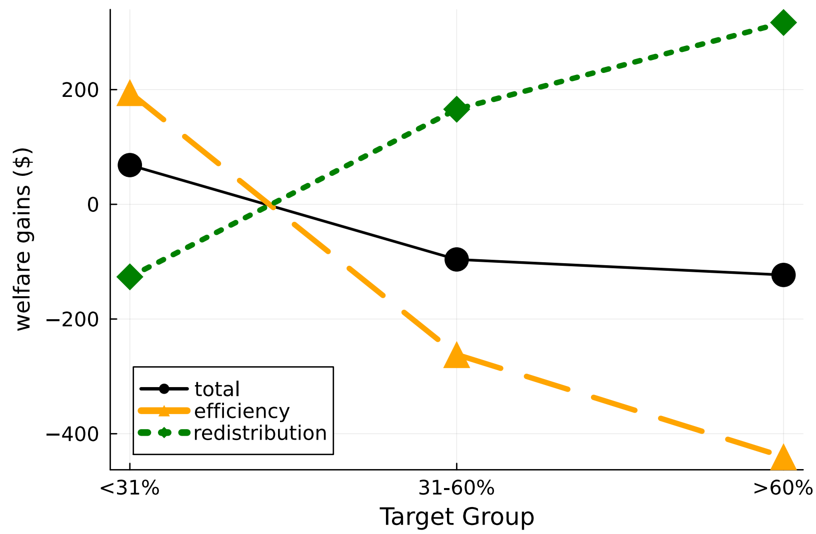
$$\widetilde{W} = \mathcal{W}(V_1, V_2, ..., V_I),$$

over the individual value V_i for agents $i = 1, 2, .., I$. Let λ_i be agent i's marginal utility of income, and $\phi_i \equiv \frac{\partial \mathcal{W}}{\partial V_i}$ be the marginal contribution of agent i's utility to social welfare. It is convenient to work with the normalized welfare $W = \widetilde{W} / \frac{\sum_i \phi_i \lambda_i}{I}$ so that policies that raise W by x units are welfare-equivalent to increasing all agents' income by x dollars. We also define the relative marginal social value of a change in the allocation of the numeraire to agent i, $\omega_i \equiv \frac{\phi_i \lambda_i}{\sum_i \phi_i \lambda_i / I}$. Note that $\sum_i \omega_i / I = 1$. The normalized welfare impact ΔW of a given policy can then be approximated by:

$$\Delta W \approx \sum_i \omega_i \frac{\Delta V_i}{\lambda_i} = \underbrace{\sum_i \frac{\Delta V_i}{\lambda_i}}_{\text{efficiency}} + \underbrace{\mathcal{C}ov\left(\omega_i, \frac{\Delta V_i}{\lambda_i}\right)}_{\text{redistribution}} \tag{D-2}$$

where ΔV_i is the change in value for agent i. Equation 5.1 shows how the aggregate welfare gain can be expressed as the sum of all agents' dollar-equivalent gains, $\frac{\Delta V_i}{\lambda_i}$, weighted by their relative marginal social value, ω_i. Simple algebra allows to further decompose the aggregate welfare gain into the sum of two terms. The first term quantifies the *efficiency* gains of the policy and it is the unweighted sum of all individual changes in dollar-equivalent welfare. This term is invariant to the choice of a specific social welfare function. The second term, quantifying the *redistribution* effect of the policy, captures

Figure D-1: Welfare Effects of School Choice Expansion as a Function of the Target Group



This figure shows the aggregate welfare gains of the school-choice expansion policy as well as the decomposition of welfare gains into efficiency and redistribution (based on Equation (5.1)) as we vary the target group.

the extent to which the policy disproportionately benefits agents who have higher contributions to aggregate welfare, either because their marginal utility of income λ_i is higher or because they have a higher weight in the social welfare function, ϕ_i.

In the analysis on Section 5, we adopt a money-metric welfare criterion which corresponds to the case in which $\phi_i = \frac{1}{\lambda_i}$. It is easy to see that, under these weights, $\omega_i = 1$ and the redistribution term in Equation D-2 is equal to zero. The total change in welfare then simplifies to the efficiency component, as in Equation 5.1. We adopt the efficiency component as our benchmark welfare criterion because it captures the relevant objective of a social planner who has access to other, more direct, tools for redistribution (e.g. a progressive fiscal policy). As a commonly used alternative, we also compute the gains implied by a utilitarian social welfare function, that is one in which all individuals are given the same weight in the social welfare function ($\phi_i = 1$). The redistribution term will then be positive if the policy creates higher dollar-equivalent benefits for agents with a higher marginal utility of income.

