On the Spatial Determinants of Educational Access*

Francesco Agostinelli† Margaux Luflade‡ Paolo Martellini§

November 2026

Abstract

This paper develops a spatial equilibrium model of residential sorting and school choice in which catchment areas, admission lotteries, and commuting costs shape each neighborhood's educational access. Preferences are identified from house-price responses to quasi-experimental boundary redrawings and cross-sectional school attendance by income, skill, and distance. Low-income families prioritize school proximity, while high-income families prioritize peer composition. A place-based school choice expansion and a people-based housing voucher granting access to high-quality schools deliver modest gains to intended beneficiaries, more than offset by losses for other families. Distance partially shields affluent neighborhoods from the first policy, and zoning from the second.

JEL Classification: I24, J13, R23, R31.

*We thank seminar participants at ASU, BFI China, Boston University, Cornell, Duke, Federal Reserve Bank of Minneapolis, NYU, Penn, Stanford, University of Arizona, UCLA, University of Minnesota, University of Oslo, Yale. We are also thankful to our audiences at workshops and conferences including the Barcelona Summer Forum, Chicago-Princeton Spatial Economics Conference, Junior Spatial Economics Conference at BFI, NBER SI 2023, Rohit and Harvanit Kumar Conference on Early Childhood Development at BFI, Saieh Family Fellows Workshop, and Urban Economics Association Conferences. 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. Julia Reynolds and 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, NYU-Stern School of Business. Email: paolo.martellini@nyu.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 has two related objectives. First, we measure preferences for school and neighborhood attributes in the presence of institutional constraints that govern residential and school choice. Second, we investigate how the geography of neighborhood sorting determines the aggregate and distributional implications of local policies that grant low-income families access to high-quality schools. We address these issues by building a spatial equilibrium model in which families' neighborhood and school choices are determined by preferences over school composition, commuting distance from home 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 Wake County, North Carolina, combining the effect of longitudinal changes in school catchment areas on house prices, and variations in enrollment by school distance and transportation provision. The estimates reveal substantial heterogeneity, by income and child skill, in how families trade off school quality against proximity. This heterogeneity shapes our policy analysis: a place-based school choice expansion and a people-based housing voucher both deliver modest gains to their intended beneficiaries and generate aggregate welfare losses. The losses are concentrated among families who, absent the policies, pay high house prices to reside near high-quality schools. Among them, inframarginal low-income families are overwhelmingly exposed to the inflow of policy beneficiaries into their schools. Geography plays a key role in both policies: high-quality schools located far from low-income neighborhoods are shielded from the school choice expansion by distance, while more stringently zoned neighborhoods are insulated from the relocation of voucher recipients.

Our first step is to develop a spatial equilibrium model of residential sorting and school choice that captures key institutional features common across U.S. school districts: residence-based school portfolios, school transportation provision, and zoning restrictions on housing demand. Conditional on location, families have guaranteed access to a neighborhood-specific public school ("base school") and may apply to a neighborhood-specific set of alternative public schools ("option schools"), with admission determined by lottery when schools are oversubscribed. Families may also opt out of the public sys-

tem and attend a private school. School choices depend on exogenous and endogenous school quality—measured by peer skill composition—and commuting costs, which vary with distance and transportation availability. Residential choice depends on the value of the associated school portfolio, house prices, neighborhood amenities, as well as zoning restrictions that impose minimum housing consumption.

Our empirical analysis focuses on Wake County, North Carolina, a natural setting for our study. The county is served by a single school district with uniform per-pupil expenditure, and it experienced several changes in base-school catchment boundaries. To map the model to the data, we construct a comprehensive dataset combining multiple sources. We use student-level administrative data from the North Carolina Education Research Data Center (NCERDC), which report each public-school student's school attendance, test scores, and residential location. We merge these records with (i) yearly information from the Wake County Public School System (WCPSS) on school catchment areas and school portfolios by address; (ii) data on transportation eligibility by address; and (iii) data on house prices and residential zoning regulations. Finally, we incorporate information from the National Center for Education Statistics (NCES) Private School Survey on private-school locations and data from the American Community Survey (ACS) on neighborhood income distribution and private-school enrollment.

We estimate the model using the method of simulated moments, exploiting both longitudinal and cross-sectional variation for identification. Changes in base-school catchment boundaries generate quasi-experimental shifts in peer composition that identify families' willingness to pay for school quality. Sorting of children across schools by income and skills further identifies heterogeneity in the valuation of school quality. Commuting costs and their heterogeneity by income and transportation availability are identified using variation in distance travelled from home to school. Parameters governing private-school costs, which may vary by location and income, are identified from spatial variation in private-school attendance conditional on income. Residual income sorting across neighborhoods identifies (heterogeneity in) families' valuation of neighborhood amenities.

Parameter estimates reveal that economically disadvantaged (henceforth, *ED*) families have a lower willingness to pay than higher-income (or *non-ED*) families for school peer quality. Conditional on income, families' valuation of school quality increases with their child's skills. This gradient is particularly steep for higher-income families. For non-ED families with a child at the 25th percentile of the skill distribution, the equivalent variation associated with a ten percent of a standard deviation increase in peer quality is \$902, and it rises to \$1,629 if the child's skills reach the 75th percentile of the skill distribution. The analogous numbers for ED families range from \$223 to \$295. For ED families, a one-mile

shorter commute to school and transportation provision to all option schools are each worth more than a ten percent increase in peer quality, regardless of children's skills. For non-ED families, the ranking is reversed, with better peers worth up to twice as much as school proximity and transportation.

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. In the place-based policy, we target low-income neighborhoods, where over 40 percent of families are ED, and expand their school portfolios to include the top 40 schools in the district by peer composition (henceforth, *receiving* schools). In the people-based policy, we provide housing vouchers to ED families conditional on residing within the catchment areas of those same 40 receiving schools.

We find that both policies have negative aggregate welfare effects, amounting to average annual losses of $1,030 and $83 per family, respectively. While beneficiaries obtain modest welfare gains—most of which are offset by higher commuting costs in the school choice expansion policy—the aggregate loss is driven by negative effects on families residing in the catchment areas of the receiving schools in the baseline equilibrium. Among them, ED families—who tend to have higher-skilled children than the typical ED family in the economy and a higher willingness to pay for accessing high-quality schools—are more exposed to the decline in school peer quality resulting from the influx of new children into their schools and are less able to insure against it.

In the school choice expansion policy, commuting costs limit applications to receiving schools far from the targeted low-income neighborhoods, despite school transportation provision. These schools, which tend to be in more affluent neighborhoods, are therefore protected from the policy by their geographic location on the outskirts of the city. In the voucher policy, the shielding role of distance is instead played by zoning restrictions. Although incoming children must reside within the catchment areas of receiving schools under the policy, minimum lot-size constraints tilt their neighborhood choice toward less regulated, less affluent neighborhoods.

Related Literature. This paper contributes to the literature on residential choice and the valuation of school quality. Black (1999) shows that elementary-school quality is capitalized into house prices. Epple and Sieg (1999) and Sieg et al. (2004) develop and estimate equilibrium models in which heterogeneous households sort across communities, and evaluate the welfare effects of changes in local public goods. Bayer et al. (2007) develop a unified framework for estimating preferences for school quality separately from preferences for correlated neighborhood attributes. We depart from this literature by allowing

neighborhoods to provide access to a portfolio of schools, using longitudinal changes in school catchment areas to estimate willingness to pay for school quality, and evaluating the welfare effects of local education policies.

Our paper also relates to the empirical literature on school choice. We focus on the demand side of the education market and contribute to the branch of the literature interested in the determinants of school choice by endogenizing families' residential location, largely treated as exogenously fixed in previous work (e.g., Hastings et al., 2009; Abdulkadiroğlu et al., 2017; Agarwal and Somaini, 2018; Kapor et al., 2020; Laverde, 2024).[1]

Our approach to modeling residential sorting in the presence of school choice is in the spirit of Nechyba (2000), Epple and Romano (2003), and Avery and Pathak (2021). Like us, these papers propose equilibrium models of school quality and house prices, and study the extent to which school choice policies increase access to high-quality schools by low-income children. We depart from those studies by showing how home-school distance determines both the take-up rate of school choice and the differential exposure across non-targeted families. Park and Hahm (2023) examine the impact of changes in the school assignment mechanism on neighborhood and school choice, but abstract from equilibrium responses of house prices and school peer composition.

This paper also contributes to the growing urban literature that studies how agglomeration and congestion forces shape city structure and individual outcomes (Ahlfeldt et al., 2015; Redding and Rossi-Hansberg, 2017). Recent contributions have focused on the effect of transportation infrastructure on commuting patterns (Heblich et al., 2020) and neighborhood sorting (Tsivanidis, 2026). Like them, we explore how transportation provision affects access to desired locations within a city. However, unlike much of the urban literature, which focuses on access to labor markets, we study how educational access varies across neighborhoods. Hsiao (2024) studies migration toward labor markets with higher returns to education in spatial equilibrium. Similar to us, Almagro and Domínguez-Iino (2025) show that preference heterogeneity is a key determinant of the distributional effects of changes in endogenous local amenities. They focus on consumption amenities, while our paper is about the quality of local schools.

Recent work in macroeconomics has focused on the impact of spatial sorting 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 devel-

[1] Important contributions to the literature on the supply side effects of school choice include Singleton (2019), Allende (2021), Dinerstein and Smith (2021), and Campos and Kearns (2023).

opment of children's human capital through the transmission of skills across generations (Aliprantis and Carroll, 2018; Fogli et al., Forthcoming; Eckert and Kleineberg, 2024; Chyn and Daruich, 2025). We contribute to this literature by showing how city structure, commuting costs, and local institutions (school assignment rules, school transportation, and zoning) shape children's exposure to school peers.

2 Model

2.1 Overview

We build a static model of neighborhood and school choice by families with an elementary school-aged child within a single county. Families are heterogeneous in terms of both their income and the skills of the child. Families choose their neighborhood of residence taking into account house prices, housing regulations, and exogenous and endogenous amenities. Endogenous amenities are given by the quality of the neighborhood-specific portfolio of schools that families can apply to. Schools are characterized by some exogenous characteristics, namely, their value added, capacity, and location in the city. Schools are also characterized by the endogenous number and skill of the children that apply to them, which determine equilibrium probability of admission and peer composition, respectively. Commuting from home to school entails disutility which varies depending on distance and on whether school transportation is provided.

The county is also inhabited by other households, o, by which we mean households without an elementary school age child. Other households, in contrast to families, are only heterogeneous with respect to their income, and only choose their residential location. Although families are the main focus of our analysis, incorporating other households into the model allows us to measure families' valuation for school quality using changes in house prices, accounting for the fact that families represent only a fraction of the housing demand in the county.

2.2 Environment

Demographics. The county, 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 (other) households does not have children in the relevant age group. Families are of type (w, a), where w is household income and a is the child's skills. The joint distribution over family types is given exogenously by $\phi_{w,a}$. The

other households have an income distribution ϕ_w^o.

Geography. The county is divided into neighborhoods, indexed by n, and is served by a set of public and private schools, indexed by s.

School Portfolios. Motivated by our empirical setting that covers the Wake County Public School System (WCPSS), each neighborhood n is associated with a portfolio of schools $\mathcal{L}_n$. The school portfolio 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 neighborhood portfolio that provides 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, $\mathcal{C}_s = \{n \mid \mathcal{B}_n = \{s\}\}$. The second subset of public schools, denoted by $\mathcal{T}_n$, does not guarantee admission but provides bus transportation to the students enrolled. The third subset of public schools does not guarantee admission or provide bus transportation ($\mathcal{NT}_n$). 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. 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 private schools, $\mathcal{P}_n$. Although 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. In summary, the portfolio of schools available in each neighborhood is $\mathcal{L}_n = \mathcal{B}_n \cup \mathcal{T}_n \cup \mathcal{NT}_n \cup \mathcal{P}_n$. With a slight abuse of notation, we also use $\mathcal{B}_n$ and $\mathcal{P}_n$ to denote the base school and the private school themselves.

Public School Capacity, Admission Probability, Applications. Base schools are mandated to admit students in their catchment area. Therefore, we assume that each school acting as the base has sufficient capacity to guarantee admission to all children who reside in its catchment area. This includes both children who applied directly to their base school and those who did not obtain admission to their preferred option schools through the application lottery.[2] Additionally, each school s reserves a limited number of seats,

[2]The WCPSS manages occasional instances in which capacity is exceeded due to large base enrollment through designated "overflow" partner schools, to which base students may be reassigned if total enrollment surpasses capacity. Such cases are extremely rare: administrative data indicate that, between 2008

denoted by q_s, to children for whom that school is an available option given the neighborhood where they reside, $\{n \mid s \in \mathcal{T}_n \cup \mathcal{NT}_n\}$. 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 $s \in \mathcal{T}_n \cup \mathcal{NT}_n$.[3] If children are not admitted to the option school they apply to, they can choose whether to enroll in their base or private school. Families can always choose to enroll their children in a private school (and be admitted with probability one) without applying to any school in the public school system.[4]

Housing Supply and Zoning Restrictions. Each neighborhood is characterized by an imperfectly elastic housing supply, $H_n = H_{0n} r_n^{\xi_n}$, where r_n is the unit price of housing in neighborhood n, ξ_n is a neighborhood-specific elasticity, and H_{0n} is an exogenous neighborhood shifter. Housing demand is constrained by neighborhood-specific housing regulation in the form of a minimum size constraint, which we simply refer to as *zoning*. In neighborhoods without housing regulations, we assume that the size of a house must meet a common lower bound that captures minimum physical space required for habitation. We denote the constraint on housing demand in neighborhood n, whether regulatory or not, by $\underline{h}_n$.

Preferences: Families. Families have additive preferences over three components of utility: a bundle of housing and non-housing consumption, school characteristics, and neighborhood characteristics.

Preferences over Consumption. Family utility over housing, h, and non-housing, c, consumption is given by

$$\left(\frac{1-\beta}{\beta}\right) \ln\left(\frac{c}{1-\beta}\right) + \ln\left(\frac{h}{\beta}\right). \tag{2.1}$$

and 2019, an average of 0.2 percent of all elementary-school students were assigned to an overflow school in a given year.

[3]Option schools in the WCPSS grant sibling priority, which we abstract from in the model. However, data indicate that only about 4.5 percent of children in our sample have an older sibling enrolled in an option school for which this priority would be relevant. Moreover, we find that base-school attendance rates are similar for children with and without siblings.

[4]At the elementary school level, option schools in the WCPSS include calendar options and magnet programs (see Appendix A.1 for further detail and discussion). As in our model, families may apply to only one calendar option. For magnet programs, during the first two years of our sample (and in all prior years), families ranked up to three schools, and assignments were made under the Boston mechanism. Because priority is given to first-ranked schools and most schools receive more first-choice applicants than they have seats available (Dur et al., 2018), choices ranked below the first were typically irrelevant in practice. In the final two years of our sample, the district switched to Deferred Acceptance (DA), which does not grant priority to families' first choices. However, since families can apply to only five out of an average of 18 options, strategic considerations remain important, consistent with our model.

Families are subject to the following budget and housing constraints

$$\hat{w} = (1+\tau)w \geq c + r_n h + \mathbb{I}_{\text{private}} P_n \tag{2.2}$$

$$h \geq \underline{h}_n \tag{2.3}$$

where the price of c is normalized to one, while each unit of housing costs r_n. The term P_n represents the price paid by families who enroll a child in the private school associated with neighborhood n ($\mathbb{I}_{\text{private}} = 1$).[5] 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 rents, distributed proportionally to households' income w, while the complementary fraction covers the production cost of the structure.

Preferences over School Characteristics. Conditional on living in neighborhood n, family i with income w and a child of skill level a obtains the following utility from attending public school $s \in \mathcal{L}_n \setminus \mathcal{P}_n$,

$$\gamma_{w,a} \left(\ln \bar{a}_s + \theta A_s\right) - \kappa_{s;w} d_{s,n} + \sigma_S \varepsilon_{s;i} \,. \tag{2.4}$$

The variable $\bar{a}_s$ denotes the average skill level of children enrolled in school s. In contrast, A_s captures an exogenous school value-added (henceforth, VA). The joint contribution of peer quality and school VA determines the overall school quality from the parents' perspective.[6] The parameter $\gamma_{w,a}$ governs how families value school quality. We allow this preference to be heterogeneous across families by income and child skills. The parameter θ captures the (relative) valuation for school VA.[7] The term $\kappa_{s;w} d_{s,n}$ denotes the utility cost

[5] Throughout, we treat P_n as a pure resource cost, which amounts to assuming a competitive market for private schools with a constant marginal cost per child and no fixed costs. We explore the robustness of our policy analyses to endogenous changes in private school prices in Section 5.3.

[6] A change in a school's peer composition is likely to trigger changes in other school attributes that respond endogenously to it. For example, more effective teachers tend to be disproportionately drawn to schools with a positively selected pool of students (e.g., Jackson, 2009). Hence, in our counterfactual analysis, while school VA is held constant, we interpret changes in peer composition as reflecting both the direct impact of peers on school demand and their indirect effect through other endogenous school attributes. In treating peer composition as a key determinant of school quality, we join a vast literature, including Bayer et al. (2007), Epple and Romano (2003), and Avery and Pathak (2021), among others.

[7] Similarly to Allende (2021), we account for both peer quality and school VA in families' demand for schools. We also follow an analogous strategy to recover school VA, net of peer composition, from data on student test scores. Differently from her work, our baseline specification does not include an additional unobserved fixed effect for public schools: attendance at those schools is constrained by admission lotteries and school capacity, so attendance shares are only partially informative about school-specific preferences. Section 5.3 assesses the sensitivity of our results to introducing public-school fixed effects. In our baseline specification, we do include unobserved school attributes in the demand for *private schools*, which constitute the main focus of Allende (2021).

associated with commuting. Such cost depends on the home-school distance, $d_{s,n}$, family income w, and whether 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, $\varepsilon_{s;i}$ represents a school-specific idiosyncratic preference shock, which is independent and identically distributed across families and schools according to an Extreme Value Type 1 distribution. The parameter σ_S governs dispersion in preferences.

The value of attending a private school is instead given by

$$\gamma_{w,a} \ln \bar{a}_{\mathcal{P}_n} + \chi_{\mathcal{P}_n;w} - \kappa_{NT;w} d_{\mathcal{P}_n,n} + \sigma_P \varepsilon_{\mathcal{P}_n;i}, \tag{2.5}$$

which is analogous to the value of attending a public school, except for two differences. First, in Wake County, the decision of private schools to offer transportation is decentralized, and we lack detailed geographical data on individual private school transportation provision. Given that the vast majority of private schools do not offer transportation, we make the assumption that none do. Second, school VA is replaced by the unobserved characteristic $\chi_{\mathcal{P}_n;w}$. The different specification is motivated by our lack of data on test scores in private schools, needed to estimate school VA, and the presence of unmodeled features of such schools, with valuation that may vary by family income. Recall that attending private schools also entails a tuition payment P_n, introduced above.

In our empirical context, parents are required to present verifiable proof of residency before applying to schools in their neighborhood portfolio. We assume that families first choose their neighborhood of residence—and their housing consumption—forming expectations over the value of public and private schools, and then choose which school to apply to after observing the realization of their idiosyncratic preference shocks for schools.

Preferences over Neighborhood Characteristics. Family i has utility over exogenous neighborhood characteristics given by

$$\alpha_{0n} + \ln(w)\alpha_{1n} + \sigma_N \varepsilon_{n;i} \,. \tag{2.6}$$

The first term is the value of neighborhood amenities that is common across households, α_{0n}. The second term captures amenities that are valued differently by (ln) family income. Finally, $\varepsilon_{n;i}$ represents a neighborhood-specific idiosyncratic preference shock, which is independent and identically distributed across families and neighborhoods according to

an Extreme Value Type 1 distribution. The parameter σ_N determines the dispersion of this shock.

Preferences: Other Households. The other households have the same type of preferences over consumption and neighborhood characteristics as families, but place no value on school characteristics. We allow amenities, α_{0n}^o and α_{1n}^o, to be different from the analogous parameters for families, and we denote them by the superscript o. These households may place different values on specific local amenities—such as restaurants and gyms—compared to families with children in elementary school. Families and the other households interact with each other via their participation in the housing market.

2.3 Equilibrium

The spatial equilibrium is characterized by individual choices and endogenous neighborhood and school characteristics such as house prices r_n, school peers $\bar{a}_s$, and admission probabilities p_s, together with the transfer rate τ from land income.

Individual Decisions. The housing demand for a family with a child enrolled in a public school is

$$h_{n;w} = \max\{\beta\hat{w}/r_n, \underline{h}_n\}. \tag{2.7}$$

The indirect utility from consumption is then equal to

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

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, as we highlight in our voucher policy analysis in Section 5.

We denote by $u_{\mathcal{P}_n;w}$ the utility from consumption for families whose child attends the private school in neighborhood n. We assume that this utility is identical to that in Equation (2.8), except for $\hat{w} - P_n$ replacing $\hat{w}$. That is, housing demand in Equation (2.7) does not respond to tuition, which instead crowds out non-housing consumption. Consistent with this specification, we show that families with children in private school devote the same share of income to housing as families with children in public school (Table A-4 in Appendix A.2.5). Our results are unchanged under the opposite assumption, with fami-

lies spending on housing out of post-tuition income.[8]

Combining utility from consumption in Equation (2.8) and school attributes, we obtain the value of attending public school s from neighborhood n for family i of type (w, a),

$$v_{s|n;w,a,i} = u_{n;w} + \gamma_{w,a} \left(\ln \bar{a}_s + \theta A_s \right) - \kappa_{s;w} d_{s,n} + \sigma_S \varepsilon_{s;i}. \tag{2.9}$$

The value of attending a private school is instead given by

$$v_{\mathcal{P}_n|n;w,a,i} = u_{\mathcal{P}_n;w} + \gamma_{w,a} \ln \bar{a}_{\mathcal{P}_n} + \chi_{\mathcal{P}_n;w} - \kappa_{NT;w} d_{\mathcal{P}_n,n} + \sigma_P \varepsilon_{\mathcal{P}_n;i}. \tag{2.10}$$

Families who are unsuccessful in their school application receive a value

$$v^{fallback}_{n;w,a,i} \equiv \max\{v_{\mathcal{B}_n|n;w,a,i}, v_{\mathcal{P}_n|n;w,a,i}\}$$

from attending their preferred fallback school. Applying to school s is thus equivalent to choosing a lottery that delivers $v_{s|n;w,a,i}$ with probability p_s and $v^{fallback}_{n;w,a,i}$ with probability $1 - p_s$, where $p_s = 1$ for base and private schools. The implied conditional probability of applying to a given public or private school s is

$$\pi_{s|n;w,a} = \Pr\left(p_s v_{s|n;w,a,i} + (1 - p_s) v^{fallback}_{n;w,a,i} \geq p_{\tilde{s}} v_{\tilde{s}|n;w,a,i} + (1 - p_{\tilde{s}}) v^{fallback}_{n;w,a,i} \;\; \forall \tilde{s} \in \mathcal{L}_n \right),$$

where the probability is induced by idiosyncratic preference shocks over schools.[9] It follows that the expected value of the school portfolio accessible to families in neighborhood n is

$$\bar{v}_{n;w,a}(\mathcal{L}_n) = \mathbb{E}\left[\max_{s \in \mathcal{L}_n} \left\{ p_s v_{s|n;w,a,i} + (1 - p_s) v^{fallback}_{n;w,a,i} \right\} \right]. \tag{2.11}$$

We deem the expression in Equation (2.11) our measure of *educational access*, which is reminiscent of the commuter market access in Ahlfeldt et al. (2015) and Tsivanidis (2026). Intuitively, a neighborhood has higher educational access when its portfolio has many

[8] An intermediate case consists in allowing families to choose housing expenditure based on the expected utility of doing so, given the uncertainty over their attended school due to idiosyncratic school preferences and admission lotteries to option schools. Doing so creates a substantial computational challenge since families' own housing choices would in turn influence the probability distributions over school lotteries in their subsequent discrete choice. The two cases we consider sidestep this unnecessary complication, and effectively bracket the impact of private school prices on families' utility from consumption in each neighborhood.

[9] 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. 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. In addition, we allow the variance of idiosyncratic preference shocks to be different between private and public schools.

schools, with high admission probability if public or low price if private, high-quality peers, and located at short distance.