Figure D-1 shows utilitarian welfare (black solid line), the money-metric component (orange long-dashed line) and the redistribution component (green short-dashed line) for three school choice expansion policies that differ only with respect to the target set of neighborhoods that gain access to the receiving schools. The rightmost point is our benchmark policy in Section 5.1, that offers additional school choice to neighborhoods with a share of ED families above 60%. As reported in the main text, the money-metric component is negative, corresponding to a loss of more than $400 per family. Since this policy benefits lower-income families, the redistributive term is positive, but not enough to generate an overall gain even under a utilitarian welfare function. The two other policies we consider provide increasing efficiency gains as we target higher-income neighborhoods– which are also located closer to receiving schools. For those policies, efficiency and redistribution roughly cancel each other out, delivering higher, although overall small, utilitarian welfare gains.

D.3 Additional Details on the Cost of the School-Choice Policy

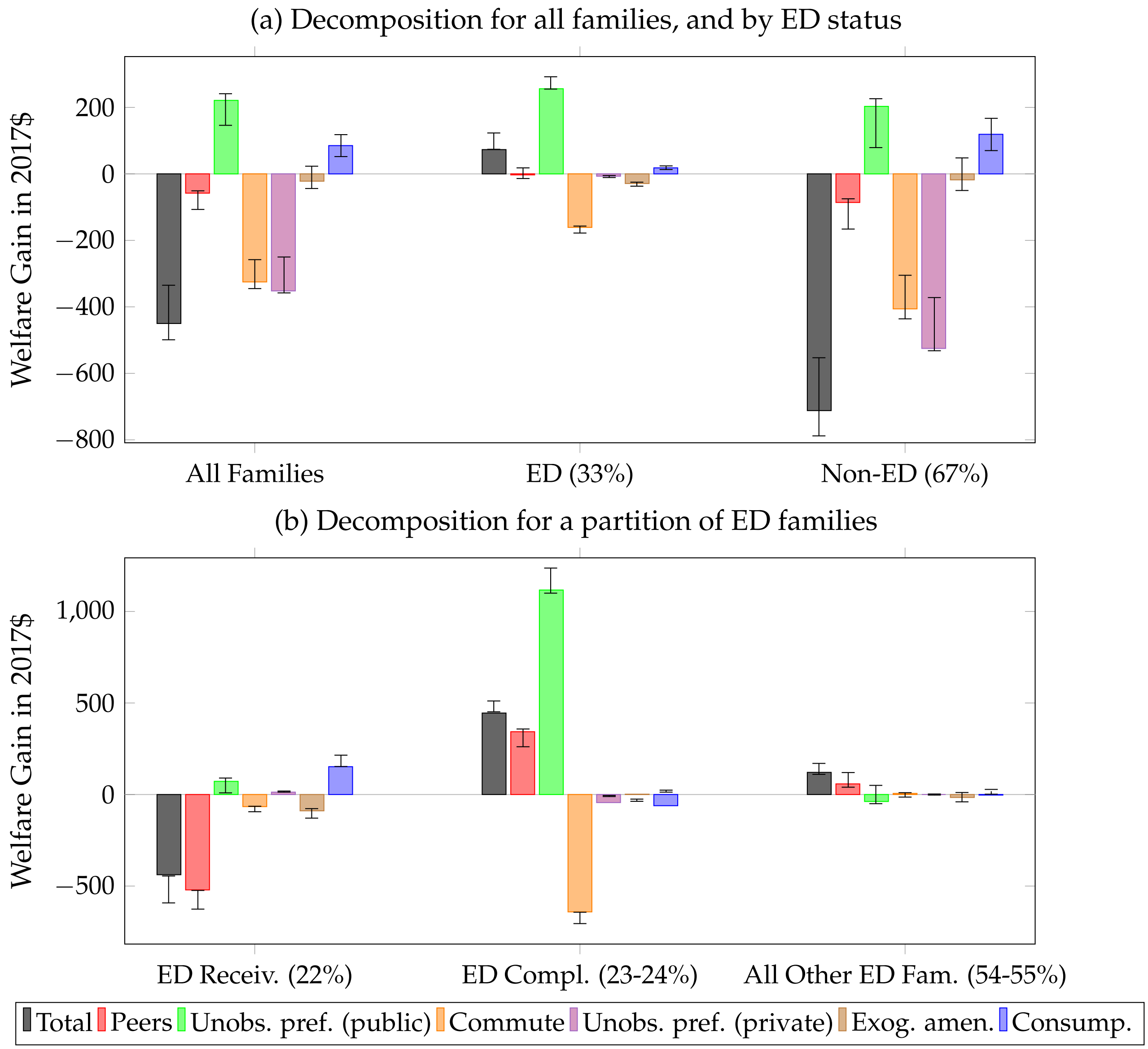
To derive the per-mile cost of school transportation, we use data from the North Carolina Department of Public Instruction which report that total expenditures for school transportation in the WCPSS amounted to $57,658,097.15 in the 2014-15 school year for a total mileage of 17,091,229.[36] These numbers imply a cost of $3.37 per mile.

For this per-mile cost, we derive the total cost of our counterfactual school-choice expansion policy. At baseline, children in the sample who attend a school providing transportation commute on average 2.56 miles to their attended school, and the policy increases the average commuting distance by 21.8 percent. These extra miles are traveled twice a day, five days a week, 36 weeks per year; they amount to an additional school transportation cost of $3.37 \times 0.20 \times 2.56 \times 2 \times 5 \times 36 =$$678.30 per child.

D.4 Exploring Equilibrium Multiplicity

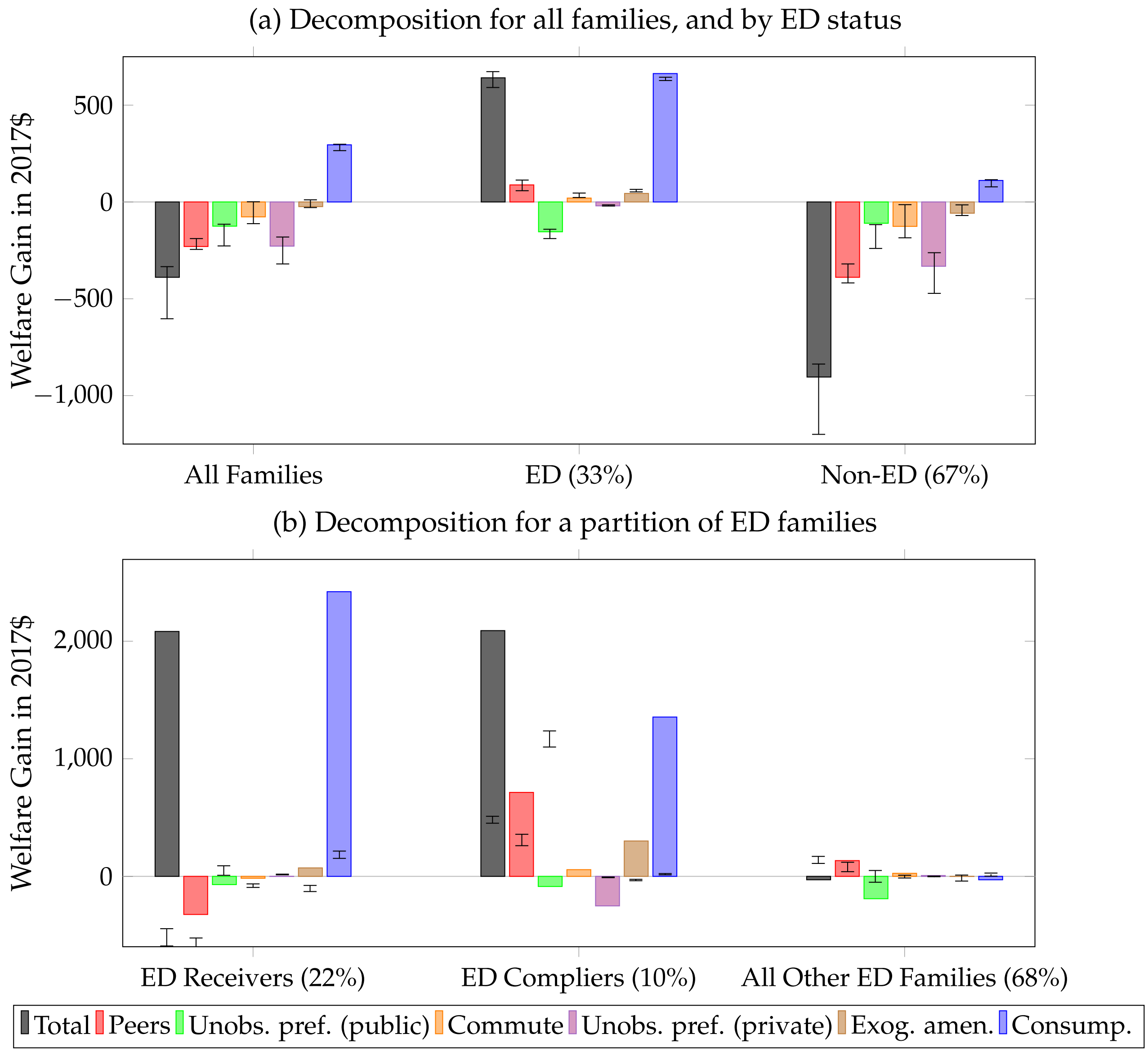
[36]Source: https://www.ncbussafety.org/documents/2014-15_nc_pupiltransportationdata.pdf, p.11 (expenditures) and p.3 (total miles).