The indirect utility over neighborhood n for family i of type (w, a) is then given by

$$v_{n;w,a,i} = \hat{v}_{n;w,a} + \sigma_N \varepsilon_{n;i}, \qquad \hat{v}_{n;w,a} = \alpha_{0n} + \ln(w)\alpha_{1n} + \bar{v}_{n;w,a}(\mathcal{L}_n).$$

With Extreme Value Type 1 idiosyncratic preferences, the share of families living in neighborhood n satisfies the canonical logit formula

$$\pi_{n;w,a} = \frac{\exp\left(\frac{\hat{v}_{n;w,a}}{\sigma_N}\right)}{\sum_{\tilde{n}} \exp\left(\frac{\hat{v}_{\tilde{n};w,a}}{\sigma_N}\right)}.$$

It is then straightforward to compute the share of families living in neighborhood n and applying to school s

$$\pi_{s,n;w,a} = \pi_{n;w,a}\pi_{s|n;w,a}\,. \tag{2.12}$$

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

$$\begin{aligned} \hat{v}^o_{n;w} &= u_{n;w} + \alpha^o_{0n} + \ln(w)\alpha^o_{1n} \\ \pi^o_{n;w} &= \frac{\exp\left(\frac{\hat{v}^o_{n;w}}{\sigma_N}\right)}{\sum_{\tilde{n}} \exp\left(\frac{\hat{v}^o_{\tilde{n};w}}{\sigma_N}\right)}\,. \end{aligned} \tag{2.13}$$

Aggregates. Building toward our aggregate outcomes, we define the following auxiliary variables. Let

$$\begin{aligned} \pi_{n;w} &= \Sigma_a \pi_{n;w,a}\phi_{w,a} + m\pi^o_{n;w}\phi^o_w \\ \pi_s &= \Sigma_{w,a}\Sigma_{n\notin\mathcal{C}_s}\pi_{s,n;w,a}\phi_{w,a} \\ \pi^{attend}_{s;a} &= \Sigma_{n,w}\left[p_s\pi_{s,n;w,a}\,\mathbb{I}_{s\in\mathcal{T}_n\cup\mathcal{NT}_n} + \left(\pi_{s,n;w,a} + \pi^{fallback}_{s,n;w,a}\right)\mathbb{I}_{s\in\mathcal{B}_n\cup\mathcal{P}_n}\right]\phi_{w,a} \end{aligned}$$

be the measure of households of type w who live in neighborhood n, the measure of applicants to public school s as an option school, and the measure of children of type a who *attend* school s, respectively. For base and private schools, the measure of attendees must account for children who turn to school s as a fallback school after losing the admission

lottery to their preferred option,

$$\pi^{fallback}_{s,n;w,a} = \sum_{\tilde{s} \in \mathcal{T}_n \cup \mathcal{NT}_n} \pi_{\tilde{s},n;w,a}(1 - p_{\tilde{s}}) \Pr\left(v^{fallback}_{n;w,a,i} = v_{s|n;w,a,i} \,\middle|\, i \text{ applies to } \tilde{s}\right), \quad s \in \mathcal{B}_n \cup \mathcal{P}_n \,.$$

The market clearing condition for housing in neighborhood n reads

$$H_n = \sum_w \pi_{n;w} h_{n;w} \,, \tag{2.14}$$

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

$$\tau = \mu \frac{\sum_n r_n H_n}{\sum_w w \left(\sum_a \phi_{w,a} + m\, \phi^o_w\right)} \,, \tag{2.15}$$

where the denominator is total income in the city (including families and other 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,

$$p_s = \min\left\{\frac{q_s}{\pi_s}, 1\right\}. \tag{2.16}$$

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 a lottery.

The quality of peers is given by the average skills of children in school s,

$$\bar{a}_s = \mathbb{E}_s[a], \tag{2.17}$$

where the expectation $\mathbb{E}_s$ is taken with respect to the (conditional) distribution $\pi^{attend}_{s;a}$.

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

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

- *The choice probabilities $\{\pi_{s,n;w,a}\}$ are induced by families' choice of school, consumption, and neighborhood, and are given by (2.12); the choice probabilities $\{\pi^o_{n;w}\}$ are induced by the other households' choice of consumption and neighborhood, and are given by (2.13);*

- *The housing markets clear,* (2.14)*, the admission probability is consistent with school capacity and applications,* (2.16)*, school composition is consistent with individual choices and admission probabilities,* (2.17)*;*
- *The transfer rate is consistent with the land share of housing expenditure as in* (2.15)*.*

3 Mapping the Model to the Data

Our empirical analysis focuses on Wake County, North Carolina, which serves as a natural setting 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 areas 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 Appendix A.2.

3.1 Data Sources and Measurement

Our dataset combines several sources. Student-level data, which include school attended, end-of-grade test scores in reading and mathematics from grades 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, the NCERDC also reports scores obtained at literacy assessments administered at the beginning of the kindergarten year. Our sample includes children entering kindergarten between fall 2013 and fall 2016, the four cohorts for which we observe both residential addresses and kindergarten test scores. In what follows, cohort-specific variables (e.g., the distribution of ED status and skills, the number of option seats by school, house prices) are further indexed by t.

To construct our measure of child skills (a), we use the DIBELS Composite score of the beginning-of-year kindergarten literacy assessment, standardized over the pooled cohorts and rescaled to unit standard deviation within each cohort. We discretize the resulting index of log-skills into a 10-point support common across all years by computing the median log-skill within each decile of the pooled distribution. While the support points

are fixed over time, the empirical distribution associated with each point is allowed to vary by year. We then exponentiate these points to define children's skill types in levels, $a \in \{a_1, \ldots, a_{10}\}$. For each public school s, the number of seats reserved for option (as opposed to base) students in year t, $q_{s,t}$, is measured as follows. If s is oversubscribed in year t, then $q_{s,t}$ is set to the count of kindergarten students attending s as an option in that year. If it is not oversubscribed in year t but is in other years, then $q_{s,t}$ is set to the larger of its option enrollment in year t and the average number of kindergarten students attending s as an option in the years it is oversubscribed. If it is never oversubscribed, we set $q_{s,t}$ to the school's option enrollment in year t.

The 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 this institutional design to define neighborhoods in our city. Specifically, a neighborhood n is the set of all contiguous addresses that share, in all 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-5 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. 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 annuitized average price per square foot observed in neighborhood n in year t as our measure of neighborhood-level cost of housing services, $r_{n,t}$. 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 in charge of regulating different land areas within Wake County.[10] In the model, we assume that 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, $\underline{h}_n$, via a linear quantile regression of the size of the smallest houses sold in a neighborhood on its MLS, which traces the lower envelope of house sizes across levels of regulation.

Because the NCERDC data provide information about children's ED status, rather than family income, we use the American Community Survey (ACS) 5-year estimates (2013–17) to measure income for families and other households. We aggregate the sixteen income brackets provided by the ACS into ten discrete household income groups,

[10] Our zoning map reflects 2020 boundaries, the earliest year with consistent GIS records across zoning authorities. Historical data indicate that residential zoning was largely stable between 2012 and 2020, with changes affecting at most 2.7 percent of total land area (3.2 percent of residential or mixed-use land).

$w \in \{w_1, \ldots, w_{10}\}$.[11] We convert each group's mean gross income into net-of-tax income using the NBER TAXSIM calculator. Mapping our neighborhoods to 2010 Census tracts, we use tract-level information from the ACS to obtain the distribution of the other households' types, $\phi^o_{w,t}$, as well as their sorting across neighborhoods, $\pi^o_{n;w,t}$. 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, $\phi_{w,a,t}$, and their sorting across neighborhoods, $\pi_{n;w,a,t}$. Namely, we derive the distribution of family income conditional on residential neighborhood and an ED status imputed from ACS income bins, which we then combine with the distribution of children's skills conditional on residential neighborhood and ED status available in the NCERDC. The number of families in each neighborhood comes from the kindergartners living there, adjusted to include private-school families.

Finally, we draw on the biennial National Center for Education Statistics (NCES) Private School Survey for 2013, 2015, and 2017, together with the private school registry maintained by the North Carolina Department of Administration, to identify the universe and locations of private schools in Wake County. Because private school tuition is not systematically recorded in existing datasets, we compile tuition information from two complementary sources. First, from the 2014–15 school year onward, we obtain tuition data for schools participating in the Opportunity Scholarship Program from the North Carolina State Education Assistance Authority (NCSEAA). Second, we augment these data by hand-collecting tuition information from private school websites and archived webpages accessed through the Wayback Machine. For each neighborhood n, we let the private school associated with the neighborhood, $\mathcal{P}_{n,t}$, be the nearest private school to the neighborhood centroid that is open in year t. We use Public Use Microdata Area (PUMA)-level information from the ACS 1-year samples (2013–17) to construct, for each of the eight PUMAs constituting Wake County, the share of kindergartners attending private school as a function of income in our sample period.

[11] 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 2017–18 for reduced-price lunch were $45,510 annual income for a household of four.

3.2 Final Sample

Our final sample consists of 34,785 students across four kindergarten cohorts.[12] Detailed descriptive statistics about the samples of schools, students, and neighborhoods are reported in Table A-3. Twenty-eight percent of families with a kindergartner are economically disadvantaged. 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 74 percent of students attend their base school.

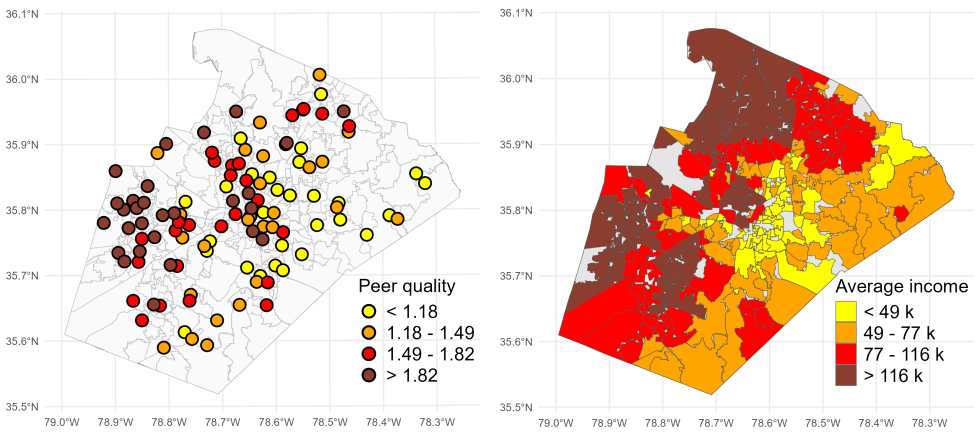
There are 111 public elementary schools in our sample. While all 111 schools are base schools, 94 (85 percent) are also option schools for some neighborhoods. On average, 64 percent of option schools are oversubscribed each year. The share of ED students ranges from zero to 0.90 across schools, with a mean of 0.38 and a standard deviation of 0.22. In addition, 11.1 percent of Wake County kindergartners attend private school. Our final sample of private schools consists of 36 schools with a median annual tuition of $7,512 over the sample period (see Figure A-4 for their location).

Figure 1 shows the geography of neighborhood income and public school peer quality. The two maps align notably well—higher-income neighborhoods and schools with higher-skilled peers are located in the west and in parts of the center of the county, while lower-income neighborhoods and schools with lower-skilled peers can be found downtown and east. Average house prices follow a very similar pattern (Figure A-6a). Finally, zoning regulations are mapped in Figure A-6b. There is significant heterogeneity in MLS regulations across neighborhoods, with higher-density neighborhoods concentrated in the urban center of the county, the city of Raleigh. Accordingly, there is significant variation in our measure of minimum house size $\underline{h}_n$ across neighborhoods (mean of 619 square feet, ranging from 466 to 1,474 square feet).

On average, each neighborhood is associated with a choice set of 18 option schools in addition to its base. Option schools tend to be much farther away from the neighborhood than the base is (11.1 versus 3.6 miles on average). Twenty-two percent of neighborhoods

[12] We drop student observations for which school attended, address, kindergarten literacy Composite score, or ED status is missing, which leaves 36,690 students. Students attending a school that was in the choice set of their address within the six years prior, but not in their kindergarten year (638 students), are assumed to be grandfathered to follow an older sibling: we keep them as residents of their neighborhood but exclude them from statistics on school attendance. When the school attended was never in the choice set of the student's address, we reassign the student to a neighborhood within half a mile whose choice set includes that school, if there is one (the NCERDC adds noise to residential coordinates), and drop the student otherwise (1,311 students). Finally, we drop neighborhoods with fewer than ten students over the four years or with no student in one of the years (594 students). This leaves 34,785 students as residents. Statistics on school attendance further exclude the grandfathered students remaining in these neighborhoods (also 594 students: the 638 above, less 44 in dropped neighborhoods).

Figure 1: Geography of School Quality and Neighborhood Income



(a) Public school peer quality (b) Neighborhood income

Each map in this figure shows Wake County, NC. The left map shows school-level peer quality for each school in the sample. Our measure of peer quality is the average skill level a among enrolled kindergartners in the sample. 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 after-tax household income by neighborhood. 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.

experienced a change in base school over the sample period, while all neighborhoods experienced a change in eligibility 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. The existing literature documents that catchment-area reassignments in Wake County were essentially random conditional on observable neighborhood traits and were not anticipated by residents (Hill et al., 2023). In Appendix A.1.2, we review this literature and formally show that these changes are not systematically related to pre-policy trends in neighborhood house prices or to the characteristics of the housing stock.

4 Model Identification and Estimation

We set some parameters outside the model using external sources or established values from the literature, and we estimate the remaining parameters using the Simulated Method of Moments (SMM). Appendices C.1 and C.2 detail the construction of each targeted moment in the data and in the model, respectively.

The utility weight on housing, β, and the land share of housing expenditure, μ, are both set equal to 0.25, in line with typical values in the literature (Davis and Ortalo-Magné, 2011, and Ahlfeldt et al., 2015). School VA, A_s, is obtained by regressing third-

grade test scores on children's own kindergarten test scores, school-peer average kindergarten test scores, and a school fixed effect—this fixed effect capturing the school VA. Neighborhood-level housing supply elasticities, ξ_n, are obtained by cross-walking the census-tract level housing-supply elasticity estimates from Baum-Snow and Han (2024). Appendix A.2 provides further details on the estimation of A_s and the construction of ξ_n.

We estimate the following parameters within the context of the model: preferences for school quality, $\gamma_{w,a}$, and relative contribution of value added, θ; commuting costs, $\kappa_{T;w}$ and $\kappa_{NT;w}$; amenities for families and other households, α_{0n} and α_{1n}, and α^o_{0n} and α^o_{1n}, respectively; unobserved utility of attending a private school, $\chi_{\mathcal{P}_n;w}$; and the variance of preference shocks, σ_S, σ_P, and σ_N. We additionally obtain housing supply shifters for each neighborhood, H_{0n}.

While there is no one-to-one mapping between parameters and moments, we offer an intuitive argument for how each parameter affects a particularly informative key moment. In Appendix C.3, we report a set of plots that illustrate how the targeted moments are strictly monotone in the associated parameters, in line with the arguments presented in this section. This connection is also the basis of our SMM algorithm.

4.1 Identification

Preferences for School Quality. School quality in our setting is captured by two elements: peers $\bar{a}_s$ and value-added A_s. We let preferences over this school-quality bundle vary by ED status and skill type a, yielding two vectors of parameters, $\{\gamma_{\text{ED},a}\}$ and $\{\gamma_{\text{non-ED},a}\}$, with $a \in \{a_1, \ldots, a_{10}\}$. These parameters are identified by two main sources of variation.

First, the extent to which house prices vary with changes in school quality induced by the redrawing of school catchment areas is informative of the *average* willingness to pay for school quality. Our setting allows us to exploit this source of variation by leveraging policy-induced boundary changes affecting the same neighborhood over time. This approach contrasts with previous studies, which primarily rely on cross-sectional variation across catchment area boundaries (see, for example, Black, 1999; Bayer et al., 2007). We believe longitudinal variation to be more robust to potential biases arising from unobserved preference heterogeneity across families residing on opposite sides of the boundary. 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.

Formally, we estimate the following regression, both in the data and in the model:

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

where $\Delta_n \ln r_{n,t}$ represents the within-neighborhood change in (log) 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 with school $\mathcal{B}_n$ between two consecutive cohorts of kindergarten children, in $t-1$ and t. Recall the definition of the catchment area of school s, $\mathcal{C}_{s,t} \equiv \{n|\mathcal{B}_n = s \text{ in school year } t\}$. The policy-induced change in quality is defined as follows,

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

The variable $\widehat{\bar{a}}_{\mathcal{C}_{\mathcal{B}_n,t},t-1}$ is the mean skill level 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 $\mathcal{C}_{\mathcal{B}_n,t}$ if the distribution of children's skills across neighborhoods remained 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 a base school's 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 sorting) constant. This variation is cleansed from changes in catchment area composition that are due to endogenous families' relocation in response to the policy. To recover the aggregate effects on prices, regressions are weighted by the number of house sales in the neighborhood and year in the Wake County real estate transaction records. Our preferred estimate of the price-response coefficient b is 0.073 (see Column (2) in Appendix Table B-1), suggesting that a 10 percent increase in school peer quality leads to a 0.73 percent increase in house prices.

Next, we compute the average within-neighborhood difference in the mean peer quality of schools attended by ED and non-ED students. This moment isolates systematic differences in school choices across income groups while holding residential location fixed, and is therefore directly informative about heterogeneity in preferences for school quality by ED status. Last, within ED-group, we construct the ratio of the mean peer quality of the schools attended by skill level a children to that of the lowest-skill level, a_1.[13]

Valuation for School Value Added. The parameter governing the weight placed by families on school VA relative to peer composition in the overall school-quality bundle, θ, is identified by targeting the correlation between school peer composition and VA, which is equal to 0.68 in the data. In the model, the heterogeneous coefficients $\gamma_{w,a}$ capture

[13] Our identification of heterogeneity in γ within ED-group uses variation in school composition by skill type induced both by school and neighborhood choice. This is because many neighborhoods contain insufficient observations for some skill types conditional on ED-group. Instead, we recover heterogeneity in γ by income from within-neighborhood variation in school choice, which is unaffected by other income-specific determinants of neighborhood choice.

families' willingness to pay for school quality. When higher-skill families place greater value on school quality—as our estimates indicate—a larger value of θ amplifies sorting of high-skilled children into high VA schools. Conversely, a smaller (or negative) θ attenuates demand for high-VA schools, especially among high-skill families, reducing the correlation between peer quality and value added.

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, $\kappa_{T;\text{ED}}$ and $\kappa_{T;\text{non-ED}}$, and the per-mile utility loss from commuting to an option school without busing or to a private school, $\kappa_{NT;\text{ED}}$ and $\kappa_{NT;\text{non-ED}}$. For each ED-status, 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 determining school choice and—all else equal—longer commutes.

Idiosyncratic Public 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 variance 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 (in peer composition, value added, and distance) play a smaller role in school choice.

Private School Valuations and Idiosyncratic Preferences. Our student-level data do not provide information on children attending private schools. Residual demand for private schools—net of tuition, peer quality, and distance—is governed by a systematic component, denoted by χ. To discipline this component empirically, we supplement our main data with information from the American Community Survey (ACS). Given that the ACS reports private-school attendance shares by income group only at the PUMA level, and each model neighborhood maps to a single PUMA, we index χ by PUMA of residence rather than by the private school associated with each neighborhood. Within the limits of the small sample of households with kindergarten-age children in private schools in the ACS, we allow χ to differ across three income groups: families earning below \$50,000; between \$50,000 and \$125,000; and above \$125,000. We then choose the values of χ to match the observed private-school attendance shares for each PUMA–income group.

Using data from the Wake County Public School System in the 2015–2016 school year, Dur et al. (2025) report that applicants who lose an option-school lottery are ten percent-

age points less likely to enroll in public schools than lottery winners. We use this moment to identify the contribution of idiosyncratic preferences to private school demand, σ_P. The intuition is that, conditional on the overall share of children enrolled in private schools, higher values of σ_P increase the likelihood that families choose private schools ex ante, rather than only after losing a public-school lottery.

Preferences for Neighborhood Characteristics. The exogenous neighborhood amenities are set to replicate the population shares and average incomes of the families and other households residing 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 neighborhood attributes like school quality, house prices, and zoning restrictions.

Idiosyncratic Neighborhood Preferences. To identify the dispersion in idiosyncratic neighborhood preferences, σ_N, we use empirical findings from the Section 8 arm of the Moving to Opportunity (MTO) experiment, as reported in the program's final impact evaluation (Sanbonmatsu et al., 2011). We replicate the MTO experiment within our model. Families with income below the poverty line ("poor") living in neighborhoods with more than 40 percent of poor households are offered a housing voucher that covers the gap between a fraction β of their income and the 40th percentile of the rent distribution. Families can use the voucher if they choose to live in a neighborhood where the share of poor households is below 40 percent. The value of σ_N is then chosen to replicate the take-up rate of 63 percent.[14]

Neighborhood Housing Supply. The housing supply function in each year t is $H_{n,t} = H_{0n} r_{n,t}^{\xi_n}$. The parameter H_{0n} is informed by the average price in each neighborhood over the sample years. Since housing is a durable good, the existing stock adjusts slowly and asymmetrically (Glaeser and Gyourko, 2005). Therefore, in our estimation based on four consecutive years of data, we assume that in each year the housing supply in the neighborhood is equal to the maximum between $H_{n,t}$, given by the long-run supply function above, and a fraction $(1-\delta)$ of the housing stock in the previous year, $H_{n,t-1}$, with $\delta = 0.02$. This asymmetry between housing expansion and contraction is quantitatively negligible in our empirical setting given the sustained population growth in Wake

[14]Our target moment is the average take-up rate across the entire MTO experiment, reported in Exhibit 1.4, p. 15, of Sanbonmatsu et al. (2011). We performed a sensitivity analysis over the range of take-up rates across the five cities where the MTO was implemented (45% in New York City, 52% in Boston, 67% in Chicago, 75% in Los Angeles, and 77% in Baltimore) and obtained quantitatively similar counterfactual results. The MTO experiment identifies σ_N for families with children but is silent about other households, for which we set $\sigma^o_N = \sigma_N$.

Table 1: Estimated Parameters

School Peer Quality	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	0.82	0.79	0.81	0.88	1.03	1.13	1.15	1.11	1.28	1.73
Non-ED Families	0.53	0.77	0.85	1.04	1.12	1.22	1.35	1.58	1.94	2.78

Commuting Cost	κ_{NT}	κ_T
ED Families	0.15	0.07
Non-ED Families	0.09	0.04

School Value-Added	θ
All Families	0.06

Idiosyncratic Preferences	σ_S	σ_P	σ_N
All Families	0.29	0.72	0.93

The table shows estimates for preferences over school peer quality, commuting costs, school VA, and dispersion parameters for utility shocks over public schools, private schools, and neighborhoods.

County during our period of analysis. We adopt the long-run housing supply function in our counterfactual analyses.

4.2 Estimation Results

Estimates of the parameters governing the valuation of school peer composition, school VA, commuting costs, and the scale of idiosyncratic preferences are shown in Table 1. Estimates for exogenous neighborhood amenities and private-school valuations are shown in Figures C-2 and C-3 in Appendix.

The valuation for school peers increases with children's own skill levels, and this gradient is substantially steeper for high-income families than for low-income (ED) families. One interpretation of this result is that children's kindergarten skills reflect innate ability, and that parents place greater weight on school peer quality depending on their child's innate ability. An alternative explanation is that the heterogeneity in skills arises from unequal parental investments during early childhood. Parents who place greater weight on the quality of kindergarten school peers might be those who invested in their children's skills early on, generating the observed positive correlation between γ's and our measure of skills, particularly among higher-income families (Heckman and Mosso, 2014). It is natural to conjecture that both initial conditions and endogenous investments in early life contribute to the realized heterogeneity in skill endowments upon entering kindergarten. Assessing the relevance of each, while certainly valuable, is beyond the scope of this paper. Turning to the estimate of θ, we find a positive valuation for school VA, although it is more than an order of magnitude smaller than the valuation of school peer quality.

With respect to commuting costs, our estimates indicate that families place substantial value on the provision of school transportation. The marginal cost of commuting an additional mile to a school without transportation is more than twice as large as the

Table 2: Equivalent Variation

	25th percentile child skills		75th percentile child skills	
	ED families	Non-ED families	ED families	Non-ED families
2017$ EQUIVALENT				
School peer quality up by 10% of std	223	902	295	1,629
School value-added up by 10% of std	8	43	12	84
Distance to public schools down by 1 mile	486	1,160	473	1,165
Transportation to all options	411	801	336	468
Private school disutility down by 10%	31	417	39	767

Equivalent variation, in 2017 dollars, of changes in school attributes (see text). Columns split families by ED status and by the child's skill (25th and 75th percentiles within ED status). The compensating variation differs from these numbers by less than 1 percent.

marginal cost when transportation is provided. Importantly, attending schools located far from home remains costly even when transportation is available, plausibly reflecting non-monetary costs such as earlier morning schedules.

All but one of the estimates of the unobserved utility of attending a private school, χ, are negative. We conclude that, holding all else constant, and net of monetary costs, families tend to prefer public to private schools. The geographic distribution of such estimates is in Figure C-3 in Appendix. In addition, we estimate the variance of idiosyncratic preferences for private schools to be larger than that for public schools (0.72 vs. 0.29 in Table 1), consistent with unmodeled determinants of private-school choice such as religious affiliation.

Preferences for neighborhoods are determined by exogenous amenities and idiosyncratic shocks. Figure C-2 shows the geographic distribution of estimated amenities for families and other households as a function of income. We highlight two features of these estimates. First, regardless of ED status, there is substantial dispersion in preferences for neighborhood amenities. Second, non-ED families exhibit a pronounced preference for neighborhoods in the western part of Wake County, where high-quality schools are also located. Preferences among ED families appear to be less spatially concentrated. As shown in Table 1, we also find that the variance of idiosyncratic preferences for neighborhoods exceeds that for both private and public schools—likely reflecting numerous idiosyncratic determinants of neighborhood attractiveness, such as proximity to the parents' workplace or to children's grandparents.

Interpreting the Estimates. To facilitate interpretation of the estimated parameters, Table 2 reports the equivalent variation of specific improvements in school characteristics. This is the increase in annual income that would make a family as well off as the improvement itself. Each change is simultaneously applied to every school in the county, holding all other variables fixed, and it is valued ex ante, before the family's idiosyncratic school and neighborhood taste shocks are realized. We report our results for ED and non-ED families

and for the 25th and 75th percentiles of the skill distribution within ED status. For ED families, these correspond to the 2nd and 6th deciles of the population skill distribution, while for non-ED families they correspond to the 4th and 9th.

The results in the first row illustrate how valuations of school peers vary across families. For high-income families, having a high-skilled child is associated with an 81 percent higher willingness to pay for a uniform increase in peer quality equal to 10 percent of its cross-sectional standard deviation ($1,629 versus $902). The skill gradient is also present, though flatter, among low-income families: their willingness to pay for the same increase rises from $223 to $295. Our finding that parents of higher-achieving children place greater value on peer quality echoes Epple et al. (2018).

The valuation for school VA is lower than for peers across the entire distribution of skills and income. Families' willingness to pay for a 10 percent standard-deviation increase in school VA ranges from $8 to $84. These results are consistent with Abdulkadiroğlu et al. (2020), who study rank-ordered school choices in New York City's centralized high school assignment system and show that parents place substantial weight on peer quality, while—controlling for peer composition—school VA plays a limited role in shaping parental preferences.

Regarding distance and bus transportation (third and fourth rows of Table 2), non-ED families with lower-skilled children value a one-mile shorter commute more than a better peer composition, while for those with higher-skilled children peers are worth about 1.4 and 3.5 times as much as distance and transportation provision, respectively. ED families instead consistently prioritize proximity and transportation: their willingness to pay for a shorter distance to school exceeds that for peer quality by 60 to 120 percent. Regardless of income, families with lower-skilled children are also willing to pay more for transportation provision to all option schools—compared with families with higher-skilled children—as transportation is more likely to be their marginal determinant of choice.

Changes in the parameters describing unobserved (dis)utility for private schools are one order of magnitude more important for non-ED families, the main consumers of such schools.

5 Policy Counterfactuals

Our model estimates shed light on the key determinants of residential and school choice. Heterogeneous parental preferences over both exogenous and endogenous school attributes (i.e., educational access), as well as neighborhood characteristics, interact with zoning regulations to generate the observed patterns of residential sorting. The pres-

ence of neighborhood-specific school catchment areas and the unequal provision of school transportation across neighborhoods translate residential sorting into school sorting.

In this section, we study the aggregate and distributional effects of two widely debated types of policy aimed at reducing inequality in educational access in the United States. We first examine a place-based policy that expands school choice for lower-income neighborhoods by adding high-quality schools, predominantly located in higher-income neighborhoods, to their choice sets. We then consider a people-based policy that offers housing vouchers to low-income families who relocate into the catchment areas of the same high-quality schools—that is, neighborhoods these schools serve as their base.

We quantify the effect of the policies on family welfare as the equivalent variation in units of annual income.[15] Specifically, let ΔV_i be the change in value for family i and $\mathcal{MU}_i$ its marginal utility of income. The welfare impact of the policy per family, with I the number of families, is approximately given by

$$\Delta W = \frac{1}{I} \sum_i \frac{\Delta V_i}{\mathcal{MU}_i}. \tag{5.1}$$

In addition, we break down ΔV_i into its components: school peers, school VA, commuting distance to school, unobserved preferences for public schools, unobserved preferences for private schools, neighborhood amenities, and consumption (both housing and non-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 components.

Before proceeding further, we make four remarks about our counterfactual analysis. First, we apply our policies to the institutional environment of the 2015–16 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 essentially the same results. Second, we report our welfare analysis only for families. The other households are affected by each policy only via its impact on house prices. While some gain and some lose from the policies, the net welfare changes for the other households are much smaller than those experienced by families. Third, while we report the monetary costs of providing additional school transportation or a housing voucher, we do not introduce a tax on households in order to balance the government budget. Since we find aggregate welfare losses from both policies, imposing a tax would only strengthen our results. Interestingly, the monetary

[15] Using the compensating variation would change average welfare effects by less than 1 percent.

cost is remarkably similar between the two policies. Last, we examine the possibility of multiple equilibria arising from complementarities between school and neighborhood choices. While in the baseline economy we naturally require our equilibrium to match a wide range of empirical moments, a similar selection device is missing in counterfactuals. In Appendix D.6 we show that multiple equilibria might exist, but our results are robust across all equilibria we identify.

5.1 Expanding School Choice

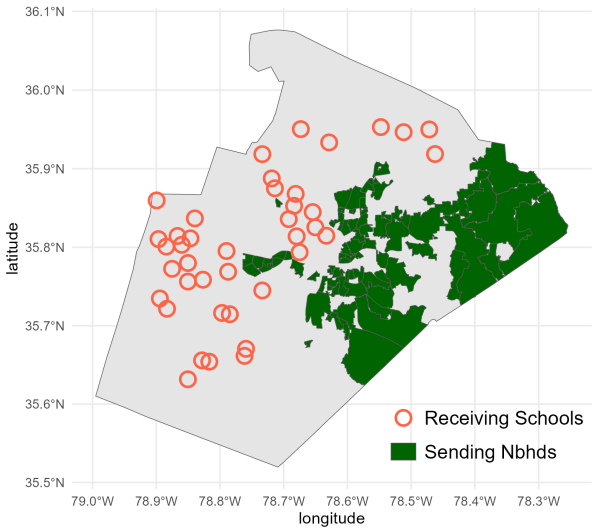
Policy Design. There has long been extensive debate surrounding school-choice policies aimed at decoupling residential location from school assignments (Epple and Romano, 2003; Cullen et al., 2006; Hastings and Weinstein, 2008; Deming et al., 2014). In our counterfactual, we first identify the 40 public schools with the best skill composition, $\bar{a}_s$, in the baseline equilibrium.[16] We refer to these schools—and to the neighborhoods in their catchment areas—as *receiving*. We then add these receiving schools as options with transportation to the school portfolio of neighborhoods with the highest shares of ED families in the county (above 39 percent). We refer to these as *sending* neighborhoods.[17] 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 enrollment. When a school receives more applications than available option seats, those seats are allocated via lottery, as described in the model section. 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 28 percent. Finally, to (partially) mute the well-known "love for variety" effect—namely, the mechanical welfare increase in random utility models that arises when the choice set expands—families in sending neighborhoods draw idiosyncratic preference shocks for newly added receiving schools, with replacement, from the subset of shocks already realized (at baseline) for schools in their original portfolios.[18]

[16] Given our findings on the central role of peer composition in parental preferences, selecting schools according to $\ln \bar{a}_s + \theta A_s$—the combined value of peers and value added—would select the same set of schools.

[17] Some lower-income neighborhoods in the city center have one of the receiving schools as base school at baseline. These neighborhoods are not included in the *sending* group despite their large share of ED families because their choice set is not changed by the policy. Some receiving schools are already in the choice set of some sending neighborhoods as options at baseline; in this case, transportation is added if it is not available. At baseline, 1.8 percent of ED (3.3 percent of non-ED) children in sending neighborhoods attend a receiving school as option.

[18] Simple adjustments for the number of alternatives—such as those used in standard logit models—are inadequate in our setting, where each extra option represents a lottery rather than a standalone choice. We have also performed this counterfactual with families drawing a new i.i.d. utility shock for each receiving school added to the choice set of sending neighborhoods. As expected, the welfare gains attributable to

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.

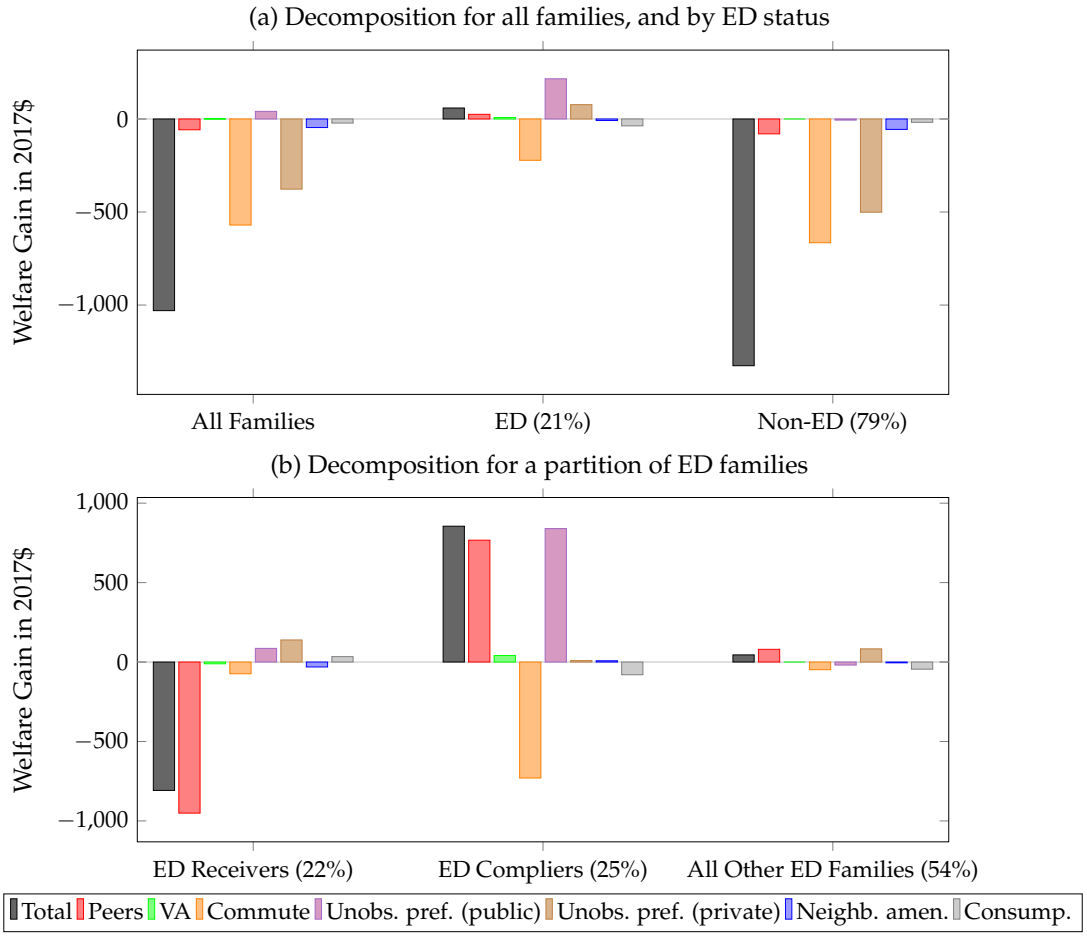
Figure 2 illustrates the spatial distribution of sending neighborhoods and receiving schools. Sending neighborhoods are concentrated in the central and eastern parts of the county, whereas receiving schools are predominantly located in the west. This east–west spatial divide closely mirrors the income distribution shown in Figure 1, with central and eastern areas featuring lower average income, and the west being more affluent.

Results. ED families account for 53 percent of families in sending neighborhoods at baseline (Table 3). Under the policy, their share declines to 50 percent, reflecting the increased attractiveness of these neighborhoods to higher-income families ("gentrification"). Take-up is substantial, with 19 percent of ED (27 percent of non-ED) children in sending neighborhoods enrolling in receiving schools, despite an average admission rate of 35 percent. While non-ED families in sending neighborhoods have children with skills that are 17 percent below the typical child in receiving neighborhoods, the corresponding gap for ED families is 59 percent.

Figure 3 illustrates the impact of the policy on overall welfare (leftmost bar, in black) and on its components (colored bars), as defined in Equation (D-1), across different family groups. Figure 3a reports the average welfare gains for all families and separately by ED status. The policy induces an average welfare loss of $1,030 per family per year (left panel). Losses are particularly large for non-ED families (−$1,326, right panel) whereas

idiosyncratic public-school preferences are larger under this alternative, yet the impact on all other welfare components remains very similar to our baseline policy.

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



The top chart shows the welfare changes induced by the school choice expansion for all families (left), ED families (center), and non-ED families (right); the bottom chart, for a partition of ED families into *Receivers* (living in receiving neighborhoods at baseline, left), *Compliers* (living in sending neighborhoods at baseline and induced by the policy to apply to receiving schools, center), and all other ED families (right). For each group, the black bar is the total average welfare change, and the colored bars decompose it into the policy-induced changes in, from left to right: school peers ($\gamma_{w,a} \ln \bar{a}_s$), school value added ($\gamma_{w,a}\theta A_s$), commuting costs ($\kappa_{s;w} d_{s,n}$), idiosyncratic preferences for public schools ($\sigma_S \varepsilon_s$), net value of private schools ($\chi_{\mathcal{P}_n;w} + \sigma_P \varepsilon_{\mathcal{P}_n}$), neighborhood amenities ($\alpha_{0n} + \ln(w)\alpha_{1n} + \sigma_N \varepsilon_n$), and consumption ($u_{n;w}$). The formal decomposition is reported in the Appendix.

ED families experience a very small gain (+\$59, middle panel). Strikingly, improvements in peer quality contribute little to the average welfare gains of ED families.

The average outcome masks significant heterogeneity among ED families, in particular as a function of their location choice in the baseline equilibrium. Figure 3b shows how such heterogeneity in sorting across neighborhoods maps into unequal welfare changes for three distinct subsets of ED families—those residing in receiving neighborhoods at baseline ("ED receivers"); those residing in sending neighborhoods at baseline who apply to a receiving school after the school choice expansion ("ED compliers"); and all other

Table 3: Heterogeneity among ED Families across Baseline Neighborhood Choices

	Nbhd. chosen at baseline		
	Receiving	Sending	Other
Average share of ED families (%)	10	53	20
Average house price for ED fam. ($ per sqft)	129.16	99.73	120.79
Average child skills cond. on ED (std. child skill distr.)	0.96	0.85	0.90

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 (level, a) among ED families conditional on their baseline neighborhood choice, as well as the average house price they pay (per square foot) at baseline.

ED families.

On the one hand, as one would expect, ED compliers experience a sizeable increase in welfare. Notably, however, the welfare gains from access to public schools with better peers are largely offset by the disutility associated with the longer commute to receiving schools (averaging $+2.3$ miles). Recall that non-ED families in sending neighborhoods are equally eligible given the place-based nature of the policy, hence they too incur longer and costly commutes. 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. As detailed in Table 3, these families spend more on housing to access higher-quality base schools and are positively selected in terms of their children's skills compared with ED families who choose sending neighborhoods at baseline. By and large, ED receivers fare more like the average non-ED family than like the target ED families. Indeed, they experience an even larger decline in peer quality than non-ED receivers (Table 4), owing to both their greater exposure to the policy and their higher cost of mitigating its adverse effects.

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

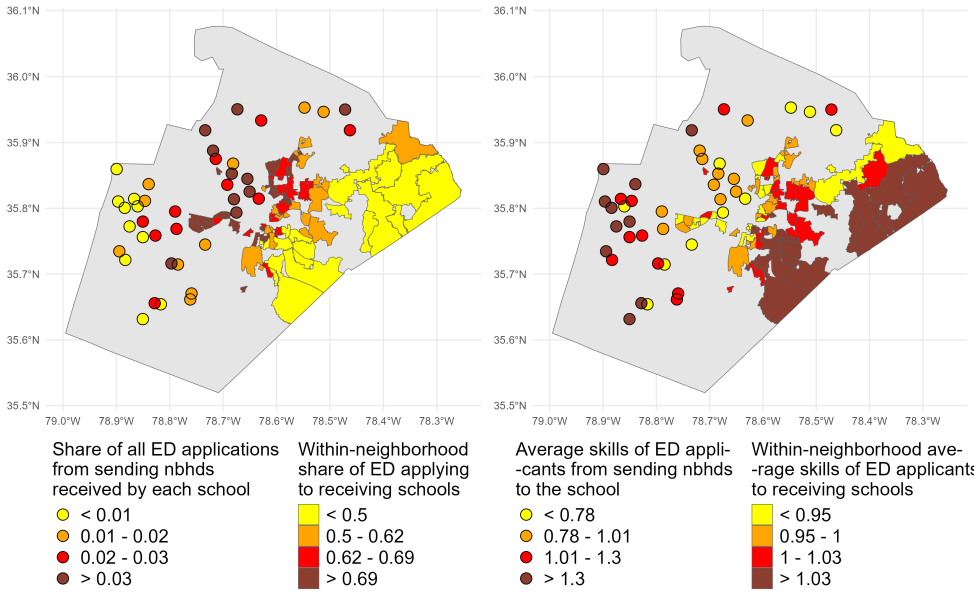
High- and low-income families in receiving neighborhoods also differ in terms of *responses* to the policy. To mitigate their welfare loss from decreased peer quality, non-ED

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

	ED Receivers	Non-ED Receivers
Change in average school peers (% std)	−40.0	−13.3
Change in share attending private school (pp)	1.9	4.9
Share locating outside receiving nbhds (%)	13.4	8.6

The table 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.

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



The figure shows the spatial distribution of applications to receiving schools by ED children in sending neighborhoods. Applicants are ED children in sending neighborhoods whose first choice is a receiving school; not all are admitted. In the left map, each sending neighborhood is shaded by the share of its ED children who apply to a receiving school, and each receiving school (dot) is shaded by its share of all such applications. In the right map, sending neighborhoods and receiving schools are shaded by the average skill of these applicants, among those who live in the neighborhood or apply to the school, respectively. Figure D-1 shows analogous maps for non-ED takers.

receivers turn to private schools more than ED families (+4.9 pp vs. +1.9 pp, see Table 4), bearing the associated monetary cost and disutility from commuting. Instead, 13.4 percent of ED families residing in receiving neighborhoods at baseline relocate to gain access to alternative public schools, typically closer to home than private schools and more often providing transportation.

The policy we consider in this section is just one possible approach to school choice expansion. In Figure D-2, in Appendix, we illustrate the welfare effects of two alternative policies that differ from the one considered here in the set of targeted sending neighborhoods. Specifically, we offer receiving schools as options with transportation to two alter-

native groups of neighborhoods, defined by their share of ED families (between 19 and 39 percent, or below 19 percent).[19] Less progressive policies yield smaller welfare losses because they target higher-ability children, involve lower commuting costs (as sending neighborhoods are closer to receiving schools), and induce less switching of receiving families to private schools.[20] Finally, adopting a utilitarian criterion that incorporates redistributive concerns preserves the welfare ranking between the three policies, with the most progressive one still generating the largest losses.

5.2 Housing Vouchers

Policy Design. Our second policy counterfactual involves housing vouchers, a people-based policy that has received considerable attention among both academics and policymakers (Kling et al., 2007; Chetty et al., 2016; Aliprantis and Richter, 2020; Bergman et al., 2024). 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-peer quality receiving schools identified in the school choice expansion policy.[21] Housing vouchers have the potential to address a key shortcoming of the school choice expansion: by allowing low-income families to reside near high-quality schools—including in the more affluent areas at the outskirts of the city—they remove the commuting barrier that reduces compliers' gains and that concentrates the inflow into city-center schools.

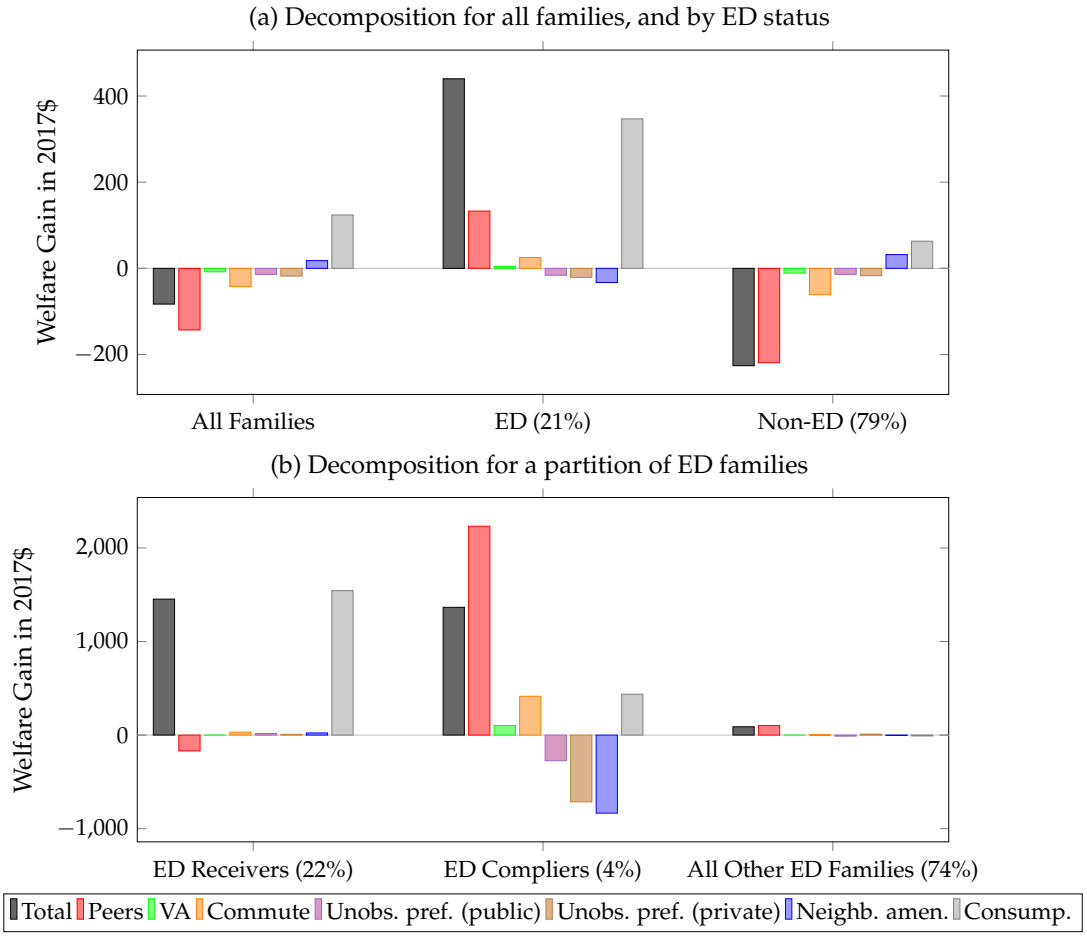
The size of the housing voucher is calibrated to ensure that every eligible family can afford housing of a typical size for its income bracket—defined as the average house size in the baseline equilibrium—while spending no more than 25 percent of income on housing. By scaling with income, the voucher attenuates the adverse income selection (conditional on ED status) that would emerge under a flat transfer. Since voucher amounts are larger in more expensive neighborhoods, the policy disproportionately facilitates access to those areas.

[19] The set of receiving neighborhoods is the same in all three versions of the policy. We rank the remaining neighborhoods by their share of ED families and choose the cutoffs so that each group of sending neighborhoods contains one third of the families living outside receiving neighborhoods.

[20] 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 west side of the city. However, the increased home-school distance would result in a much lower take-up rate and still negligible overall welfare gains for ED families.

[21] Some ED families choose to live in receiving neighborhoods even absent the voucher. Our policy designates also those families as voucher recipients on implementability grounds, as it is typically hard for a policymaker to identify the set of inframarginal families.

Figure 5: Decomposition of Welfare Effects for the Voucher Policy



This figure is the exact analogue of Figure 3 for the housing voucher policy; see the notes to that figure for the definition of each component. The only difference is the definition of *Compliers* (center group, bottom chart): ED families who do not live in receiving neighborhoods at baseline but do under the policy.

Results. We present our results as in the school choice expansion counterfactual and focus the discussion on comparing the two policies. Average welfare effects and their decomposition are shown in Figure 5. Figure 5a presents the decomposition for all families, as well as separately for ED and non-ED families. Figure 5b shows the decomposition of welfare changes for a partition of ED families: those families who live in receiving neighborhoods at baseline (*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 generates an average welfare loss of $83 (leftmost black bar in Figure 5a).

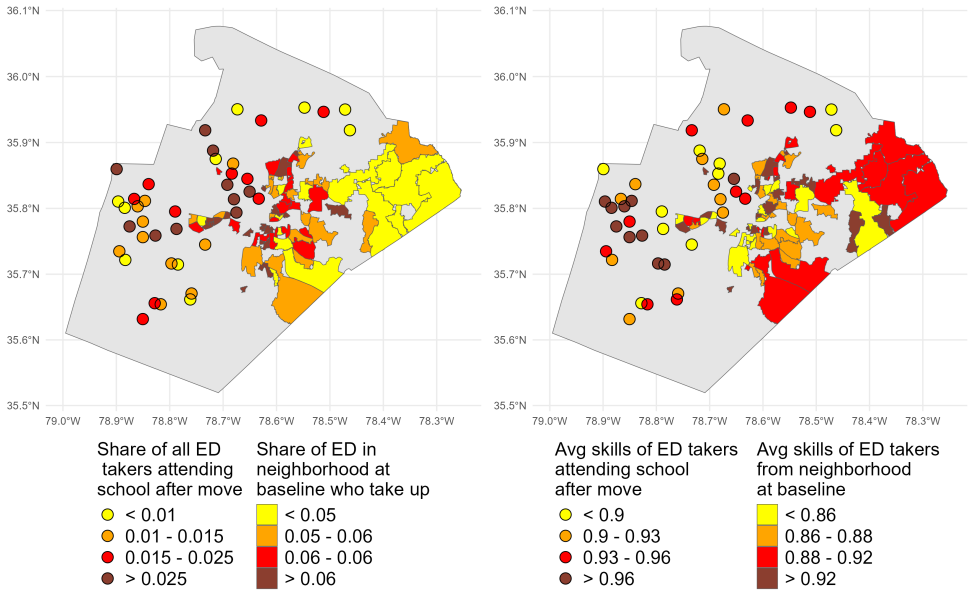
Under the policy, ED families account for 12 percent of the family population in receiving neighborhoods, up from 10 percent at baseline. As the share of ED children increases, their average skill level declines by about one percent relative to baseline, reflecting the lower skill of marginal ED entrants. A similar decline in average skill characterizes non-

Table 5: Outcomes of the Voucher Policy for Receiving Families

	ED Receivers	Non-ED Receivers
Change in average school peers (% std)	−8.4	−7.5
Change in share attending private school (pp)	0.1	1.0
Share locating outside receiving nbhds (%)	0.3	2.2

This table is the exact analogue of Table 4 for the housing voucher policy; see the notes to that table.

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



The figure shows the spatial distribution of voucher take-up across sending neighborhoods and receiving schools. Voucher-takers are ED families who live in a sending neighborhood at baseline and move to a receiving neighborhood under the policy. In the left map, each sending neighborhood is shaded by the share of its baseline ED families who take up the voucher, and each receiving school (dot) is shaded by the share of all voucher-takers whose child attends it. In the right map, sending neighborhoods and receiving schools are shaded by the average skill of the takers' children, among those who leave the neighborhood or attend the school, respectively.

ED families residing in receiving neighborhoods after the policy.

Overall, ED families gain much more, and non-ED families lose less, under the voucher policy than under the school choice expansion. Focusing on ED families, the larger gains for compliers are primarily driven by school peers and, in smaller part, by the direct effect of the voucher on consumption. Part of this gain is offset by the relocation to neighborhoods whose amenities are less well aligned with their preferences. The absence of significant losses from commuting reflects the fact that compliers now reside in receiving neighborhoods, close to their typical school, rather than commuting from distant sending neighborhoods as under the school choice expansion policy.