Figure D-2: Decomposition of Welfare Effects of the School Choice Policy Across Equilibria



The figure is analogous to Figure 3 in the main text and, in addition, it shows the range of values taken by each factor in the welfare decomposition across equilibria. The top chart illustrates the welfare changes induced by the housing voucher policy for all families (left), ED families (center), and non-ED families (right). The bottom chart illustrates the welfare changes induced by the housing voucher for a partition of ED families —those who live in receiving neighborhoods at baseline (or *Receivers* for short, left), those who do not live in receiving neighborhoods at baseline but do under the policy (*Compliers*, center), and all other ED families (right). For each set of families, the first bar on the left (black) shows the total average change in welfare resulting from the policy. The second to seventh (colored) bars show the decomposition of this average welfare change into the changes induced by the policy in the following determinants of family utility (from left to right): school peers ($\gamma \bar{a}$), idiosyncratic preference received from attending a public school if doing so ($\sigma_S \varepsilon_s$), commuting costs ($\kappa\tau$), net value from attending a private school if doing so ($\chi_{w,n} + \sigma_P \varepsilon_{\mathcal{P}_i}$), exogenous neighborhood amenities ($\alpha_{0n} + \alpha_{1n} \log(w) + \sigma_N \varepsilon_{ni}$), and value from consumption ($u_{w,n}$). The formal decomposition of welfare into its components is shown in Equation D-1.

Figure D-3: Decomposition of Welfare Effects of the Voucher Policy Across Equilibria



The figure is analogous to Figure 3 in the main text and, in addition, it shows the range of values taken by each factor in the welfare decomposition across equilibria. The top chart illustrates the welfare changes induced by the housing voucher policy for all families (left), ED families (center), and non-ED families (right). The bottom chart illustrates the welfare changes induced by the housing voucher for a partition of ED families —those who live in receiving neighborhoods at baseline (or *Receivers* for short, left), those who do not live in receiving neighborhoods at baseline but do under the policy (*Compliers*, center), and all other ED families (right). For each set of families, the first bar on the left (black) shows the total average change in welfare resulting from the policy. The second to seventh (colored) bars show the decomposition of this average welfare change into the changes induced by the policy in the following determinants of family utility (from left to right): school peers ($\gamma\bar{a}$), idiosyncratic preference received from attending a public school if doing so ($\sigma_S\varepsilon_s$), commuting costs ($\kappa\tau$), net value from attending a private school if doing so ($\chi_{w,n} + \sigma_P\varepsilon_{\mathcal{P}_i}$), exogenous neighborhood amenities ($\alpha_{0n} + \alpha_{1n}\log(w) + \sigma_N\varepsilon_{ni}$), and value from consumption ($u_{w,n}$). The formal decomposition of welfare into its components is shown in Equation D-1.