ED receivers experience substantial welfare gains under the voucher policy, in contrast to the losses they incur under school choice expansion. Under both policies, they are exposed to lower-skilled peers, but the voucher more than offsets this loss. Only 0.3 percent choose to live outside receiving neighborhoods, compared with 13.4 percent under the school choice expansion (see Table 5 versus Table 4), since relocating entails forfeiting the voucher.

The gap in peer-quality losses between ED and non-ED receivers (-8.4 vs. -7.5 percent) is substantially smaller under the voucher policy than under the school choice expansion (-40.0 vs. -13.3 percent). To shed light on this result, Figure 6 illustrates the average share and average skill of ED voucher-takers from sending neighborhoods attending each receiving school under the voucher policy (left and right panels, respectively). A comparison with Figure 4 shows that the magnitude and quality of the ED influx are more evenly distributed across receiving schools under the voucher policy: commuting costs no longer operate as a selection device for few, high-skilled, compliers into schools in the affluent west side.

Under the voucher policy, housing regulations—rather than commuting costs—become the key determinant of the differential exposure between ED and non-ED receivers. Voucher recipients tend to relocate to areas with lower minimum lot-size requirements: in receiving neighborhoods with below-median zoning regulation, the share of ED families increases by 3.0 percentage points on average, more than twice the average increase in neighborhoods with above-median regulation (1.2 percentage points). Figure D-3 in Appendix D.3 shows the full distribution of ED inflows. Residing in neighborhoods with stringent zoning requires recipients to supplement the voucher with their own resources, partially shielding these neighborhoods from the effects of the policy.

5.3 Robustness to Model Assumptions

We conclude our policy analysis by assessing the robustness of our results to extensions and alternative modeling assumptions. For each specification, Table 6 reports our measure of welfare change for all families and separately by ED status. It also reports the average cost of each policy per family in the county, whether due to additional transportation expenditures (see Appendix D.2.3) or the voucher subsidy. The complete counterparts of Tables 4 and 5 and Figures 3 and 5 are provided in Appendix D.5, together with additional estimation results where applicable.

Endogenous Private School Prices. Throughout the paper, we assume that private-school prices remain fixed at their baseline levels. In doing so, we abstract from supply-side

responses in the private-school sector, which we treat as an outside option to the public system and which accounts for a modest 11 percent of baseline enrollment.

We assess the robustness of our findings to allowing private-school prices to adjust in response to the policies, reflecting potential changes in marginal costs (as enrollment rises) or markups set by school owners. To keep things simple, we impose that in the new equilibrium one of two conditions must hold for each private school: either enrollment remains at its pre-policy level, or the price varies by up to 25 percent relative to baseline, increasing if enrollment is higher (decreasing if lower).[22]

Results. Welfare changes are similar to, but uniformly more negative than, those in the baseline counterfactuals. Private-school prices rise by about 16 percent on average, and multiple schools reach the upper bound of the price change. Higher prices dampen the shift toward private schools, particularly among ED receivers, whose attendance rises by about 0.5 to 1 percentage points less than in the baseline scenario.

Migration In and Out of the County. Throughout the analysis, we treat the distributions of family and other household types as exogenous and directly observed in the data for each year. While this approach accommodates changes in county composition during the estimation period, it does not allow households' county of residence to respond to policies. In our baseline counterfactuals, the demographic composition of Wake County is therefore held fixed at its 2015-2016 level.

Because the policies we study affect not only the relative attractiveness of neighborhoods and schools but also the overall value of residing in the county, in this model extension we allow for endogenous migration responses. Specifically, households first choose a county based on the expected value associated with residing there, and then choose a neighborhood—and, for families, a school. The first choice set includes all 100 counties in North Carolina. We estimate the value of residing in alternative counties from observed choice shares, and the dispersion of idiosyncratic county preferences using estimates from the migration literature. The formal model extension and measurement details are provided in Appendix D.4.

Results. Allowing for migration responses leads to modest changes in aggregate population in our counterfactuals: the total number of families in Wake County declines by 1.4 percent under the school choice expansion and increases by 0.4 percent under the voucher policy. In both cases, the composition shifts toward ED families, whose mass increases by

[22] We abstract from private-school profits and endogenous market entry, as well as from how these outcomes might adjust in response to the policies we consider. A full model of private-school supply, while certainly of interest, is beyond the scope of this paper. We also note that the mean and standard deviation of observed year-to-year private-school price changes in our sample are 0.7 percent and 4.8 percent, respectively—well below the pricing responses we allow for in our robustness exercise.

Table 6: Welfare Change from Counterfactual Policies under Alternative Assumptions

	(1) Baseline	(2) Endogenous P_n	(3) Migration	(4) Nbhd. Peers	(5) School FE
School Choice Expansion					
All Families	−1,030	−1,193	−1,069	−865	−905
ED Families	59	−96	51	54	89
Non-ED Families	−1,326	−1,491	−1,373	−1,115	−1,176
Cost per Family	128	129	132	128	137
Housing Vouchers					
All Families	−83	−305	−192	−70	−73
ED Families	440	281	406	437	443
Non-ED Families	−226	−465	−355	−208	−213
Cost per Family	130	132	134	131	130

The table shows the welfare gains (in 2017$) by family type generated by our two counterfactual policies (school choice expansion in the top panel; housing vouchers in the bottom panel), under alternative model assumptions. For convenience, Column (1) reproduces the welfare gains obtained under our baseline assumptions (i.e., the levels of the black bars in Figures 3a and 5a). Columns (2)–(5) show welfare gains under each of the alternative assumptions considered in Section 5.3. The last row of each panel reports the average fiscal cost of the policy per family in the county (in 2017$), equal to the total cost divided by the number of families: additional transportation expenditure (top panel) and voucher subsidy (bottom panel).

0.5 percent under expanded school choice and 3.3 percent under vouchers. At the same time, between 0.4 and 1.9 percent of non-ED families leave the county. The population of households without elementary-school children remains virtually unchanged. Welfare changes for incumbent families are consistent with the baseline counterfactuals, though they become more negative for non-ED families under the voucher policy, where the inflow of ED families is largest.[23]

Preferences over Neighborhood Income Composition. Our baseline specification abstracts from preferences for neighborhood income composition, or neighborhood "peers". Parental income is positively correlated with children's outcomes, leading to well-known challenges in disentangling preferences over those attributes. To gauge the robustness of our results to the inclusion of preferences for (endogenous) neighborhood peers, in this extension we assume that family i has utility over neighborhood characteristics given by

$$\alpha_{0n} + \ln(w)\left(\alpha_{1n} + \eta \ln(\bar{w}_n)\right) + \sigma_N \varepsilon_{n;i}\,, \tag{5.2}$$

[23]When migration is allowed, a total welfare analysis must account not only for changes in utility among residents who remain in the county, but also for the gains or losses associated with households relocating into or out of the county in response to the policy. In our counterfactuals, this component is negligible.

which augments Equation (2.6) with preferences over average neighborhood income, $\bar{w}_n$.[24] We fix a value for η and re-estimate the rest of the model targeting the same moments as in our baseline specification, and then perform our counterfactual analysis in this expanded model.

For η, we choose a value that implies an average willingness to pay of \$1,200 per year for a uniform 10 percent increase in the average income of all neighborhoods. While the comparison is necessarily imperfect, the estimates in Bayer et al. (2007) imply a willingness to pay for a similar variation of about \$500 per year (also in 2017 dollars), and similarly increasing with respect to the household's own income. We see our choice as a conservative upper bound. Other plausible values do not alter our conclusions.

Results. Both our results and the underlying mechanisms are essentially unchanged. With preferences over neighborhood composition, the estimated valuations of school attributes are only slightly smaller. Losses are therefore more muted under the school choice expansion, the counterfactual with the larger take-up and hence the larger deterioration in the peer composition of receiving schools.

Unobserved School Heterogeneity. Last, we assess the sensitivity of our results to the presence of systematic unobserved school attributes, potentially valued differently by ED status. We augment the utility from public school s in Equation (2.9) with a fixed effect $\lambda_g \psi_s$, where $g \in \{\text{ED}, \text{non-ED}\}$ and λ_{ED} is normalized to 1. We restrict ψ_s to be common within 8 clusters of geographically proximate schools. We re-estimate the model, recovering the cluster effects from observed school attendance shares, and the loading $\lambda_{\text{non-ED}}$ from the differential attendance of ED versus non-ED children at high-ψ schools. As argued in Section 2, attendance shares are only partially informative about such attributes in our setting, because observed attendance reflects not only preferences but also capacity constraints and the assignment of lottery losers to their base schools. The estimated school effects are somewhat negatively correlated with observed school quality ($\text{corr}(\psi_s, \ln \bar{a}_s) = -0.10$), are valued more by non-ED families ($\lambda_{\text{non-ED}} = 2.36$), and slightly raise the estimated preferences for school peer composition relative to the baseline. That is, once the valuation of peers is pinned down by the house-price response, attendance patterns reveal no residual sorting toward unobservably better schools that also have better peers.

Results. The last column of Table 6 reports the welfare effects. Both the aggregate losses and their distribution between ED and non-ED families are quantitatively very close to the baseline.

[24] We similarly augment the utility specification for the rest of households, using the same value of η.

6 Conclusion

We build and estimate a spatial equilibrium model of neighborhood and school choice that incorporates the key institutional determinants of educational access in the United States: school catchment areas, transportation provision, and zoning regulations. The model is identified using observed patterns of residential sorting and school choice, both in the cross-section and through quasi-experimental evidence on how demand responds to changes in catchment-area composition.

The model sheds light on the trade-offs involved in designing large-scale interventions aimed at reducing inequality in educational access. School-choice expansions and housing vouchers are moderately effective at improving outcomes for their intended beneficiaries, but they impose losses on families that have already invested in their children's education through their residential location. Among these families, the low-income ones who pay more for housing to access a better peer composition are particularly harmed by school choice expansions, as their geographic proximity to compliers disproportionately exposes them to additional inflows into their schools. Relocation or private schools offer only imperfect and costly insurance to families with a high valuation of school quality seeking to mitigate these adverse effects. In the aggregate, both policy counterfactuals generate welfare losses.

Our analysis can be extended along several dimensions. First, we abstract from dynamic considerations in families' choices of schools and neighborhoods, which may matter when school portfolios are expected to evolve over time (Bayer et al., 2016). Second, the model could incorporate additional parental responses to integration policies—such as seeking special classroom assignments (Crema, 2023) or adopting parenting strategies that regulate children's social interactions (Agostinelli et al., 2026)—which would shape children's effective exposure to policy-induced changes in peers. Third, a natural extension would focus on how changes in school composition affect skill formation and, ultimately, labor market outcomes. Finally, the design of optimal local institutions—such as school boundaries and zoning regulations—represents a promising, albeit challenging, endeavor that builds on our counterfactual analysis. We view these extensions 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," *Journal of Political Economy*, 2026, *134* (1), 313–365.

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," 2021. Unpublished Manuscript, Stanford University.

Almagro, Milena and Tomás Domínguez-Iino, "Location Sorting and Endogenous Amenities: Evidence From Amsterdam," *Econometrica*, 2025, *93* (3), 1031–1071.

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

Baum-Snow, Nathaniel and Lu Han, "The Microgeography of Housing Supply," *Journal of Political Economy*, 2024, *132* (6), 1897–1946.

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.

_ **, Robert McMillan, Alvin Murphy, and Christopher Timmins**, "A Dynamic Model of Demand for Houses and Neighborhoods," *Econometrica*, 2016, *84* (3), 893–942.

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*, May 2024, *114* (5), 1281–1337.

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," *American Economic Review*, December 2025, *115* (12), 4476–4522.

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.

Dávila, Eduardo and Andreas Schaab, "Welfare Assessments with Heterogeneous Individuals," *Journal of Political Economy*, 2025, *133* (9), 2918–2961.

Davis, Morris A and François Ortalo-Magné, "Household expenditures, wages, rents," *Review of Economic Dynamics*, 2011, *14* (2), 248–261.

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.

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 G. Hammond, Matthew A. Lenard, Melinda Morrill, Thayer Morrill, and Colleen Paeplow, "The attraction of magnet schools: Evidence from embedded lotteries in school assignment," *Economics of Education Review*, 2025, *107*, 102663.

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

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

Eckert, Fabian and Tatjana Kleineberg, "The Geography of Opportunity: Education, Work, and Intergenerational Mobility Across US Counties," 2024. Unpublished Manuscript, University of California San Diego.

Epple, Dennis, Akshaya Jha, and Holger Sieg, "The superintendent's dilemma: Managing school district capacity as parents vote with their feet," *Quantitative Economics*, 2018, *9* (1), 483–520.

_ **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, Veronica Guerrieri, Mark Ponder, and Marta Prato, "The End of the American Dream? Inequality and Segregation in US Cities," *Journal of Political Economy*, Forthcoming.

Glaeser, Edward L and Joseph Gyourko, "Urban decline and durable housing," *Journal of political economy*, 2005, *113* (2), 345–375.

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," 2024. Unpublished Manuscript, Stanford University.

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.

Kennan, John and James R Walker, "The effect of expected income on individual migration decisions," *Econometrica*, 2011, *79* (1), 211–251.

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

Laverde, Mariana, "Distance to Schools and Equal Access in School Choice Systems," 2024. Unpublished Manuscript, Boston College.

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

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, migration, and local employment elasticities," *American Economic Review*, 2018, *108* (12), 3855–3890.

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

Notowidigdo, Matthew J, "The incidence of local labor demand shocks," *Journal of Labor Economics*, 2020, *38* (3), 687–725.

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. Unpublished Manuscript, University of Michigan.

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

Sanbonmatsu, Lisa, Jens Ludwig, Lawrence F. Katz, Lisa A. Gennetian, Greg J. Duncan, Ronald C. Kessler, Emma Adam, Thomas W. McDade, and Stacy Tessler Lindau, *Moving to Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation*,

Washington, DC: US Department of Housing and Urban Development, Office of Policy Development and Research., 2011.

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 Bogotá's TransMilenio," *American Economic Review*, February 2026, *116* (2), 418–63.

A Data appendix

A.1 Additional Institutional Details

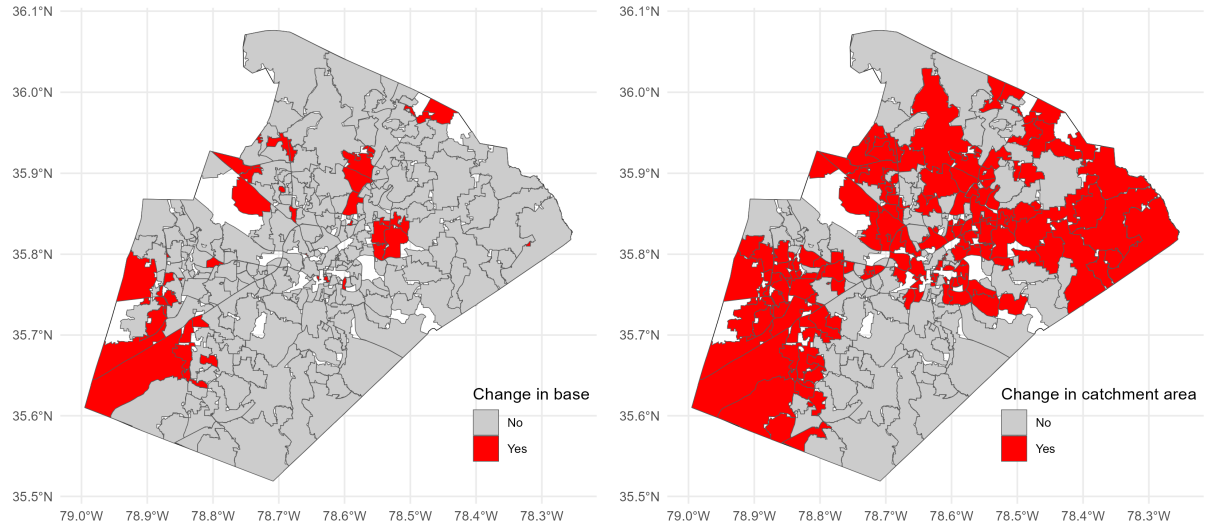
The Wake County Public School System (WCPSS) is the county-wide school district covering Wake County, North Carolina, which is the county of the state capital, Raleigh. The WCPSS was, in 2019–20, the fourteenth largest school district in the United States, with more than 161,000 students. The geography 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 1a.

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 were 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 has been the main driver of school reassignments (Parcel and Taylor, 2015, p. 53). Figure A-1 illustrates changes in catchment areas over the sample period. The map on the left shows in red all neighborhoods that 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

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



The map on the left highlights in red the neighborhoods reassigned to a different base school between 2013–14 and 2016–17. A reassignment modifies the catchment areas of both the initial and new schools: the neighborhood is removed from the former and added to the latter. As a result, neighborhoods that remain assigned to either school experience a change in their catchment area, even if their base school does not change. Accounting for this, the right map marks in red all neighborhoods that experienced either a base-school reassignment or a change in their catchment area during the sample period. Unshaded areas are out of our sample and essentially non-residential.

the neighborhood is removed from the former and added to the latter. This means that the neighborhoods that remained assigned to the initial base school and those that were already assigned to the final base school experienced a change in 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 child. Also based on their residential address, parents may or may not be offered school transportation to the magnet school.[25] Magnet choice sets and transportation provision do not only change cross-sectionally but also over time: during the period of interest, all neighborhoods experienced *eligibility* changes, in the sense that new options were added to all choice sets over the sample period.

[25] 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.[26] 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 with 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 rank the school lower on their appli-

[26] 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.

cation list.[27] Because priority is given to first-ranked schools and most schools receive more first-choice applicants than seats (Dur et al., 2018), choices ranked below the first one were typically irrelevant in practice. In that regard, our single-application assumption is a reasonable approximation of the Boston Mechanism. In the final two years of our sample, the district switched to Deferred Acceptance (DA). However, since families can apply to only five out of an average of 18 options, strategic considerations remain important, consistent with our model.

A.1.2 On the Exogeneity of School Catchment Area 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 discuss institutional background that sheds light on the drivers of these changes; we review existing evidence of the conditional exogeneity of these changes; and we provide additional evidence that these institutional changes were not anticipated by households.

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."[28] In addition, while families knew that changes in catchment areas were likely over the period of interest, and while the population expanded more in some parts of the county than in others, it was still extremely hard for families to predict which catchment area boundaries would be redrawn and how. As Parcel and Taylor (2015) note, "[h]ousing development policy, frequently described as 'sprawling', compounded the problem of acquiring reasonably inexpensive real estate to build schools proximate to large numbers of students. With housing development the responsibility of municipalities and schools governed at the county level, residences were constructed with little regard to the geographic placement of existing public schools ... Some campuses, moreover, were better able to expand physically than others. This accelerated and uneven growth forced the school board to undertake yearly efforts to reassign children to different, but not always brand new, schools." Parcel and Taylor (2015) document, ac-

[27] 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."

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

cordingly, that families generally expected that some boundaries would change in any given year, but could not predict which neighborhoods would be affected.

In addition, 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 the same school in one year, would be assigned to different schools the following year." They formally assess the conditional exogeneity of catchment area changes by comparing the demographic characteristics of reassigned and non-reassigned students, using data on seventh- and eighth-graders from school years 2005–06 to 2011–12. Specifically, they regress individual student characteristics (race, gender, prior absences, and prior test scores) on an indicator for whether the student resides in a reassigned neighborhood—distinguishing between reassignments to existing versus newly opened schools—while controlling for grade-by-school-by-year fixed effects and neighborhood-by-grade-by-school-by-year covariates (including the share of ED students, average prior-year ELA scores, measures of proximity to newly opened schools, and neighborhood student counts). Across specifications, the estimated coefficients on reassignment status are small in magnitude and statistically insignificant at conventional levels.

To further show that changes were not anticipated by households, we test whether policy-induced changes in catchment-area peer quality (defined in Equation (4.2)) correlate with pre-trends in house prices. Specifically, we estimate the same specification of the regression in Equation (4.1), but changing the left-hand-side outcome variable. We regress prior average changes in within-neighborhood house prices—from $(t-2)$ to $(t-1)$, from $(t-3)$ to $(t-1)$, and from $(t-4)$ to $(t-1)$—on the current policy-induced changes in catchment-area peer quality. Results are shown in Table A-1, and indicate that policy-induced changes in catchment-area peer quality do not predict pre-policy trends in house prices.

We also examine whether neighborhoods that subsequently experienced catchment boundary changes and those that did not differed in terms of their housing stock characteristics before the redrawing occurred. Specifically, we estimate regressions similar to Equation (4.1), but using the following as the left-hand-side variable: average house size, year built, and lot size for properties in the neighborhood sold at $(t-1)$ and $(t-2)$. Results are shown in Table A-2 and provide no evidence that our measure of policy-induced changes in catchment-area peer quality is correlated with neighborhood characteristics.

This suggests that neighborhoods experiencing different changes in school peer quality do not systematically differ from other neighborhoods along these dimensions.

Table A-1: Policy-induced Changes in School Peer Quality and Pre-trends in House Prices

	Change in Log House Price per Sq. Ft.		
	(t-2) to (t-1)	(t-3) to (t-1)	(t-4) to (t-1)
Change in School Peer Quality (Log)	-0.020	0.002	0.059
	(0.026)	(0.033)	(0.050)
Observations	801	789	779
Sale-Year F.E.	Yes	Yes	Yes

The table reports tests of whether policy-induced changes in base-school peer quality (defined in Equation (4.2)) are correlated 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.

Table A-2: Policy-induced Changes in School Peer Quality and House Characteristics

	Characteristics of Houses Sold in Previous Years					
	Housing Size (sq.ft.)		Year Built		Lot Size (acres)	
	(t-1)	(t-2)	(t-1)	(t-2)	(t-1)	(t-2)
Change in School Peer Quality (Log)	-42.026	-151.557	3.571	1.019	-0.029	-0.013
	(435.374)	(464.405)	(9.389)	(9.027)	(0.197)	(0.178)
Observations	807	803	807	803	807	803
Sale-Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes

The table reports tests of whether policy-induced changes in base-school peer quality (defined in Equation (4.2)) are correlated with the characteristics of houses sold in prior years. We define policy-induced changes as reassignments of school catchment areas implemented by the WCPSS between the 2013–14 and 2016–17 school years (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.

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

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

A.2.1 Students

Data source. Student-level data were obtained from the North Carolina Education Research Data Center (NCERDC).[29] The data show, for each year and each student enrolled

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

in a North Carolina public school, the school the child is enrolled in and a set of demographic variables (gender, economically disadvantaged (*ED*) status). 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, kindergarten literacy Composite score (see below). This yields 36,690 students.
2. 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 their 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 six years prior; and (3) children attending a school that was never part of the choice set attached to their address. We assume the 638 students of type (2) were grandfathered into attending the same school as an older sibling. They remain in the sample as residents of their neighborhood (594 of them are in the final sample, 1.7 percent), but we exclude them from all school-side variables (option capacities, school peer quality, school shares), since their school was not chosen from their neighborhood's current choice set. 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 half a 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 their choice of school cannot be explained given the choice set —we drop 1,311 students through this step.

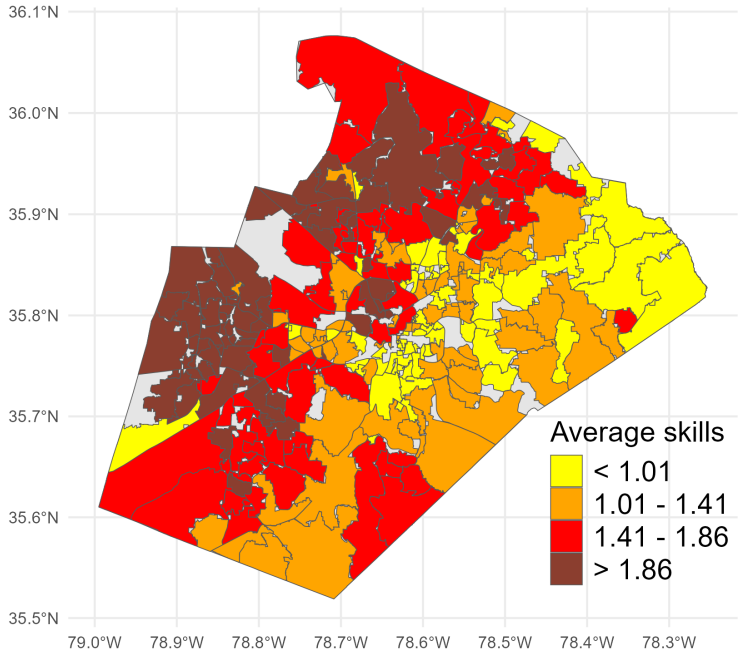
3. Finally, we count the number of students in each neighborhood, and exclude neighborhoods with fewer than ten students over the four years of data, as well as neighborhoods with no student in at least one of the four years. Our final sample consists of 310 neighborhoods and 34,785 students.

Construction of the skills measure. We use beginning-of-year kindergarten literacy assessment test scores as a measure of a student's baseline skills. We use DIBELS (Dynamic Indicators of Basic Early Literacy Skills) assessment scores from the mCLASS assessment system (https://amplify.com/northcarolina/) available in the NCERDC data, in particular the Composite score. We standardize the Composite score (mean zero and unit standard deviation over the Wake County public-school kindergartners of 2013–14 to 2016–17 with a known address) and use it as the baseline log-skill measure. The resulting baseline log-skill measure is divided by its cohort-specific standard deviation (unit standard deviation in each cohort; it is not demeaned by cohort). We then discretize the log-skill distribution. Specifically, we define a common 10-point support across all years by computing the median log-skill within each decile of the pooled distribution. While the support points are fixed over time, the empirical mass associated with each point is allowed to vary by year. After discretizing the log-skill distribution, we exponentiate the support points to obtain skill types in levels, $a \in \{a_1, \ldots, a_{10}\}$. Figure A-2 reports average skill levels by neighborhood.

A.2.2 Schools

Public schools: location, peer quality, capacity. Figure 1a 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 skill level a (the discretized skill types in levels defined above) for kindergarten students enrolled in school s in year t. It is computed over the kindergarten students of our final sample, excluding grandfathered students. For each school s, the number of seats reserved for option (as opposed to base) students in year t ($q_{s,t}$) is measured as follows. If s is oversubscribed in year t, then $q_{s,t}$ is set to the count of kindergarten students attending s as an option in that year. If s is not oversubscribed in year t but is in some other year in the sample, then $q_{s,t}$ is set to the maximum between its option enrollment in year t and the average number of kindergarten students attending s as an option in the years it is oversubscribed. Finally, if s is never oversubscribed over the sample period, $q_{s,t}$ is set to the option enrollment in s in year t. Option enrollment counts exclude grandfathered students, and $q_{s,t}$ is expressed as a share of the year's (non-grandfathered) kindergartners

Figure A-2: Geography of Children's Kindergarten-entry Skills



The map shows the average children skills by neighborhood over the four sample years. The average skill level thresholds used to build the color legend are the quartiles of the distribution of average skills across neighborhoods in the sample.

in our final sample. While we do not observe individual applications, school-level information provided by the WCPSS indicates whether option schools are oversubscribed (kindergarten admission rate below one). Ninety-four of the 111 public schools in sample are an option school for at least one year. Of these 94 option schools, 78 are oversubscribed at least once over the sample period. The rules above follow from the fact that observed option enrollment equals capacity only when the program is oversubscribed; when all applicants are admitted, it is a lower bound. Schools that do not offer an option program in year t have zero capacity.[30] In the model, the mass of public-school families is normalized to one, so $q_{s,t}$ is the number of option seats per public-school kindergartner, and the admission probability to program s is $\min\{1, q_{s,t}/\pi_{s,t}\}$, where $\pi_{s,t}$ is the model's mass of applicants to s. Over 2013–2016, there are 329 option program-years. Of these, 211 are oversubscribed (median admission rate 0.58). Of the 118 in which all applicants are admitted, 32 use the average over oversubscribed years, 82 use observed option enroll-

[30] Four program-years with an admission rate of one but no option kindergartner in our sample (all in programs that are never oversubscribed) are assigned one seat, rather than their observed option enrollment of zero, so that the program remains open to applicants. We also correct two admission rates by hand: we fill in one missing rate from the district's records of applications and admissions, and we drop one rate recorded for a school that was not offered as an option that year.

ment (32 of them because the program is never oversubscribed), and 4 are assigned one seat. Total option capacity amounts to 25–28 percent of kindergartners each year (2,080 to 2,386 seats).

Public schools: school value added. We measure school value added as the contribution of schools to their students' third-grade test scores, conditioning on students' (log) initial skills as well as the (log) average skills of their peers. The econometric model is as follows:

$$Y_{i,s,t} = b_0^{VA} + b_1^{VA} \ln a_{i,s,t} + b_2^{VA} \ln \bar{a}_{s,t} + A_s + \delta_t + \epsilon_{i,s,t}^{VA},$$

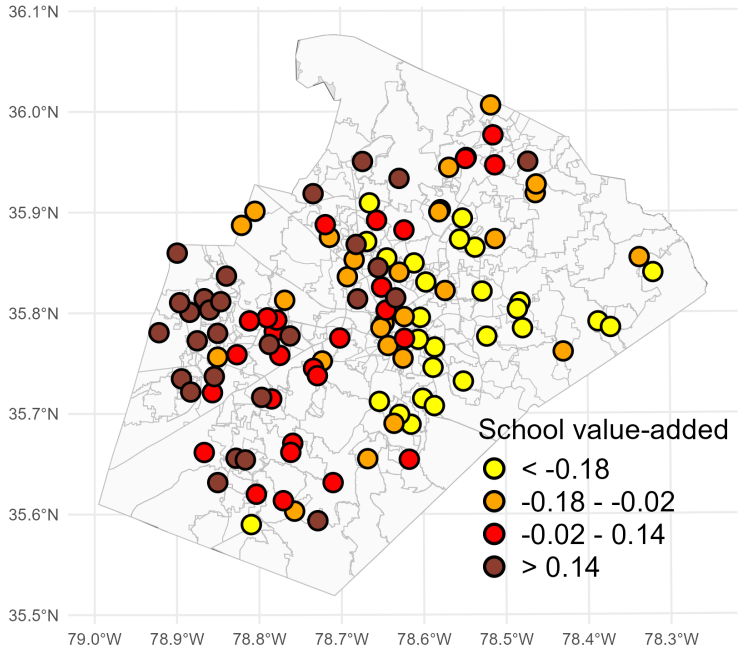
which represents the determinants of a child's third-grade skill measure for a child who attended kindergarten in school s during school year t, with skill level $a_{i,s,t}$ and peers with average skill $\bar{a}_{s,t}$. The value-added term A_s represents the school's contribution to child i's test-score performance, conditional on the child's initial skill level at kindergarten entry and the average skill level of peers at kindergarten. The term δ_t is a kindergarten-cohort fixed effect.

Endline skills. The outcome $Y_{i,s,t}$ is built from the child's end-of-third-grade standardized test scores in mathematics and reading. These are available for the cohorts entering kindergarten in the fall of 2013, 2014, and 2015, who take the test in the spring of 2017, 2018, and 2019 respectively; there were no tests in the spring of 2020. Each score is standardized within the year in which the test was taken, using the mean and standard deviation of the students who took it on time (three years after kindergarten). For these students, this amounts to standardizing within kindergarten cohort. Students who repeated a grade before third grade take the test a year later and are thus compared with the students who took the same test. This standardization also accounts for the new mathematics score scale introduced in the spring of 2019. The two standardized scores are combined into a single endline skill measure using Bartlett factor scores. The resulting measure is divided by its standard deviation within each kindergarten cohort, and discretized as the baseline measure: each child is assigned the median of their decile of the pooled distribution.

Initial and peer skills. Initial skills $a_{i,s,t}$ are constructed as the skill types defined above, but on the value-added sample described below, and with the baseline log-skill measure also demeaned within each cohort. Peer skills $\bar{a}_{s,t}$ are the average of $a_{i,s,t}$ over the kindergartners of school s in year t in the value-added sample.

Sample and estimation. Since value added measures the school, we use all kindergartners enrolled in the 111 public schools of our sample in 2013–14 to 2016–17 with non-missing ED status and baseline skills, including grandfathered students and students outside our

Figure A-3: Geography of School Value Added



The map shows the estimated school value added for each school in the sample. The value added thresholds used to determine the colors of the dots represent the quartiles of the distribution of school value added in the sample.

sample neighborhoods: all of a school's students are informative about its value added. The estimation sample consists of the 30,671 students with third-grade scores, in 107 schools. A_s is estimated as the school fixed effect (within-school estimator, robust standard errors). The coefficient on peer skills is identified from year-to-year changes in a school's peers, and is consistently small. The four schools that opened in 2016–17 have no students with third-grade scores; their value added is predicted from a cross-school regression of the estimated A_s on the school's average (log) peer skill.
Figure A-3 maps the value-added estimates.

Private schools: data sources, location, attendance, and prices. To get information about the set and location of private schools offering kindergarten in Wake County over the period of interest, we combine several sources of data. First, we obtained through a public record request school-and-grade-level enrollment counts for all registered private schools in Wake County from the North Carolina Department of Administration (a version of these statistics aggregated at the county level is available at https://www.doa.nc.gov/divisions/non-public-education/private-school-directory-statistics, accessed August 2025). Because the NC DOA collects these data by survey, they suffer from non-response (for instance, enrollment counts for the school year 2015–16 were not

available). We therefore combined these data with the biennial 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. In total, the consolidated dataset includes 68 private schools in Wake County that reported at least one student enrolled in kindergarten between fall 2013 and spring 2017 (after dropping two short-lived schools, open for only two and three years respectively, one of which also changed address while open). Kindergarten enrollment can be difficult to interpret, as it may occur in both elementary schools and in pre-schools or daycare centers. Whether a school also reports enrollment in first grade serves as a useful indicator of whether it operates at the elementary level. Of the 68 schools with some kindergarten enrollment during 2013–2016, 54 also enrolled first-grade students, while 14 did not.

Tuition. Private school tuition rates are not systematically recorded in existing datasets. However, since the 2014–15 school year, North Carolina has operated the Opportunity Scholarship Program, which provides vouchers that allow eligible students to attend participating private schools. As part of this program, the North Carolina State Education Assistance Authority (NCSEAA) (www.ncseaa.edu) maintains annual records of the tuition charged by private schools attended by voucher recipients. We obtained NCSEAA data for private schools in Wake County covering the 2014–15 through 2019–20 school years. These records include only schools that participated in the Opportunity Scholarship Program, so coverage is incomplete, especially in the program's early years when participation was limited. To supplement these data, we hand-collected additional tuition information from private school websites (typically available only for the current year) and from archived pages on the Wayback Machine (https://web.archive.org/). Remaining missing tuition information for the school years 2013–14 to 2016–17 was imputed from more recent years obtained from the NCSEAA using the school's average year-to-year growth rate of tuition (in 2017 dollars): a missing year is predicted backward from the next available year, divided by that average growth factor.

Final private school sample. Because our analysis focuses on entry into elementary school, we restrict the sample to schools offering elementary-grade levels beyond kindergarten —excluding daycares and institutions that operate solely at the pre-school level. We further drop schools that opened only in the last year of our sample (fall 2016), which removes one school. Next, we exclude schools for which tuition information over the sample period was not recorded and could not be predicted (13 schools). Finally, we drop schools with an average annual kindergarten enrollment of fewer than five students during the analysis period (four schools removed). Figure A-4 shows the location of our final sample of 36 private schools offering kindergarten in Wake County between 2013–14 and

2016–17, their annual tuition level, as well as the boundaries of the eight PUMAs partitioning Wake County.

Attendance by income. We use data at the PUMA level from the ACS (1-year samples for survey years 2013 to 2017, pooled) to construct the share of households with a child enrolled in kindergarten to grade four whose children attend private school, by PUMA and income level.[31] A household with one child in public school and one in private school counts as public. Households are identified by survey year and household serial number, and income is converted to 2017 dollars. In particular, we use the variables "School attendance", "Grade level attending (detailed version)", "Public or Private school", "Total household income", and "PUMA" (downloaded from https://usa.ipums.org/usa/index.shtml, accessed August 2021). From this, 11.1 percent of Wake County households with children in grades K–4 send them to private school.

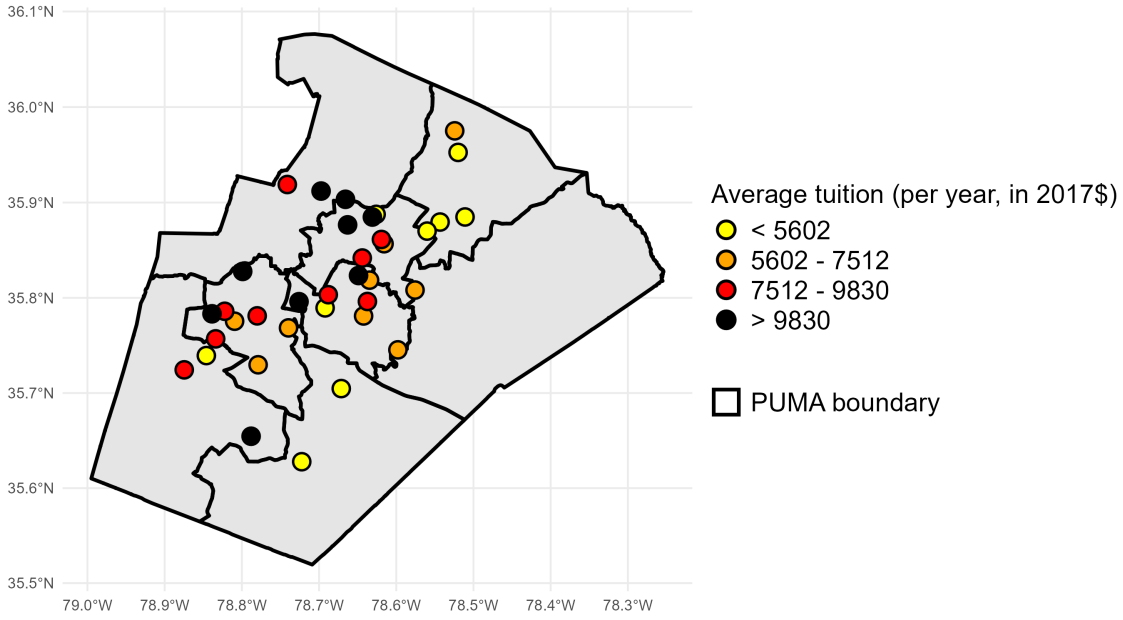
Mapping the model to the data. The private school $\mathcal{P}_{n,t}$ in the choice set of residents of neighborhood n in year t is the private school closest to n, that is, the school with the shortest road distance to the centroid of n among the private schools in our sample open in year t. We set $d_{\mathcal{P}_{n,t},n}$ to that distance. The price $P_{n,t}$ paid by families in neighborhood n who attend private school is the tuition charged by $\mathcal{P}_{n,t}$ in year t, in 2017 dollars. For each of the eight PUMAs constituting Wake County, we construct the share of households with children in kindergarten to grade four who attend private school separately for three income levels —at most \$50,000 ($w \in \{w_1, \ldots, w_4\}$, which corresponds to our ED group), between \$50,000 and \$125,000 ($w \in \{w_5, w_6, w_7\}$), and above \$125,000 ($w \in \{w_8, w_9, w_{10}\}$).

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.

[31] Private-school shares are measured over kindergarten to grade four for consistency with the tract-level private shares used in Section A.2.3. See footnote 38. In addition, Unlike the other ACS 1-year statistics (yearly county household counts, the county income distributions of families and other households, and the ratios of private- to public-school families by ED status), which use the 2014–2017 surveys so that each survey is matched to the school year starting the previous fall (the 2014 survey to 2013–14, through the 2017 survey to 2016–17), the private-school attendance targets use all five 1-year surveys from 2013 to 2017. The NCERDC data cover only public-school kindergartners, so to account families whose children attend private school we weight each public kindergartner by the private-school share of children in their census tract, taken from ACS 5-year table B14002 for 2013–2017. Tract-level shares exist only in the 5-year release. Together, the five 1-year samples 2013–2017 contain exactly the households in that 5-year file, so using all five makes the attendance targets describe the same households as the weights.

Figure A-4: Private School Locations and PUMA Boundaries



The map shows PUMA boundaries and the location and annual tuition of private schools offering kindergarten between fall 2013 and spring 2017 in Wake County. The tuition thresholds used to determine the legend are the quartiles of the tuition distribution, expressed in 2017 dollars.

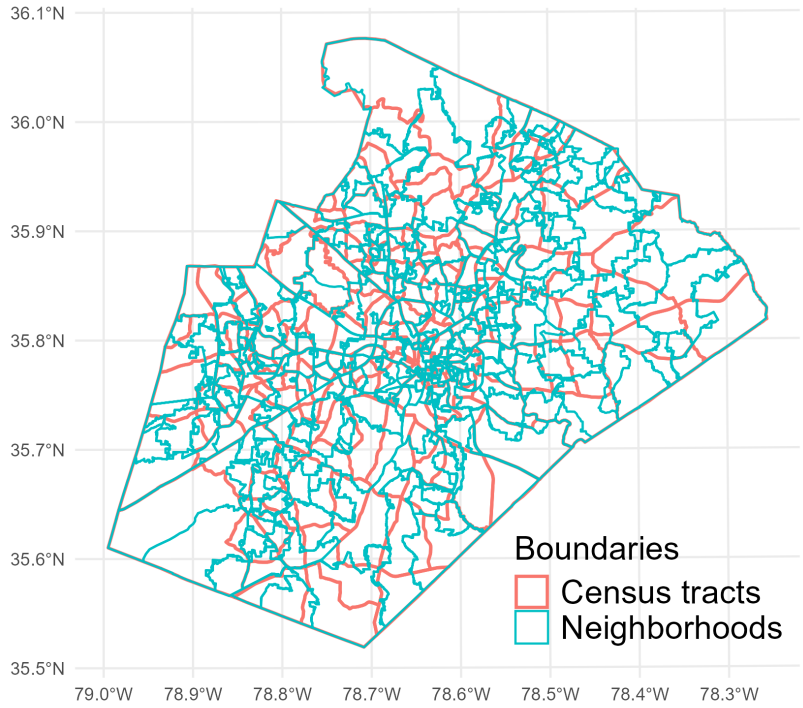
Construction of neighborhoods. Each neighborhood n is characterized by a sequence of public school portfolios from school year 2013–14 to school year 2016–17,

$$\{ (\mathcal{B}_{n,t}, \mathcal{T}_{n,t}, \mathcal{NT}_{n,t}) \mid t = 2013, \ldots, 2016 \},$$

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 set of option schools in the choice set of neighborhood n that do not provide transportation. Neighborhood n is the union of all *contiguous* points with this sequence of portfolios. Formally, let us denote each residential address by its coordinates (x, y). Then, $(x, y) \in n$ if and only if the following points are satisfied for each t:

1. (x, y) has base school $\mathcal{B}_{n,t}$ in school year 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.
4. 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.

Figure A-5: Neighborhood Boundaries and 2010 Census Tracts



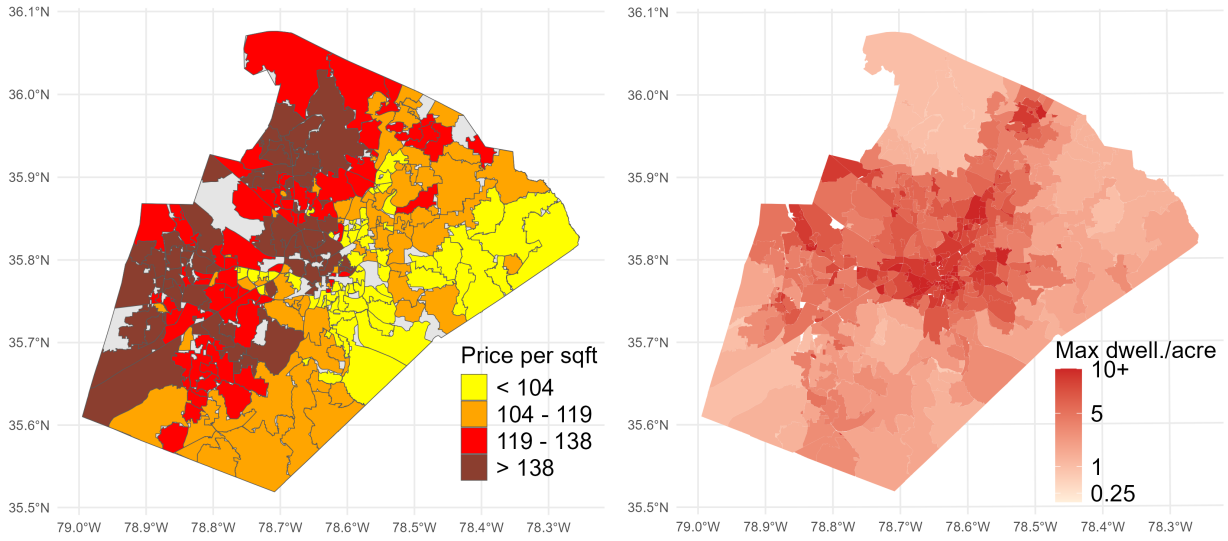
The map shows the boundaries of the constructed neighborhoods (teal), as well as the boundaries of the 2010 Census tracts (red).

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-5 shows the obtained partition of Wake County into neighborhoods, with 2010 Census tract boundaries for comparison.

Distance between schools and neighborhoods $d_{s,n}$. 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.[32]

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

Figure A-6: House Prices (Left) and Residential Density Regulations (Right)



(a) House prices

(b) Zoning

The map on the left shows average house prices per square foot by neighborhood (expressed in 2017 dollars, source: Wake County Real Estate Transaction data, sale years 2014–2017). The dollar thresholds used to determine the colors of the neighborhoods are 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.

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—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 are 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. Residential zones that cap density but set no minimum lot size enter with an MLS equal to the inverse of the density cap (acres per dwelling unit). Our zoning map is built from the 2020 GIS layers, the earliest year with consistent records across zoning authorities. The exception is planned developments: their rules are set project by project rather than by the zone code, and we code them from the approved plan in force in 2013–2016. Non-residential zones are excluded. Figure A-6b maps the corresponding density caps (dwelling units (du) per acre, the inverse of MLS) by neighborhood in Wake County.

Lighter areas in Figure A-6b 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 acres per lot. There is a relatively wide range of MLS regulations throughout Wake County—from more than 10 du/acre in the urban center of the county, to less than 1 du/acre in parts of the periphery.

To each neighborhood, we attach an MLS. For neighborhoods that overlap multiple zoning areas with distinct MLS restrictions, the neighborhood-level MLS restriction is constructed as the average MLS over the neighborhood's residentially zoned area, weighted by the area of each zone. Formally, we set

$$\underline{\text{mls}}_n = \frac{\int_{(x,y)\in n} \underline{\text{mls}}(x,y)\, d(x,y)}{\int_{(x,y)\in n} d(x,y)},$$

integrating over residential land only, 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, however, households choose, and are constrained in, *house* size rather than *lot* size, so we map neighborhood restrictions on minimum lot size ($\underline{\text{mls}}_n$) into a minimum available house size ($\underline{h}_n$). We first group neighborhoods by their value of $\underline{\text{mls}}_n$ and, for each group g, compute h_g^{min}, the fifth percentile of the heated area of the houses sold in the neighborhoods of that group (for this, we restrict the sample to houses built in 1980 or later and sold between 1990 and 2019). We then regress h_g^{min} on the group's MLS, $\underline{\text{mls}}_g$ (measured in acres), with a linear quantile regression at the first percentile, weighting each group by its number of sales. This yields the mapping $\underline{h}_n = 408 + 637 \times \underline{\text{mls}}_n$, which traces the lower envelope of observed house sizes across regulation levels. This implies an essential minimum housing of 408 square feet in an unregulated neighborhood ($\underline{\text{mls}}_n = 0$). The smallest MLS across sample neighborhoods is $\underline{\text{mls}}_n \approx 0.091$ acres, which implies a minimum house size in the neighborhood sample of $\underline{h}_n \approx 465.68$ square feet. Rounded down, it is the lower bound on house size used to trim house sales below.

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.[33] 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

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

was built, whether the building is for residential use, and its heated area. We use heated area as our measure of house size.

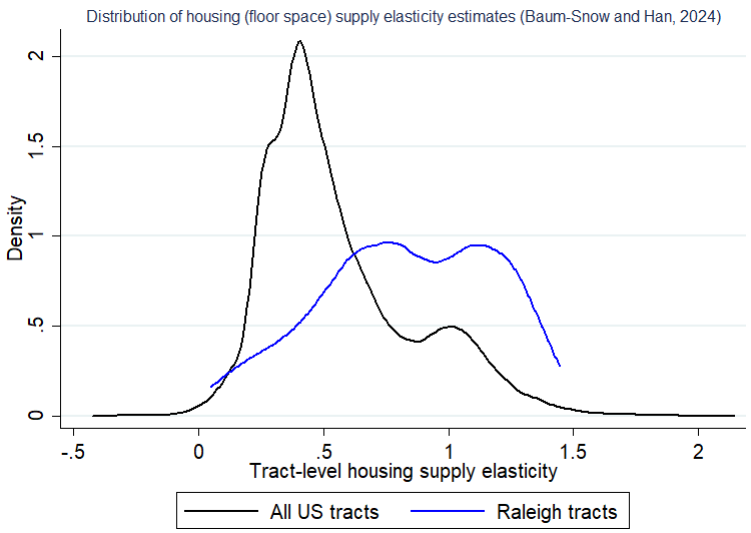
To construct a measure of average house price by neighborhood and year, we proceed in three steps. First, we convert all prices into 2017 dollars to be consistent with household income provided in 2017 dollars in the ACS.[34] Second, for each neighborhood and year, we compute the average of the log sale price per square foot and take its exponential. We trim outliers by excluding properties with a heated area smaller than 465 square feet (the smallest minimum house size across neighborhoods, rounded down) or larger than 5,000 square feet, with a sale price exceeding \$1,000,000 or below \$40,000 (in 2017 dollars), or with a price per square foot below \$25 or above \$500. We exclude sales the county flags as *disqualified*, such as transfers between relatives, sales involving government or lending institutions, forced sales, and non-warranty deeds. Third, 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 the sale years 2011 to 2018) 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 dollars per square foot) are depicted in Figure A-6a.

As our measure of the neighborhood-level cost of housing services ($r_{n,t}$), we use the annuitized average price per square foot in neighborhood n in year t. We compute the annuitized price by dividing the average price per square foot by 15. We choose this factor for two reasons. First, it implicitly corresponds to an annual user cost of approximately six to seven percent of the house value, consistent with standard estimates of depreciation, maintenance, property taxes, and the opportunity cost of capital. Second, it generates an average housing size at the county level that is consistent with the value in the data.

Housing supply elasticity. Baum-Snow and Han (2024) provide housing supply estimates for housing units, new construction housing units, land development, floor space, and floor space in new construction units. We focus on floor-space estimates. While they provide estimates from several empirical specifications, we follow the recommendation they give in the 'Appropriate usage' appendix, and use estimates from the quadratic specification based on the finite-mixture model estimator. We use the estimates using the 2011 fraction of tract developed. Finally, Baum-Snow and Han (2024) provide two levels of estimates: tract-level, and metropolitan region-level. To use in our empirical analysis, we map the neighborhoods in our data to 2000 Census tracts (the tract definition used by

[34] We use end-of-period annual CPI for all urban consumers https://fred.stlouisfed.org/series/CPIAUCSL, accessed August 2021.

Figure A-7: Housing Supply Elasticity in Raleigh and the Rest of the US



The figure shows the distribution of Baum-Snow and Han (2024)'s estimates of tract-level housing supply elasticities in the Raleigh metro area, and in the U.S. overall.

Baum-Snow and Han, 2024) and use the tract-level estimates to get neighborhood-level housing supply elasticities. To provide some context, we document how the elasticity of housing supply in Raleigh compares to that of other metropolitan regions in the U.S. Looking at metropolitan region-level estimates, we find that Raleigh is at the 90th percentile of the distribution of 306 elasticities of housing supply estimated by Baum-Snow and Han (2024). It is similar to Tucson, AZ; State College, PA; and Enid, OK. Since we make use of the tract-level estimates in our empirical analysis, rather than the metro region-level estimates, Figure A-7 shows the distribution of tract-level housing supply elasticities in the Raleigh metro area, and in the U.S. overall.

Household income w, average neighborhood income $\bar{w}_n$, and joint distribution of children skills and household income. When it comes to income, the empirics of the paper require the distribution of household income every year across neighborhoods of the county, and the joint distribution of household income and children skills. To construct these objects, we combine three sources of data. First, the ACS 1-year for 2014 through 2017 —the ACS 1-year is a repeated cross-section of households that reveals changes in distribution of household income for the county from year to year. Second, because the ACS 1-year does not contain any geographical level finer than the PUMA, we rely on the ACS 5-year 2013–2017 to get information about the distribution of income across space

within the county. In particular, the ACS 5-year gives counts of households by census tract and by brackets of household income. The ACS 5-year, however, does not allow us to study changes in income from year to year within the period. Finally, to get at the joint distribution of child skills and household income, we also need the NCERDC data, as the ACS does not contain any information about children skills. The NCERDC also gives the residential location and the ED status of public-school kindergartners in every year. We use the NCERDC data to locate families across neighborhoods and to set the share of families below \$50,000 in each neighborhood and year, and we use their pooled locations to split the tract-level ACS 5-year counts across the neighborhoods that overlap each tract. Here, we provide the details of the construction of the variables used in the empirical part of the paper.

1. For each of the years 2014 to 2017, the ACS 1-year 20xx gives the total number of households in the county in year 20xx, as well as household income in the past 12 months, expressed in 20xx\$. Since this income refers mostly to the previous calendar year, the ACS 1-year survey 20xx is used for the sample year 20xx-1 (that is, the school year starting in the fall of 20xx-1): the 2014 survey for 2013–14, through the 2017 survey for 2016–17. We convert household income to 2017\$ using annual end-of-period CPI given at https://fred.stlouisfed.org/series/CPIAUCSL. For the purpose of this paper, in the ACS 1-year, we define *families* as households with at least one child enrolled in kindergarten, first grade or second grade (public or private school); households are identified by survey year and household serial number. We use these three grades because they are the cohorts closest to the kindergarten decision we model, and pooling three grades rather than using kindergarten alone yields a reasonable number of households in each year and income bin. Based on this definition, we obtain the ratio of families to other households in the county for each year. We assign families and the rest of households to ten income bins based on their income expressed in 2017\$: household 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. From this, using household weights provided in the ACS 1-year, we can compute the share of families and other households in each bin in the county. Net household income w for households of each type ($w \in \{w_1, \ldots, w_{10}\}$) is constructed in two steps. First, gross income is assumed to be the mean income within the bracket. The average point in each bracket for families and other households is computed directly from the household-level information in the ACS 1-year using reported household income and household weights. Next, net-of-tax income is obtained from gross income us-

ing the NBER TAXSIM program, assuming the following household characteristics: married couple, spending 28 percent of their income on a mortgage.[35] Families are assumed to have one five-year-old (that is, kindergarten age) child; other households are assumed not to have children. Net average income in the bin is then used as the representative income for households of each type (families; other households) in the income bin in all subsequent steps.

2. Since the ACS 1-year does not have any tract-level information, we also use the ACS 5-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 code AIJA; https://www.nhgis.org/, accessed August 2021), which gives counts of families, separately with and without children, by census tract and bracket of household income. We use variables AIJAE004–AIJAE019, AIJAE040–AIJAE055, and AIJAE075–AIJAE090 to characterize the income distribution of our "families," and for other households we use all households (table "Household Income in the Past 12 Months (in 2017 Inflation-Adjusted Dollars)", NHGIS code AH1O) minus these families with own children, by tract and income bracket. Unlike the AIJA families without own children, this count includes non-family households, such as people living alone. We aggregate the 16 ACS 5-year brackets into the ten brackets listed above.

3. 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.[36] For reference, for the school year 2017–18, 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.[37]

4. To construct the annual distribution of families and other households across neighborhoods and income brackets, we combine three complementary data sources: tract-level income distributions from the ACS 5-year files, yearly neighborhood-

[35] https://taxsim.nber.org/taxsim35/index.html

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

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

level information from the NCERDC (for families, the number and ED shares of the kindergartners living in each neighborhood; for all households, the pooled residential locations of kindergartners, which we use to split each tract across neighborhoods), and year-specific county-level information from the ACS 1-year files. We proceed in two steps. First, for families, we take the number of families in each neighborhood and their split below and above \$50,000 from the kindergartners, and adjust the tract-level income distribution from the ACS 5-year files so that it matches the observed neighborhood-level ED shares from the NCERDC; for other households, we allocate the tract-level ACS 5-year counts, by income bracket, to neighborhoods with a fixed crosswalk from tracts to neighborhoods built from the locations of public-school kindergartners. Because the ACS 5-year data distinguish households with children under 18 but do not isolate families with kindergarten- or elementary-age children specifically, incorporating administrative ED shares allows us to better align the neighborhood income distribution with the demographic group of interest in the paper. Second, we adjust these neighborhood-level income distributions so that, in each year, their aggregation matches the county-wide income distribution observed in the ACS 1-year data across the ten income types. The ACS data cover all children residing in the county, including those enrolled in private schools, whereas the NCERDC includes only public school students. Using the ACS to discipline the aggregate distribution therefore ensures comprehensive population coverage. Our adjustment procedure preserves cross-sectional heterogeneity across neighborhoods and allows our empirical exercise to account for annual fluctuations in the county's overall income distribution.

Specifically, we proceed as follows (separately for families and other households). Let w_k be any of the ten income brackets/types, and let share$(w_k; ACS1, 20xx)$ be the share of families (resp. other households) with income w_k in the county in year 20xx, as given by the ACS 1-year 20xx. For families, the number of families in each neighborhood and their ED shares come from the NCERDC. Because the NCERDC covers public-school students only, the kindergartners are reweighted to include private-school families: each public kindergartner i with ED status e living in census tract z counts as $\omega_i = 1 + \rho_e c_z$ families, namely the kindergartner's own family plus $\rho_e c_z$ private-school families of the same ED status. Here ρ_e is the county ratio of private-school to public-school families with ED status e (ACS 1-year, survey years 2014–2017, households with a child in kindergarten to grade four; 0.056 for ED and 0.148 for non-ED families). The tract factor c_z is set so that, in tract z, private-school families make up the tract's private share of children enrolled in kindergarten

to grade four (ACS 5-year 2013–17, table B14002, shrunk toward the county share for small tracts), given the tract's public kindergartners pooled over 2013–2016.[38] The weights are thus the same in all years, and they are defined at the tract level: a neighborhood that overlaps several tracts combines the weights of the tracts its kindergartners live in. Let $f(n, 20xx)$ be the share of the county's families living in neighborhood n in year 20xx, and $e(n, 20xx)$ the share of ED families among them:

$$f(n, 20xx) = \frac{\sum_{i \in n} \omega_i}{\sum_{n'} \sum_{i \in n'} \omega_i}, \qquad e(n, 20xx) = \frac{\sum_{i \in n,\ \text{ED}} \omega_i}{\sum_{i \in n} \omega_i},$$

where the sums are over the kindergartners of year 20xx with known ED status living in n.

Let $\text{share}(w_k, n; \text{ACS5})$ be the share of the county's families (resp. other households) that have income w_k and live in neighborhood n, according to the ACS 5-year. Since the ACS 5-year counts are available by census tract and not by neighborhood, it is obtained with a crosswalk from tracts to neighborhoods. Let $N_k(z)$ be the ACS 5-year count of families (resp. other households) with income w_k in tract z. Let $\pi(n \mid z)$ be the share of the public-school kindergartners living in tract z who live in neighborhood n, pooling the four cohorts 2013–14 to 2016–17 and counting each kindergartner once (unweighted). The same crosswalk is used for families and for other households, and for all years. Then

$$\text{share}(w_k, n; \text{ACS5}) = \frac{\sum_z \pi(n \mid z)\, N_k(z)}{\sum_{k'} \sum_z N_{k'}(z)}.$$

For other households, this gives both their number in n (that of its tracts in the ACS, split in proportion to where the tracts' kindergartners live) and their income mix (the mix of its tracts, weighted by $\pi(n \mid z)$). For families, only the shape of this income mix within each side of \$50,000 is used; their number is given by $f(n, 20xx)$.

(a) First, for each year 20xx separately, initialize. For families, the share of the county's families living in n is $f(n, 20xx)$, and they are split below and above \$50,000 with the ED share $e(n, 20xx)$; within each side, they spread across

[38] Unlike the family counts above, private-school shares are measured over kindergarten to grade four, because Table B14002 does not separate grades one–two from three–four; the private share is nearly constant across these grades.

brackets like the ACS 5-year mix of n:

$$Pr(w_k, n) = \frac{\text{share}(w_k, n; \text{ACS5})}{\sum_{w \in \{w_1, \ldots, w_4\}} \text{share}(w, n; \text{ACS5})} \times e(n, 20xx) \times f(n, 20xx)$$

for $w_k \in \{w_1, \ldots, w_4\}$ and

$$Pr(w_k, n) = \frac{\text{share}(w_k, n; \text{ACS5})}{\sum_{w \in \{w_5, \ldots, w_{10}\}} \text{share}(w, n; \text{ACS5})} \times \big(1 - e(n, 20xx)\big) \times f(n, 20xx)$$

for $w_k \in \{w_5, \ldots, w_{10}\}$. The ACS 5-year enters these expressions only through the shape of the income mix of n within each side of \$50,000. At this stage, the neighborhood's share of the county's families, $\sum_k Pr(w_k, n)$, is $f(n, 20xx)$, which changes from year to year, and the share of its families below \$50,000 is $e(n, 20xx)$. If a neighborhood's tracts have no ACS family below (resp. above) \$50,000 but the neighborhood has ED (resp. non-ED) kindergartners, its families on that side are put in the bracket closest to \$50,000 (\$35–50,000, resp. \$50–75,000); this applies to one neighborhood in each year. For other households, initialize

$$Pr(w_k, n) = \text{share}(w_k, n; \text{ACS5})$$

for all $w_k \in \{w_1, \ldots, w_{10}\}$, the same in every year.

(b) In the next step, the share of each income bracket in each neighborhood combining ACS 5-year and NCERDC information (from the ACS 5-year only, for other households) ($Pr(w_k, n)$ constructed above) is adjusted so that the county-level share of the income bracket for the year coincides with the one observed in the ACS 1-year. Specifically, start with the first income bracket w_1:

- If $\sum_n Pr(w_1, n) < \text{share}(w_1; \text{ACS1}, 20xx)$, then
  - Construct the following quantity for each neighborhood n:

$$\left(\frac{\text{share}(w_1; \text{ACS1}, 20xx) - \sum_n Pr(w_1, n)}{\sum_n Pr(w_2, n)} \right) \times Pr(w_2, n)$$

  - Add this quantity to $Pr(w_1, n)$ and subtract it from $Pr(w_2, n)$.
- If $\sum_n Pr(w_1, n) > \text{share}(w_1; \text{ACS1}, 20xx)$, then
  - Construct the following quantity for each neighborhood n:

$$\left(\frac{\sum_n Pr(w_1, n) - \text{share}(w_1; \text{ACS1}, 20xx)}{\sum_n Pr(w_1, n)} \right) \times Pr(w_1, n)$$

– Subtract this quantity from $Pr(w_1, n)$ and add it to $Pr(w_2, n)$.

- At the end of this step for year 20xx, the updated county-level share of income bracket w_1 (that is, the updated $\sum_n Pr(w_1, n)$) coincides with the observed share$(w_1; \text{ACS1}, 20xx)$.

Repeat the process in step (b) for bins w_2 through w_9 (shares for w_{10} will be adjusted automatically as part of the adjustment for w_9). At the end of this iteration for year 20xx, the updated county-level share of income bracket w_k (that is, the updated $\sum_n Pr(w_k, n)$) coincides with the observed share$(w_k; \text{ACS1}, 20xx)$, for $w_k = w_1, \ldots, w_{10}$. Since households only move between adjacent brackets within a neighborhood, and by the same proportion in every neighborhood, this step leaves each neighborhood's share of the county's families (resp. other households) unchanged; it changes the income mix within neighborhoods. In particular, the adjustment between w_4 and w_5 moves families across the \$50,000 threshold, so after this step the share of a neighborhood's families below \$50,000 equals $e(n, 20xx)$ only approximately.

This procedure yields, for families, an average ED share of 28 percent over the four sample years (Table A-3), ranging from 21 to 34 percent across years (29, 34, 21 and 27 percent in 2013–2016), with no clear trend.

5. 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. $Pr(w \mid n)$ is computed as the ratio of $\frac{Pr(w,n)}{\sum_w Pr(w,n)}$, with $Pr(w, n)$ constructed in the previous step. Since $Pr(w, n)$ covers all families, including private-school families, $\bar{w}_n$ includes them. This is done separately for families and other households, each with its own net income levels. The two averages are then combined using the share of families among the households of n, $\frac{s_n^{fam}}{s_n^{fam} + R\, s_n^{oth}}$, where s_n^{fam} and s_n^{oth} are the neighborhood's shares of the county's families and other households and R is the county ratio of other households to families in the ACS 1-year. The distribution of average neighborhood income in the county is shown in Figure 1b in the main text.

6. 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. On the other hand, the ACS, which shows household counts by income bracket, does not

contain any information about children skills. Given this data limitation, we construct the joint share $Pr(w,a)$ for every skill type $a \in \{a_1,\ldots,a_{10}\}$ and every income type $w \in \{w_1,\ldots,w_{10}\}$ assuming a and w are conditionally independent given ED and n. Specifically: $Pr^{all}(w,a) = \sum_n \sum_{ED} Pr(w|n,ED)Pr(a|n,ED)Pr(ED|n)Pr(n)$ for all families, where $Pr(w|n,ED)$, the share of income type w in neighborhood n conditional on ED status; $Pr(ED|n)$, the ED share in neighborhood n; and $Pr(n)$, the share of families living in neighborhood n, are constructed as described above. In these three objects, ED stands for income below \$50,000 (brackets w_1 to w_4), so $Pr(ED|n)$ is the share of the neighborhood's families below \$50,000 after step (b). Before the ACS 1-year adjustment, it equals $e(n,20xx)$. The share of skill type a among public-school kindergartners with a given ED status eds living in neighborhood n, $Pr(a|n,eds)$, is obtained directly from the NCERDC data. In the model, the skill distribution given income refers to public-school families, since skills are observed only for public-school children. We therefore compute the skill mix within each income bin for public-school families, and keep the income totals of all families. Specifically, let $r(n,ED)$ be the share of public-school families in cell (n,ED): the number of public kindergartners in the cell divided by the weighted number of families $\sum_i \omega_i$, with the weights $\omega_i = 1 + \rho_e c_z$ defined in step 4. For cells with no kindergartner, we use the county ratio for that ED status. Then

$$Pr(w,a) = Pr^{pub}(a \mid w)\, Pr(w), \quad \text{with}$$
$$Pr^{pub}(a \mid w) = \frac{\sum_n \sum_{ED} r(n,ED)\, Pr(w|n,ED)Pr(a|n,ED)Pr(ED|n)Pr(n)}{\sum_{a'} \sum_n \sum_{ED} r(n,ED)\, Pr(w|n,ED)Pr(a'|n,ED)Pr(ED|n)Pr(n)},$$

where $Pr(w) = \sum_n \sum_{ED} Pr(w|n,ED)Pr(ED|n)Pr(n)$ is the all-family income distribution. The family income distribution is thus unchanged (it matches the ACS 1-year county distribution), and only the skill mix within each income bin is that of public-school families.

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	45.92	51.71	0.00	309.75
Avg sale price by sqft	123.71	30.82	53.66	274.70
Avg MLS regulation (acre)	0.33	0.29	0.09	1.67
Avg minimum house size (in sqft)	618.75	183.47	465.68	1473.71
Avg # of school options (excl. base)	17.75	0.64	15.75	18.25
Avg # of school options w/ transp. (excl. base)	4.06	0.70	1.50	5.75
Distance to base sch. (miles)	3.56	3.16	0.09	17.56
Avg. distance to option sch.	11.10	6.22	0.31	34.94
Has base change during period	0.22	0.41	0.00	1.00
Has catchment area change during period	0.63	0.48	0.00	1.00
Has change in option set during period	1.00	0.00	1.00	1.00
Avg # of student obs. per year	28.05	28.63	2.50	199.25
Share of ED families (income $<$ \$50k)	0.34	0.27	0.00	0.92
# of neighborhoods in sample	*310*			
# of neighborhood-year obs.	*1,240*			
PANEL B: STUDENT SAMPLE				
Is economically disadvantaged (ED, family income $<$ \$50k)	0.28	0.45	0.00	1.00
Attends base, cond. on being ED	0.76	0.43	0.00	1.00
Attends base, cond. on being non-ED	0.73	0.44	0.00	1.00
Attends option w/ transp., cond. on ED	0.16	0.37	0.00	1.00
Attends option w/ transp., cond. on non-ED	0.21	0.41	0.00	1.00
Attends option w/o transp., cond. on ED	0.08	0.27	0.00	1.00
Attends option w/o transp., cond. on non-ED	0.06	0.24	0.00	1.00
Log-skills (standardized w/in cohort) cond. on ED	-0.56	0.83	-1.44	3.43
Log-skills (standardized w/in cohort) cond. on non-ED	0.30	0.95	-1.44	4.22
# of student-yr obs.	*34,785*			
PANEL C: PUBLIC SCHOOL SAMPLE				
Avg peer quality	1.54	0.50	0.49	3.28
School value added	-0.04	0.26	-0.88	0.56
Share econ. disadv. (ED) students	0.38	0.22	0.00	0.90
# of student obs. in sample (per year)	80.07	30.05	4.00	197.00
Is option school for some address	0.85	0.36	0.00	1.00
Has catchm. area change during period (base)	0.55	0.50	0.00	1.00
Has elig./transp. change during period (options only)	0.91	0.28	0.00	1.00
Is oversubscribed (options only)	0.64	0.48	0.00	1.00
# of schools in sample	*111*			
# of school-year obs.	*427*			

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. In Panels A and B, ED shares are shares of families with gross income below \$50,000, built from NCERDC and ACS data as described in Appendix A.2.3; in Panel C, they are shares of students reported as economically disadvantaged in the NCERDC data.

A.2.5 Housing Expenditure for Families with Children in Public vs Private School

In this subsection, we examine differences in housing expenditure between families whose children attend public versus private schools. The objective is to test our assumption which states that households allocate a constant share of after-tax income to housing, regardless of the type of school their children attend. Using data from the ACS, we compare the share of after-tax income spent on housing between families with children in private school and families with children in public school groups (that is, no child in private school).

For renters, housing expenditure is measured as annual rent, which is directly observed in the ACS. For homeowners, housing expenditure is taken to be the house value divided by 15, which is consistent with an annual user cost of 6.7 percent of the house value, in line with standard estimates of depreciation, maintenance, property taxes, and the opportunity cost of capital. Note that 15 is also the factor we use to annuitize house prices (see Appendix A.2.3). We use the ACS 1-year samples for survey years 2014 to 2017 (incomes of 2013 to 2016), and restrict the sample to the eight PUMAs partitioning Wake County in and to households with a child enrolled in kindergarten to grade four. We drop renters who pay no cash rent, for whom neither rent nor a house value is available, and households with zero income.

In Table A-4, the median share of after-tax income spent on housing is shown separately for renters and owners, as well as for the full sample. We reports the p-value of the difference in the housing share between public-school and private-school families. Across the table, the shares are very similar across the two groups of families. The 25 percent share for renters is also similar to our β, which is taken from Davis and Ortalo-Magné (2011), who focus on renters. Using annual mortgage as the value of housing expenditures for owners with a loan or mortgage yields lower housing shares —consistent with mortgages including the accumulation of equity— but no difference across the two groups of families (17.9 percent for public-school families vs. 16.3 percent) for private-school families).

Table A-4: Share of Income Spent on Housing by Attended-School Type

	Renters	Owners	All
Public School	25.4	20.7	21.8
Private School	26.7	20.3	21.2
p-value (difference)	0.687	0.605	0.444

The table the median share of after-tax income spent on housing separately for families with a child in private school and families with no child in private school. the type of school attended by the child (public versus private). After-tax income is computed using TAXSIM. Fow home-owners, housing expenditure is imputed as the house value divided by 15, which implicitly assumes an annual user cost of 6.7 percent of the house value, consistent with standard estimates of depreciation, maintenance, property taxes, and the opportunity cost of capital. The sample excludes renters paying no cash rent and households with zero income.

B House-Price Capitalization of School Quality: Evidence from Policy Variation

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

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

	Change in Log House Price per Sq. Ft.		
	(1)	(2)	(3)
Change in School Peer Quality (Log)	0.073*	0.073*	0.076*
	(0.041)	(0.041)	(0.040)
Change in School Value Added			0.023
			(0.019)
Observations	807	807	807
Sale-Year F.E.	No	Yes	Yes

The table shows the effect of policy-induced changes in school peer quality and school value added on house prices. The outcome variable represents the change, between two consecutive sale years, in the neighborhood's mean log sale price per square foot. The regressor of interest is the policy-induced change in the peer quality of the neighborhood's base-school catchment area (Equation (4.2)): the average kindergarten-entry skill level of the kindergartners of our final sample living in the catchment area of the house's assigned base school, comparing, for the same cohort of kindergartners (the previous school year's), the catchment area of the base school assigned in the current school year with that of the base school assigned in the previous school year. Column (3) adds the change in the value added of the neighborhood's assigned base school between the two years. Details on the construction of the skill measure and value added are provided in Appendix A.2. All regressions are weighted by the number of house sales in the neighborhood and year in the Wake County real estate transaction records (after the sample restrictions of Appendix A.2) to recover the aggregate effects on prices. 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.

C Estimation Appendix

Skill type is discretized into $K = 10$ 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 ($T = 4$), and N the number of neighborhoods ($N = 310$). We estimate model parameters using $1 + (K-1) \times 2 + 1 + 1 + 2 + 2 + 1 + 24 + 1 + (N-1) + (N-1) + N \times 2 + 1 + N = 1{,}600$ moments, which we define formally here.

C.1 Data Moments

To describe the construction of empirical moments from the data, 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 part of (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 $\mathcal{A}$. For the empirical moments, the average peer quality of school s in year t is computed from the data as $\bar{a}_{s,t} = \frac{\sum_{i|s(i)=s,t(i)=t} a(i)}{\#\{i|s(i)=s,t(i)=t\}}$. Grandfathered students are excluded from $\bar{a}_{s,t}$ and from every school-side moment below (moments 2–7), but not from the neighborhood moments.

1. Regression coefficient b of the change in house prices on changes in associated school peer quality; see Equation (4.1) and column (2) of Table B-1 [1 coefficient].

2. Average log peer quality in the (public) school attended by a child with skill type $a \in \{a_2, \ldots, a_{10}\}$ and ED status *eds* relative to a child with skill type a_1 and ED status *eds*, for each *eds* $\in \{\text{ED}, \text{non-ED}\}$ [$(K-1) \times 2$ moments].

$$\left(\frac{1}{T} \sum_{t=1}^{T} \frac{\sum_{i \in \{i|a(i)=a,t(i)=t,i \text{ is } eds\}} \ln \bar{a}_{s(i)}}{\#\{i \mid a(i)=a, t(i)=t, i \text{ is } eds\}} \right) / \left(\frac{1}{T} \sum_{t=1}^{T} \frac{\sum_{i \in \{i|a(i)=a_1,t(i)=t,i \text{ is } eds\}} \ln \bar{a}_{s(i)}}{\#\{i \mid a(i)=a_1, t(i)=t, i \text{ is } eds\}} \right)$$

for $a \in \{a_2, \ldots, a_{10}\}$, and *eds* $\in \{\text{ED}, \text{non-ED}\}$

3. Average (across neighborhoods and across years) within-neighborhood gap across ED status in attended school log peer quality [1 moment].

$$\frac{1}{\#\mathcal{E}} \sum_{(n,t) \in \mathcal{E}} \left(\frac{\sum_{\{i|n(i)=n,t(i)=t,i \text{ is non-ED}\}} \ln \bar{a}_{s(i)}}{\#\{i \mid n(i)=n, t(i)=t, i \text{ is non-ED}\}} - \frac{\sum_{\{i|n(i)=n,t(i)=t,i \text{ is ED}\}} \ln \bar{a}_{s(i)}}{\#\{i \mid n(i)=n, t(i)=t, i \text{ is ED}\}} \right)$$

where $\mathcal{E}$ is the set of neighborhood-year pairs (n,t) with both ED and non-ED students.

4. Average (over years) correlation between school value added and school peer quality for the set $\mathcal{S}$ of public schools [1 moment].

$$\frac{1}{T} \sum_{t=1}^{T} \frac{\sum_{s \in \mathcal{S}} (\bar{a}_{s,t} - \bar{\bar{a}}_t)(A_s - \bar{A})}{\sqrt{\sum_{s \in \mathcal{S}} (\bar{a}_{s,t} - \bar{\bar{a}}_t)^2} \sqrt{\sum_{s \in \mathcal{S}} (A_s - \bar{A})^2}},$$

$$\text{with} \qquad \bar{\bar{a}}_t \equiv \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \bar{a}_{s,t}, \qquad \text{and} \qquad \bar{A} \equiv \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} A_s.$$

In year t, $\mathcal{S}$ (and hence $\bar{\bar{a}}_t$ and $\bar{A}$) covers the schools open in year t only, since a few schools open during the sample period.

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

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{\{i|t(i)=t,\, s(i)\in\mathcal{T}_{n(i),t},\, i \text{ is } eds\}} d_{s(i),n(i)}}{\#\{i \mid t(i)=t,\ s(i)\in\mathcal{T}_{n(i),t},\ i \text{ is } eds\}}\right), \qquad eds \in \{\text{ED, non-ED}\}.$$

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

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{\{i|t(i)=t,\, s(i)\in\mathcal{NT}_{n(i),t},\, i \text{ is } eds\}} d_{s(i),n(i)}}{\#\{i \mid t(i)=t,\ s(i)\in\mathcal{NT}_{n(i),t},\ i \text{ is } eds\}}\right), \qquad eds \in \{\text{ED, non-ED}\}.$$

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

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\sum_n \widetilde{var}_{n,t}\times\#\{i \mid n(i)=n \text{ and } t(i)=t\}}{\sum_n \#\{i \mid n(i)=n \text{ and } t(i)=t\}}$$

$$\text{with} \qquad var_{n,t} = \frac{\sum_{s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t}}\left[\left(\Pr(s \mid n,t) - \frac{\sum_{s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t}}\Pr(s|n,t)}{\#(\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t})}\right)^2\right]}{\#(\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t})-1}$$

$$\text{and} \qquad \Pr(s \mid n,t) = \frac{\#\{i \mid s(i)=s, n(i)=n \text{ and } t(i)=t\}}{\#\{i \mid n(i)=n \text{ and } t(i)=t\}}$$

$var_{n,t}$ is the sample variance (denominator: number of public schools in the choice set minus one), with $\Pr(s \mid n,t) = 0$ for schools in the choice set that no student of n attends, and $var_{n,t} = 0$ when the choice set has a single public school. Because $\Pr(s \mid n,t)$ is a frequency over the $\mathcal{N}_{n,t} = \#\{i \mid n(i)=n \text{ and } t(i)=t\}$ students of the neighborhood-year, $var_{n,t}$ contains multinomial sampling noise. Let $J_{n,t} = \#(\mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t})$ be the number of public schools in the choice set in year t. Let $\mathbb{P}_{s|n,t}$ denote the underlying probability that a student of (n,t) attends s, of which $Pr(s \mid n,t)$ is the sample frequency. The expectation of $var_{n,t}$ is the variance of the $\mathbb{P}_{s|n,t}$ across schools (which is what the model computes) plus $\left(1-\sum_s \mathbb{P}^2_{s|n,t}\right)/\left((J_{n,t}-1)\,\mathcal{N}_{n,t}\right)$. The moment therefore uses the noise-corrected variance

$$\widetilde{var}_{n,t} = var_{n,t} - \frac{1-\sum_{s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t}}\Pr(s \mid n,t)^2}{(J_{n,t}-1)(\mathcal{N}_{n,t}-1)},$$

where the correction term is an unbiased estimate of the noise. No correction is applied when $\mathcal{N}_{n,t} = 1$ or $J_{n,t} = 1$. The corrected variance can be slightly negative in small neighborhood-years; it is kept as is, since only the average enters the moment. The correction lowers the moment by about two percent.

8. Private-school attendance shares at the PUMA level separately for households with a child enrolled in kindergarten to grade four with annual household income (in 2017 dollars) at most \$50,000; between \$50,000 and \$125,000; and above \$125,000 —obtained directly from the ACS (one-year samples 2013–17, pooled) [8×3 moments]. A household with one child in public school and one in private school counts as attending public school.

9. Share attending private school among applicants to option schools who lose the admission lottery [1 moment] —taken from Dur et al. (2025).

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

$$\frac{\sum_{t=1}^{T} M_t \frac{\sum_{i|n(i)=n,t(i)=t} \omega_i}{\sum_{i|t(i)=t} \omega_i}}{\sum_{t=1}^{T} M_t}, \text{ from the NCERDC data.}$$

Here $\omega_i = 1 + \rho_{e(i)} c_{z(i)}$ is the private-school weight of child i (see A.2.3): each public kindergartner also stands for the private-school families of their tract and ED status. M_t is the county mass of families in year t (ACS 1-year), so years are weighted by the number of families.

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

12. Average income for families and the other households in each neighborhood [$N \times 2$ moments]—constructed for each year as explained in A.2.3 and averaged over the four years, weighting each year by the number of families (resp. other households) living in the neighborhood that year.

13. Take-up rate from the MTO experiment reported in Sanbonmatsu et al. (2011) [1 moment].

14. Average (over time) house prices in each neighborhood [N moments].

$$\frac{1}{T}\sum_t \text{price}_{n,t} \quad \text{for each } n$$

where $\text{price}_{n,t}$, the average price per square foot in neighborhood n in year t, is constructed as explained in A.2.3.

C.2 Construction of Model-implied Moments

Recall from Section 2.3 that

$$\pi_{n;w,a} = \frac{\exp\left(\frac{\hat{v}_{n;w,a}}{\sigma_N}\right)}{\sum_{\tilde{n}} \exp\left(\frac{\hat{v}_{\tilde{n};w,a}}{\sigma_N}\right)} \quad \text{with} \quad \hat{v}_{n;w,a} = \alpha_{0n} + \ln(w)\,\alpha_{1n} + \bar{v}_{n;w,a}(\mathcal{L}_n),$$

where $\bar{v}_{n;w,a}(\mathcal{L}_n) = \mathbb{E}\left[\max_{s\in\mathcal{L}_n}\left\{p_s v_{s|n;w,a,i} + (1-p_s)v^{fallback}_{n;w,a,i}\right\}\right]$. The probability of applying to school s conditional on neighborhood n and family type (w,a),

$$\pi_{s|n;w,a} = \Pr\left[p_s v_{s|n;w,a,i} + (1-p_s)v^{fallback}_{n;w,a,i} \geq p_{\tilde{s}} v_{\tilde{s}|n;w,a,i} + (1-p_{\tilde{s}})v^{fallback}_{n;w,a,i} \;\; \forall \tilde{s} \in \mathcal{L}_n\right],$$

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

$$\pi^{attend}_{s,n;w,a} = \begin{cases} \pi_{s,n;w,a} \times p_s & \text{if} \;\; s \in \mathcal{T}_n \cup \mathcal{NT}_n \\ \pi_{s,n;w,a} + \pi^{fallback}_{s,n;w,a} & \text{if} \;\; s \in \mathcal{B}_n \cup \mathcal{P}_n \end{cases}$$

where

$$\pi^{fallback}_{s,n;w,a} = \sum_{\tilde{s}\in\mathcal{T}_n\cup\mathcal{NT}_n} \pi_{\tilde{s},n;w,a}(1-p_{\tilde{s}}) \Pr\left(v^{fallback}_{n;w,a,i} = v_{s|n;w,a,i} \,\middle|\, i \text{ applies to } \tilde{s}\right), \quad s \in \mathcal{B}_n \cup \mathcal{P}_n\,.$$

Then, the probability that a family of type (w,a) who lives in neighborhood n attends school $s \in \mathcal{L}_n$ is $\pi^{attend}_{s|n;w,a} = \pi^{attend}_{s,n;w,a} / \pi_{n;w,a}$.
Finally, the joint distribution of family types (w,a) is denoted $\phi_{w,a}$. Below, we use the notation $\phi_{w,a|eds}$ to denote the joint distribution of family types (w,a) conditional on ED status eds; and $\phi_{w|a}$ (resp. $\phi_{w|a,eds}$) to denote the distribution of family income types w

conditional on skill type a (resp. skill type a and ED status eds). When needed, cohort-specific variables are further indexed by t. In the model, ED status is defined by income: ED for income types w_1–w_4 and non-ED for w_5–w_{10}. Model peer quality $\bar{a}_{s,t}$ is the average skill of the children attending public school s in year t, and only schools open in year t enter the school-side moments.
The model-implied moments are:

1. Regression coefficient b obtained when estimating Equation (4.1) on model-generated data [1 coefficient].

2. Average log peer quality in the school attended by a child with skill type $a \in \{a_2, \ldots, a_{10}\}$ and ED status eds relative to a child with skill type a_1 and ED status eds, for each $eds \in \{\text{ED}, \text{non-ED}\}$ [$(K-1) \times 2$ moments].

$$\frac{\frac{1}{T}\sum_t \left[\frac{\sum_w \phi_{w|a,eds} \sum_n \sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s,n;w,a,t} \ln \bar{a}_{s,t}}{\sum_w \phi_{w|a,eds} \sum_n \sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s,n;w,a,t}} \right]}{\frac{1}{T}\sum_t \left[\frac{\sum_w \phi_{w|a_1,eds} \sum_n \sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s,n;w,a_1,t} \ln \bar{a}_{s,t}}{\sum_w \phi_{w|a_1,eds} \sum_n \sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s,n;w,a_1,t}} \right]}.$$

3. Average (across neighborhoods and across years) within-neighborhood gap across ED status in attended school log peer quality [1 moment].[39]

$$\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s|n;\text{non-ED},t} \times \ln \bar{a}_{s,t}}{\sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s|n;\text{non-ED},t}} - \frac{\sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s|n;\text{ED},t} \times \ln \bar{a}_{s,t}}{\sum_{s \in \mathcal{L}_{n,t} \setminus \mathcal{P}_{n,t}} \pi^{attend}_{s|n;\text{ED},t}} \right)$$

$$\text{where} \quad \pi^{attend}_{s|n;\text{ED},t} = \frac{\sum_{w,a} \pi_{n;w,a,t} \pi^{attend}_{s|n;w,a,t} \times \phi_{w,a|\text{ED}}}{\sum_{w,a} \pi_{n;w,a,t} \phi_{w,a|\text{ED}}}$$

$$\text{and} \quad \pi^{attend}_{s|n;\text{non-ED},t} = \frac{\sum_{w,a} \pi_{n;w,a,t} \pi^{attend}_{s|n;w,a,t} \times \phi_{w,a|\text{non-ED}}}{\sum_{w,a} \pi_{n;w,a,t} \phi_{w,a|\text{non-ED}}}$$

4. Average (over years) correlation between school value added and model-generated school peer quality, for the set $\mathcal{S}$ of public schools [1 moment].

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\sum_{s \in \mathcal{S}} (\bar{a}_{s,t} - \bar{\bar{a}}_t)(A_s - \bar{A})}{\sqrt{\sum_{s \in \mathcal{S}} (\bar{a}_{s,t} - \bar{\bar{a}}_t)^2} \sqrt{\sum_{s \in \mathcal{S}} (A_s - \bar{A})^2}},$$
$$\text{with} \quad \bar{\bar{a}}_t \equiv \frac{1}{|\mathcal{S}|}\sum_{s \in \mathcal{S}} \bar{a}_{s,t}, \quad \text{and} \quad \bar{A} \equiv \frac{1}{|\mathcal{S}|}\sum_{s \in \mathcal{S}} A_s.$$

[39] The average is over all (n,t), which in the model is equivalent to averaging over the set $\mathcal{E}$ used for the empirical counterpart, since every neighborhood-year has both ED and non-ED families.

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

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{w,a}\phi_{w,a|eds}\ \sum_{n}\sum_{s\in\mathcal{T}_{n,t}}\pi^{attend}_{s,n;w,a,t}\ d_{s,n}}{\sum_{w,a}\phi_{w,a|eds}\ \sum_{n}\sum_{s\in\mathcal{T}_{n,t}}\pi^{attend}_{s,n;w,a,t}}, \qquad eds\in\{\text{ED},\text{non-ED}\}.$$

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

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{w,a}\phi_{w,a|eds}\ \sum_{n}\sum_{s\in\mathcal{NT}_{n,t}}\pi^{attend}_{s,n;w,a,t}\ d_{s,n}}{\sum_{w,a}\phi_{w,a|eds}\ \sum_{n}\sum_{s\in\mathcal{NT}_{n,t}}\pi^{attend}_{s,n;w,a,t}}, \qquad eds\in\{\text{ED},\text{non-ED}\}.$$

7. Average (over neighborhoods and years) variance of neighborhood-level attendance share across public schools [1 moment].

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{n}\Pi_{n,t}\,\text{Var}_{n,t}}{\sum_{n}\Pi_{n,t}},$$

$$\begin{aligned}
\text{where}\quad & \text{Var}_{n,t}\equiv\frac{\sum_{s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t}}\left(\Pi_{s|n,t}-\bar{\Pi}_{n,t}\right)^2}{\#(\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t})-1},\\
\text{and}\quad & \Pi_{n,t}\equiv\sum_{w,a}\pi_{n;w,a,t}\,\phi_{w,a}\sum_{s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t}}\pi^{attend}_{s|n;w,a,t},\\
\text{with}\quad & \Pi_{s|n,t}\equiv\frac{1}{\Pi_{n,t}}\sum_{w,a}\pi_{n;w,a,t}\,\phi_{w,a}\,\pi^{attend}_{s|n;w,a,t}, \qquad s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t},\\
\text{and}\quad & \bar{\Pi}_{n,t}\equiv\frac{1}{\#(\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t})}\sum_{s\in\mathcal{L}_{n,t}\setminus\mathcal{P}_{n,t}}\Pi_{s|n,t}.
\end{aligned}$$

$\Pi_{n,t}$ is the mass of public-school families in n, so the shares $\Pi_{s|n,t}$ sum to one over the public schools of the menu. $\text{Var}_{n,t}$ is the sample variance, as in the data. Since the mass of public-school families is normalized to one each year, $\sum_n \Pi_{n,t}=1$.

8. Private-school attendance shares for families in income brackets w_1–w_4, w_5–w_7, and w_8–w_{10} at the PUMA level [8 × 3 moments].

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{n\in\text{PUMA}}\sum_{w\in\mathcal{W}_g}\sum_{a}\phi_{w,a}\ \sum_{s\in\mathcal{P}_{n,t}}\pi^{attend}_{s,n;w,a,t}}{\sum_{n\in\text{PUMA}}\sum_{w\in\mathcal{W}_g}\sum_{a}\phi_{w,a}\ \pi_{n;w,a,t}}, \qquad \text{for each PUMA and } g\in\{1,2,3\},$$

with $\mathcal{W}_1 = \{w_1, \ldots, w_4\}$, $\mathcal{W}_2 = \{w_5, w_6, w_7\}$, and $\mathcal{W}_3 = \{w_8, w_9, w_{10}\}$.

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

$$\frac{\sum_n \sum_{w,a} \phi_{w,a} \ \sum_{s \in \mathcal{P}_{n,t}} \pi^{fallback}_{s,n;w,a}}{\sum_n \sum_{w,a} \phi_{w,a} \ \sum_{\tilde{s} \in \mathcal{T}_n \cup \mathcal{NT}_n} (1 - p_{\tilde{s}}) \, \pi_{\tilde{s},n;w,a}}.$$

All terms are for $t = 2015$ (the year of the lottery data). A lottery loser goes to private school when their fallback (the better of their base school and their private school) is the private school.

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

$$\frac{\sum_t M_t \, s_{n,t}}{\sum_t M_t}, \quad s_{n,t} \equiv \frac{\sum_{w,a} \pi_{n;w,a,t} \, \phi_{w,a,t}}{\sum_{w,a} \phi_{w,a,t}} \text{ for each } n$$

where M_t is the county mass of families in year t relative to 2013 (ACS 1-year). Shares include both public- and private-school families.

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

$$\frac{\sum_t M^o_t \, s^o_{n,t}}{\sum_t M^o_t}, \quad s^o_{n,t} \equiv \sum_w \pi^o_{n;w,t} \, \phi^o_{w,t}, \ \text{ for each } n$$

where $\phi^o_{w,t}$ is the year-t income distribution of other households (summing to one) and M^o_t is their county mass in year t.

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

$$\frac{1}{\sum_t M_t s_{n,t}} \sum_t M_t s_{n,t} \frac{\sum_w w_t \sum_a \pi_{n;w,a,t} \phi_{w,a,t}}{\sum_w \sum_a \pi_{n;w,a,t} \phi_{w,a,t}} \text{ and}$$
$$\frac{1}{\sum_t M^o_t s^o_{n,t}} \sum_t M^o_t s^o_{n,t} \frac{\sum_w w^o_t \pi^o_{n;w,t} \phi^o_{w,t}}{\sum_w \pi^o_{n;w,t} \phi^o_{w,t}} \text{ for each } n$$

The years are weighted by the number of families (resp. other households) living in n, and w_t and w^o_t are the year-t after-tax incomes of families and other households in each income bin (TAXSIM).

13. Share of eligible families who take up the voucher in the following policy: Families in income brackets $\{w_1, w_2, w_3\}$ (that is, with gross household income below

$35,000 per year in 2017 dollars) living in neighborhoods with more than 40 percent of households in brackets $\{w_1, w_2, w_3\}$ are offered a housing voucher that pays the difference between the 40th percentile of the distribution of housing expenditures in Wake County (across all households, in the 2015 equilibrium)[40] and the family's own housing budget βw, when positive, which they can use if they choose to live in a neighborhood where the share of households in brackets $\{w_1, w_2, w_3\}$ is below 40 percent (at baseline).[41] [1 moment]. The experiment is simulated in 2015. Eligibility is based on where families live in the baseline equilibrium, and take-up is the share of eligible families who use the voucher to live in a neighborhood below the 40 percent threshold.

14. Average (over time) equilibrium house prices in each neighborhood [N moments].

$$\frac{1}{T}\sum_t r_{n,t} \quad \text{for each } n$$

Here $r_{n,t}$ is the annual rent per square foot, compared with the data price per square foot divided by the price-to-rent ratio of 15.

C.3 Identification Plots

In this section, we show that each of our key moments moves monotonically with specific parameters, in line with the identification arguments presented in Section 4.1. Each panel of Figure C-1 moves one parameter away from its estimated value by the percentage reported on the horizontal axis, holds every other parameter at its estimated value, resolves the equilibrium of the model, and reports on the vertical axis the value taken by the moment that parameter informs.

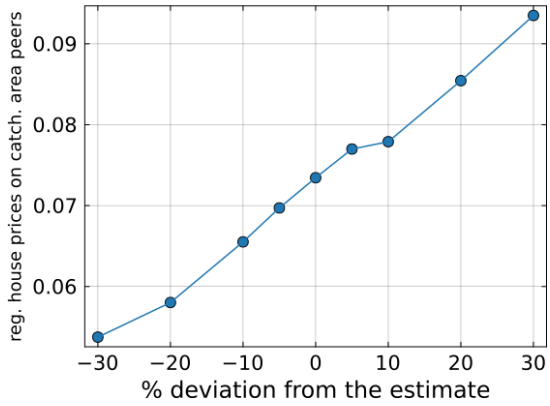
Three of the plots move more than one parameter at a time. The coefficient b of Equation (4.1) measures the value of peers relative to house prices. The effective importance of peers in school choice depends on γ relative to the other determinants of school utility

[40] The fair market rent for Wake County in our sample years is given by HUD's Fair Market Rent Documentation: https://www.huduser.gov/portal/datasets/fmr/fmrs/FY2016_code/2016summary.odn.

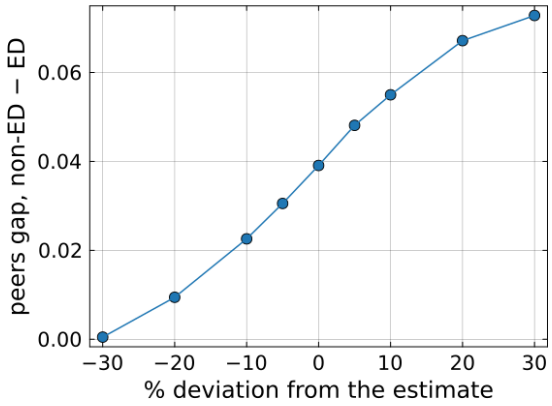
[41] Our choice of brackets $\{w_1, w_2, w_3\}$ as a criterion for eligibility is guided by the following facts. To be eligible for the MTO vouchers, families needed to be eligible for public housing and extremely-low or very-low income, that is, with an income below 30 or 50 percent of the local median income. In 2017, these limits were $24,600 and $40,100 for a family of four in Wake County (https://www.huduser.gov/portal/datasets/il.html). The former corresponds to the 2017 poverty line for a family of four. In practice though, recipients turned out to be overwhelmingly extremely low income — Kling et al. (2007) report that 75 percent of their sample receives Aid for Families with Dependent Children (AFDC), implying that they met the AFDC eligibility criterion, which, while being state-specific, tended to be close to the federal poverty line.

(commuting costs, idiosyncratic tastes, private schools' unobserved characteristics). Figure C-1(a) therefore scales γ, κ, σ_S, σ_P and χ by a common factor, in order to keep the relative importance of peers in school choice constant, and vary only their value relative to house prices. The peer-quality gap by ED status informs the overall level of non-ED families' valuations relative to ED families', so Figure C-1(b) scales all $\gamma_{\text{non-ED},a}$ by a common factor. Similarly, the share of students attending private school after losing an option lottery informs σ_P conditional on the overall share of students enrolled in private school. Figure C-1(j) therefore re-estimates the private-school valuations χ at each value of σ_P, so that private-school enrollment shares continue to match their empirical counterparts.

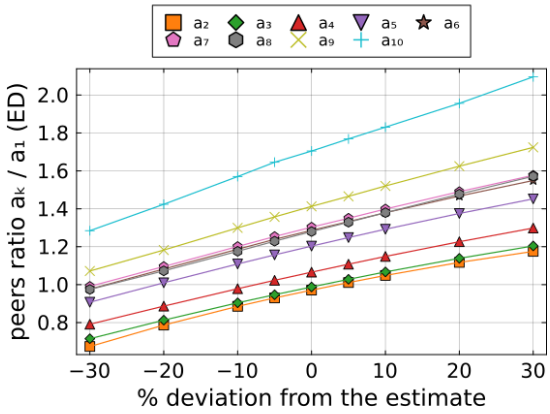
Figure C-1: Monotonicity of Key Moments in the Parameters They Inform



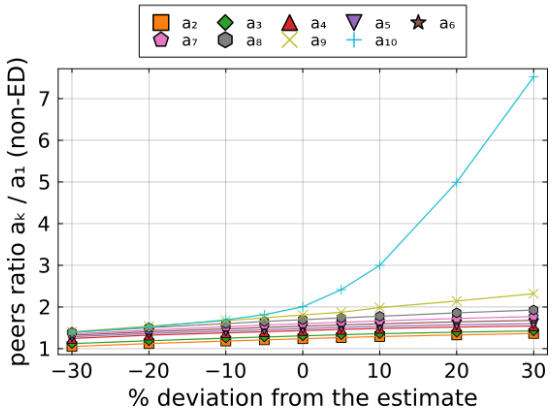
(a) School valuation level and b, Eq. (4.1)



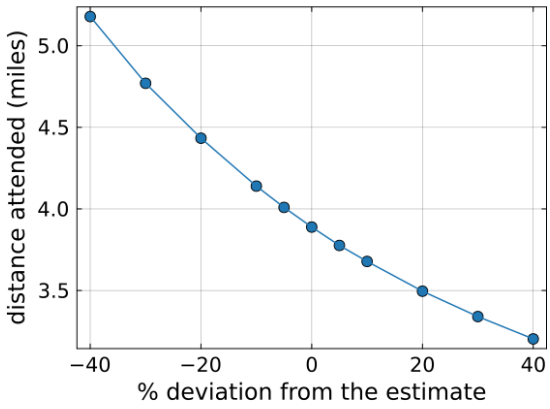
(b) $\gamma_{nonED,a}$ and peer-quality gap by ED status



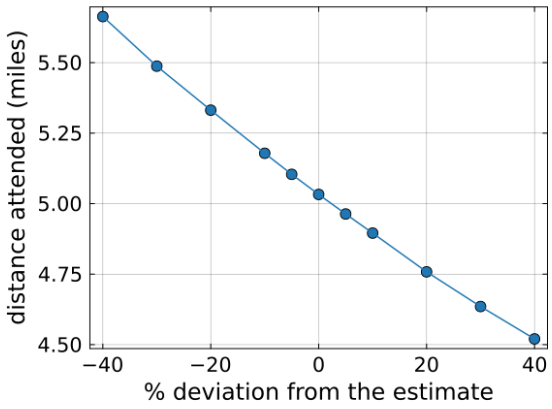
(c) $\gamma_{ED,a}$ and relative peer composition



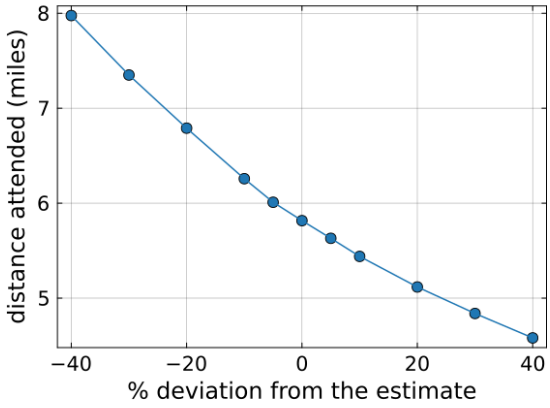
(d) $\gamma_{nonED,a}$ and relative peer composition



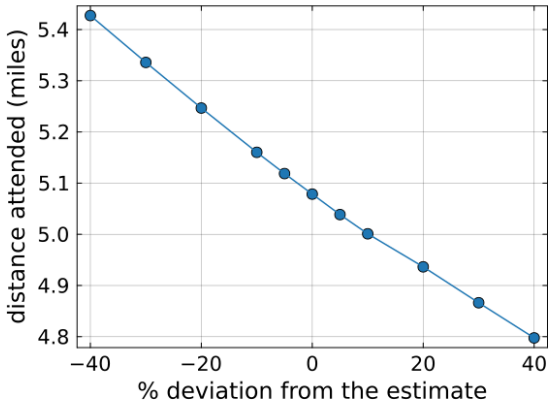
(e) $\kappa_{NT,ED}$: distance, ED, no transportation



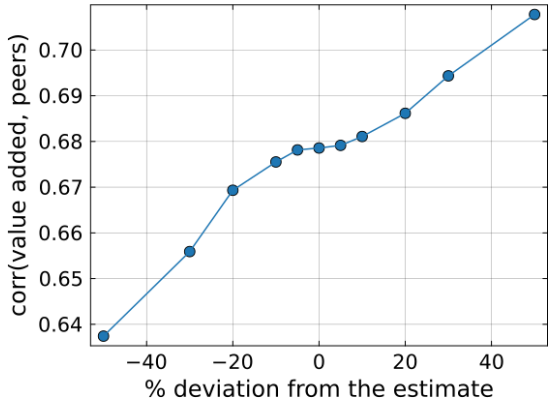
(f) $\kappa_{T,ED}$: distance, ED, transportation



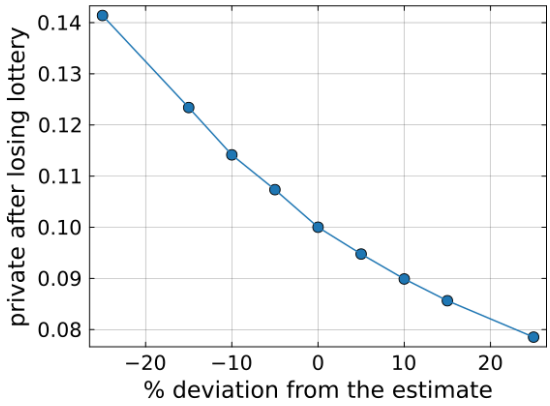
(g) $\kappa_{NT,nonED}$: distance, non-ED, no transportation



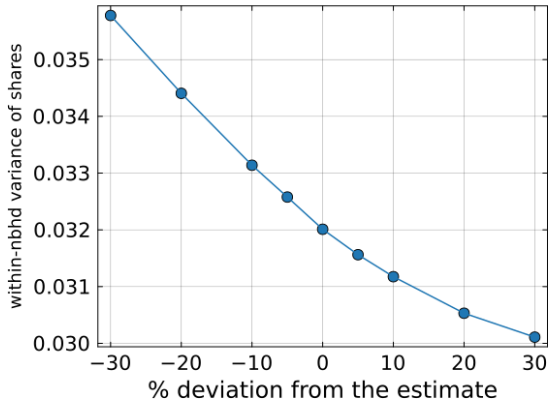
(h) $\kappa_{T,nonED}$: distance, non-ED, transportation



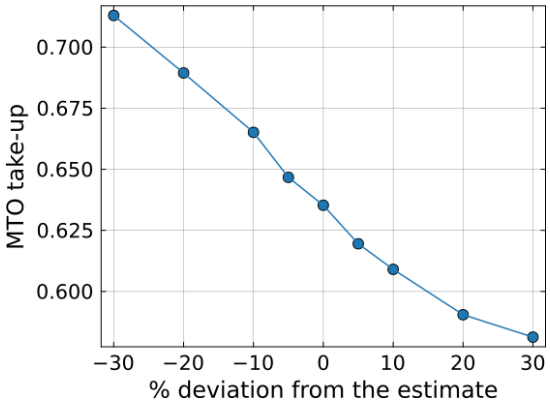
(i) θ and corr(value added, peer quality)



(j) σ_P and private enrollment after losing the lottery



(k) σ_S and within-nbhd. variance in school shares



(l) σ_N and MTO take-up

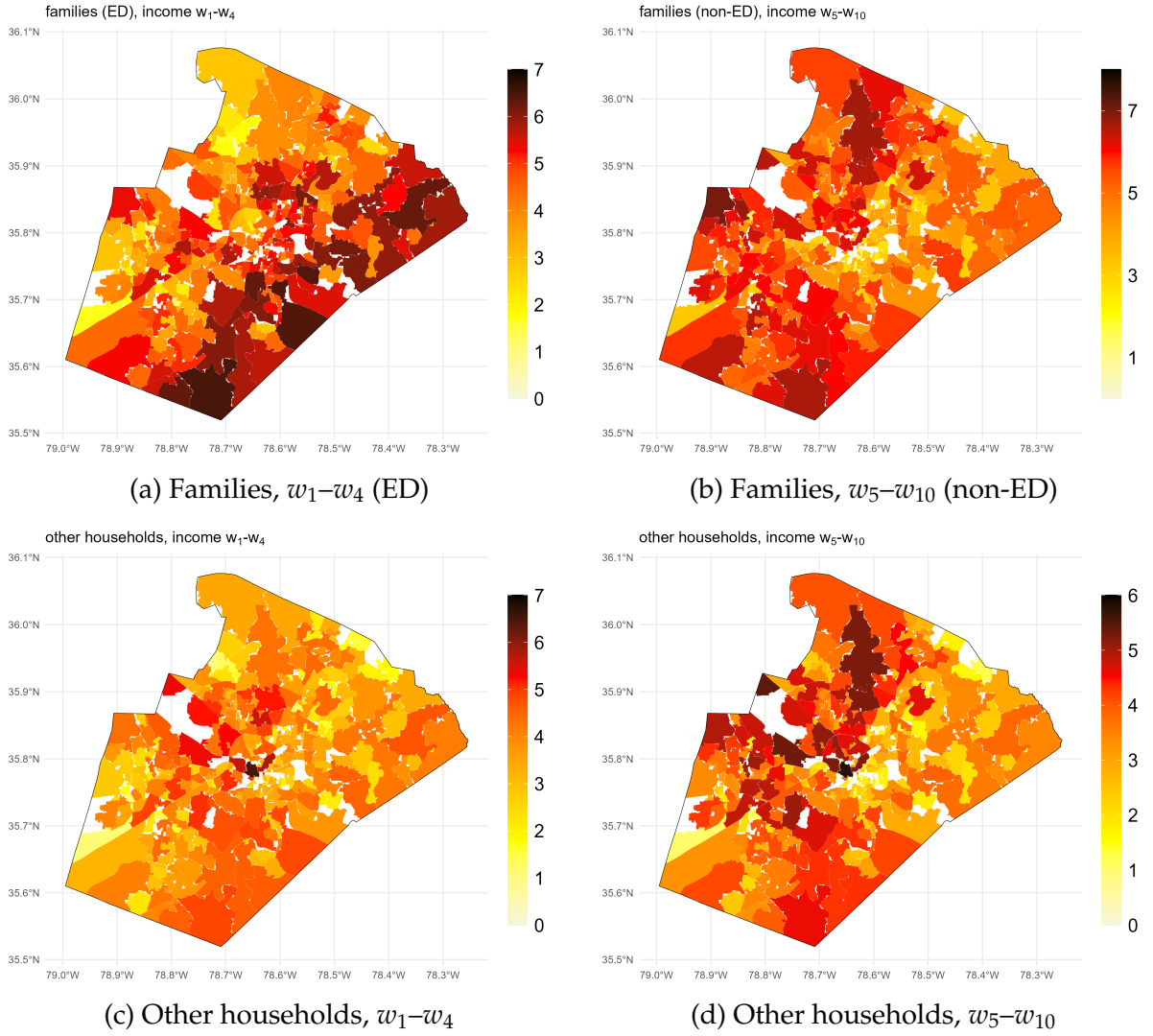
Horizontal axis: deviation of the parameter from its estimated value, in percent. Vertical axis: the moment that parameter informs, with all other parameters at their estimated values and the equilibrium re-solved. Panel (a) scales γ, κ, σ_S, σ_P and χ by a common factor. Panel (b) scales all $\gamma_{nonED,a}$ by a common factor. Panels (c) and (d) move one $\gamma_{w,a}$ at a time; each series is plotted against its own moment. In panel (j), the private-school valuations χ are re-estimated at each value of σ_P, so that private-school enrollment shares continue to match the data.

C.4 Additional Estimates

Figure C-2 shows the value of neighborhood amenities implied by our estimates of α_{0n}, α_{1n}, α^o_{0n}, α^o_{1n}. The top two maps of Figure C-2 show the value of neighborhood amenities for families with average income conditional on being ED (top left panel) or non-ED (top right panel). Consistent with Equation (2.6), the value of neighborhood n amenities for a family with income w is given by: $\alpha_{0n} + \alpha_{1n} \ln(w)$. The income levels w used to produce Figure C-2 are \$27,603 (top left panel) and \$116,299 (top right panel), which correspond to the average after-tax income in Wake County conditional on the family being ED and non-ED, respectively (averaged over 2013–2016). The bottom two maps in Figure C-2 are the analogous representations for other households. They show the value of neighborhood amenities for other households with average income conditional on income being within brackets w_1–w_4 (bottom left panel) or w_5–w_{10} (bottom right panel). The value of neighborhood n amenities for an other household with income w is given by: $\alpha^o_{0n} + \alpha^o_{1n} \ln(w)$. The income levels w used to produce Figure C-2 are \$24,508 (bottom left panel) and \$105,629 (bottom right panel), which correspond to the average after-tax income in Wake County conditional on household income being within brackets w_1–w_4 or w_5–w_{10}, respectively.

Figure C-3 shows estimates for the PUMA-level parameter χ for families as a function of their income.

Figure C-2: Estimated Value of Neighborhood Amenities by Household Type and Income



(a) Families, w_1–w_4 (ED)

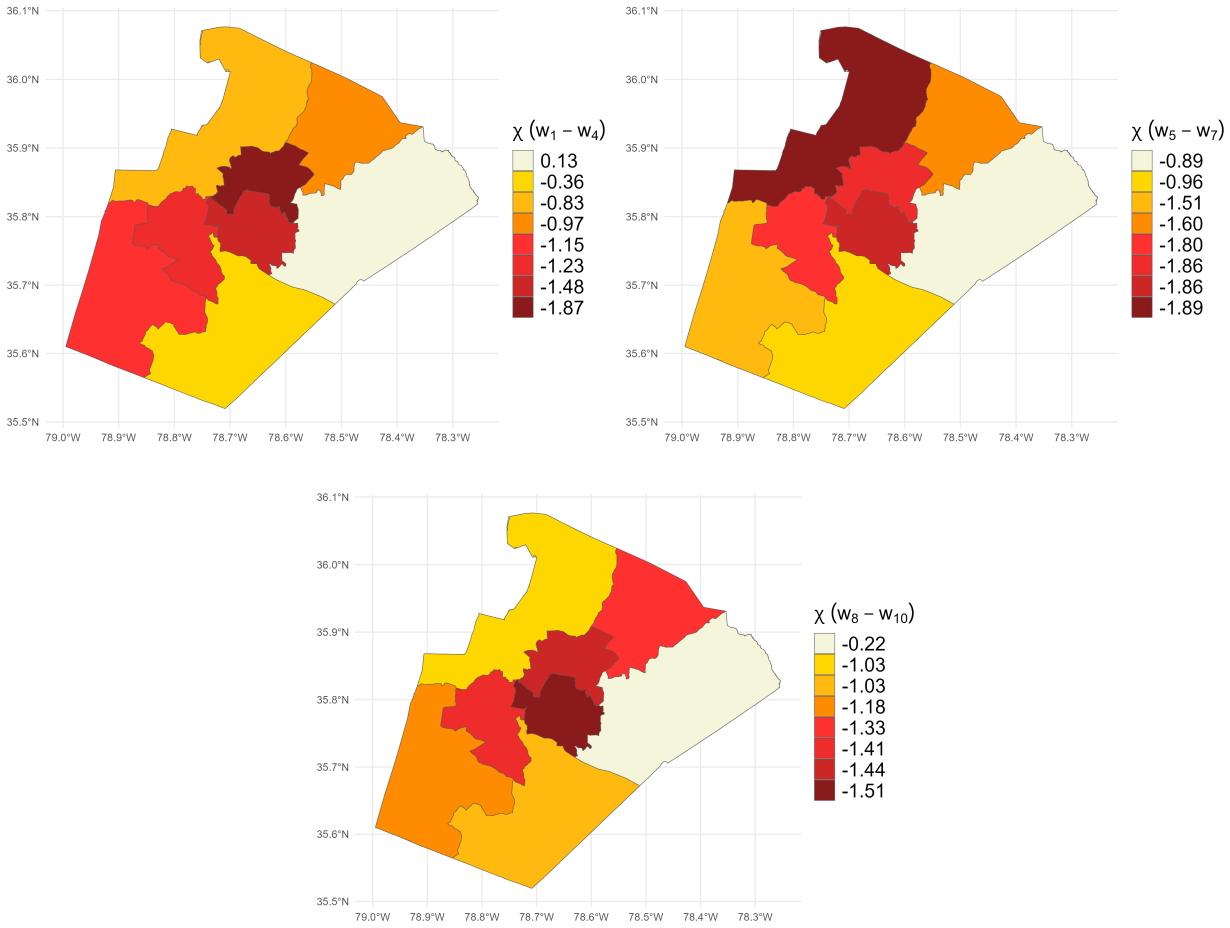
(b) Families, w_5–w_{10} (non-ED)

(c) Other households, w_1–w_4

(d) Other households, w_5–w_{10}

The maps on the top panels show estimated amenities for families; the maps on the bottom panels show estimated amenities for other households. The map on the top left (right) shows the neighborhood amenities valuation (in utils) for the average-income ED (non-ED) family implied by our estimates for α_{0n} and α_{1n}. The map on the bottom left (right) shows the neighborhood amenities valuation (in utils) for the average-income other household with income in brackets w_1–w_4 (w_5–w_{10}) implied by our estimates for α^o_{0n} and α^o_{1n}. The value of neighborhood n amenities for a family (other household) with income w is given by: $\alpha_{0n} + \alpha_{1n} \ln(w)$ ($\alpha^o_{0n} + \alpha^o_{1n} \ln(w)$). The income levels w used to produce the top-panel maps are \$27,603 (left panel) and \$116,299 (right panel), which correspond to the average after-tax income in Wake County over the time period of study conditional on the family being ED and non-ED, respectively. The income levels w used to produce the bottom-panel maps are \$24,508 (left panel) and \$105,629 (right panel), which correspond to the average after-tax income in Wake County over the time period of study conditional on household income being within the w_1–w_4 and w_5–w_{10} income brackets, respectively.

Figure C-3: Estimated Values of χ by Family Income



The map on the top left shows estimated values for the PUMA-level private-school parameters $\chi_{\text{PUMA};w}$ for ED families (income groups w_1–w_4); the maps on the top right and in the bottom show estimated values for the PUMA-level private-school parameters $\chi_{\text{PUMA};w}$ for non-ED families in income groups w_5–w_7 and w_8–w_{10}, respectively.

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:

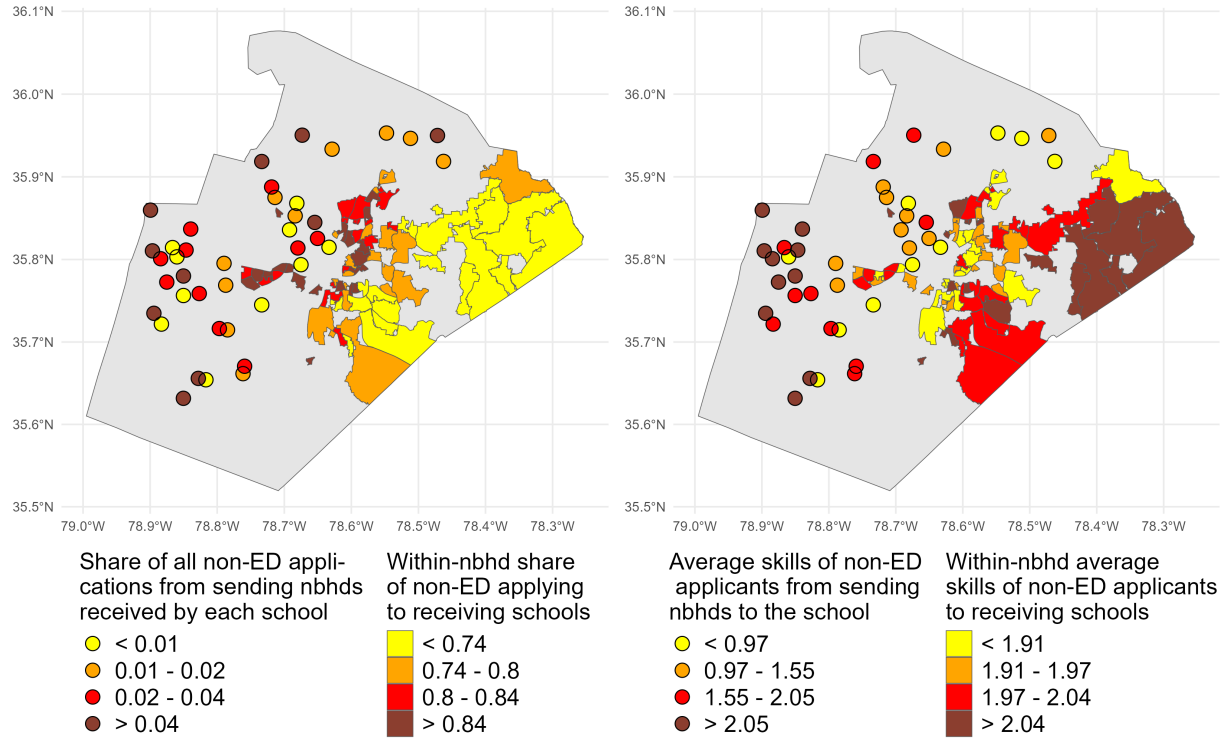
$$
\begin{aligned}
\underbrace{V_i}_{\text{total}} = & \underbrace{p_s\, u_{s;w} + (1-p_s)\, u_{\mathcal{O}_n;w}}_{\text{consumption}} + \underbrace{\gamma_{w,a}\left[p_s \ln(\bar{a}_s) + (1-p_s)\ln(\bar{a}_{\mathcal{O}_n})\right]}_{\text{peers}} + \\
& \underbrace{\gamma_{w,a}\theta\left(p_s A_s \mathbb{I}_{s\notin\mathcal{P}_n} + (1-p_s) A_{\mathcal{B}_n}\mathbb{I}_{\mathcal{O}_n=\mathcal{B}_n}\right)}_{\text{school value added}} + \\
& \underbrace{p_s\left(-\kappa_{s;w} d_{s,n}\right) + (1-p_s)\left(-\kappa_{\mathcal{O}_n;w} d_{\mathcal{O}_n,n}\right)}_{\text{commute}} + \\
& \underbrace{p_s\left(\sigma_S \varepsilon_{s;i}\right)\mathbb{I}_{s\notin\mathcal{P}_n} + (1-p_s)\left(\sigma_S \varepsilon_{\mathcal{B}_n;i}\right)\mathbb{I}_{\mathcal{O}_n=\mathcal{B}_n}}_{\text{unobs. pref. (public)}} + \\
& \underbrace{p_s\left(\chi_{\mathcal{P}_n;w} + \sigma_P \varepsilon_{\mathcal{P}_n;i}\right)\mathbb{I}_{s\in\mathcal{P}_n} + (1-p_s)\left(\chi_{\mathcal{P}_n;w} + \sigma_P \varepsilon_{\mathcal{P}_n;i}\right)\mathbb{I}_{\mathcal{O}_n=\mathcal{P}_n}}_{\text{unobs. pref. (private)}} + \\
& \underbrace{\alpha_{0n} + \ln(w)\left(\alpha_{1n} + \eta \ln(\bar{w}_n)\right) + \sigma_N \varepsilon_{n;i}}_{\text{nbhd. amenities}}
\end{aligned}
\tag{D-1}
$$

where (w, a) denotes agent i's exogenous type, and n and s denote the neighborhood they live in and the school they apply to, respectively. Recall that children who are not admitted to their option school of choice enroll in their preferred fallback alternative—either their base or their local private school—here denoted by $\mathcal{O}_n$. The consumption term $u_{s;w}$ equals $u_{n;w}$ for a public school and $u_{\mathcal{P}_n;w}$, which accounts for the tuition, for the private school (Section 2.3).

Under the specification with preferences over neighborhood peers, the component "nbhd. amenities" includes the term $\eta \ln(w)\ln(\bar{w}_n)$ of Equation (5.2). Under the specification with school fixed effects, a term $\lambda_g \psi_s$ enters the value of each public school and is weighted like school value added. The welfare component of school fixed effects is small throughout, and it is reported separately in the notes to Figures D-5 and D-6.

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: $\Delta V_i = V_i^{\text{counterfactual}} - V_i^{\text{baseline}}$. The change in each component of welfare is defined analogously.

Figure D-1: 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 by non-ED residents of sending neighborhoods. In the left map, receiving schools (dots) are colored as a function of the share of all applications from non-ED families in sending neighborhoods they receive; and sending neighborhoods are colored as a function of the share of their non-ED population who apply to receiving schools. In the right map, receiving schools (dots) are colored as a function of the average skill level of the non-ED applicants they receive from sending neighborhoods; and sending neighborhoods are colored as a function of the average skill level of their non-ED applicants to receiving schools.

D.2 Additional Results and Discussion on the School Choice Policy

D.2.1 Geography of non-ED Takers

Figure D-1 mirrors Figure 4 in Section 5.1, but focuses on non-ED takers. It exhibits several patterns similar to those observed for ED takers: sending neighborhoods located farther from receiving schools—particularly in the eastern part of the county—display lower take-up rates and stronger positive selection on skills, and receiving schools that are more distant from sending neighborhoods enroll systematically higher-skilled non-ED compliers.

A comparison of Figures 4 and D-1 highlights two key differences between ED and non-ED takers. First, take-up rates are higher among non-ED residents of sending neighborhoods, and non-ED takers are, on average, higher-skilled than ED takers. This is con-

sistent with willingness to pay for school quality increasing in family income and baseline skills (Table 1), as well as with higher baseline skills among non-ED children (Table A-3). Second, whereas distance shields more remote schools from ED takers, these same schools attract a larger share of non-ED takers. This pattern reflects higher commuting costs for ED families (Table 1) and their lower willingness to pay for school quality. Together, these mechanisms reinforce unequal exposure to the policy across receiving schools: schools in the central and north-central parts of the county primarily attract ED and lower-skilled takers, while schools in the western part of the county attract predominantly non-ED and higher-skilled takers.

D.2.2 Utilitarian Welfare and Alternative School Choice Policy Target

We show how to compute the welfare gains of a counterfactual policy using a general welfare criterion. We follow Dávila and Schaab (2025) and pose a social welfare function,

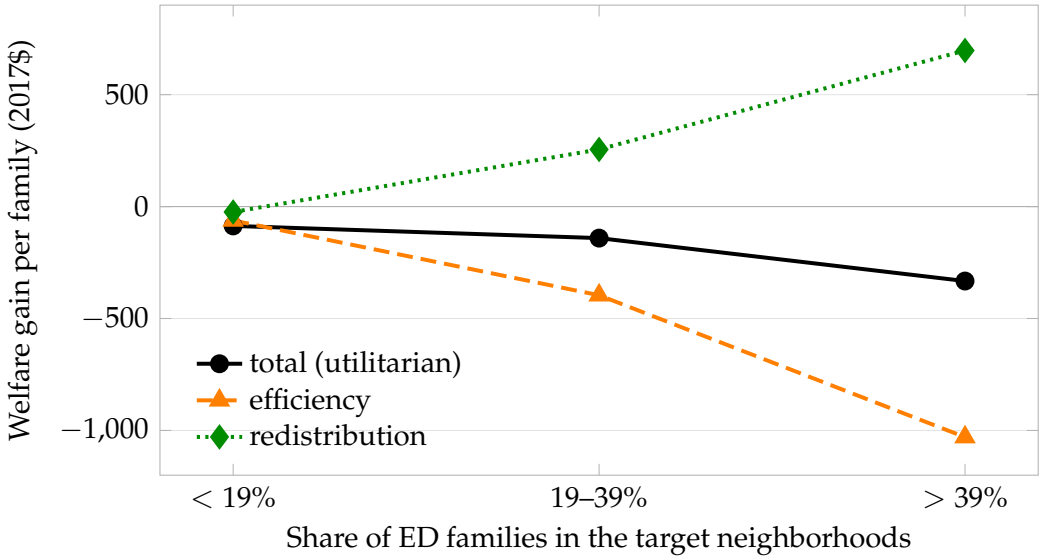
$$\widetilde{W} = \mathcal{W}(V_1, V_2, \ldots, V_I),$$

over the individual value V_i for agents $i = 1, 2, \ldots, I$. Let $\mathcal{MU}_i$ 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 $W = \widetilde{W} / \sum_i \phi_i \mathcal{MU}_i$ 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, $\omega_i \equiv \frac{\phi_i \mathcal{MU}_i}{\sum_j \phi_j \mathcal{MU}_j / I}$. Note that $\sum_i \omega_i / I = 1$. The normalized welfare impact ΔW of a given policy is then approximately given by:

$$\Delta W = \frac{1}{I} \sum_i \omega_i \frac{\Delta V_i}{\mathcal{MU}_i} = \underbrace{\frac{1}{I} \sum_i \frac{\Delta V_i}{\mathcal{MU}_i}}_{\text{efficiency}} + \underbrace{\mathcal{C}ov\left(\omega_i, \frac{\Delta V_i}{\mathcal{MU}_i}\right)}_{\text{redistribution}} \tag{D-2}$$

where ΔV_i is the change in value for agent i. Equation (D-2) shows how the aggregate welfare gain can be expressed as the average of all agents' dollar-equivalent gains, $\frac{\Delta V_i}{\mathcal{MU}_i}$, weighted by their relative marginal social value, ω_i. Simple algebra allows us 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 average 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 the extent to which the policy disproportionately benefits agents who have higher contributions to aggregate welfare, either because their marginal utility of income

Figure D-2: Welfare Effects of School Choice Expansion as a Function of the Target Group



This figure shows the welfare gains per family of the school choice expansion policy as well as the decomposition of welfare gains into efficiency and redistribution (based on Equation (D-2)) as we vary the target group.

$\mathcal{MU}_i$ is higher or because they have a higher weight in the social welfare function, ϕ_i.

In the analysis in Section 5, we adopt a money-metric welfare criterion which corresponds to the case in which $\phi_i = \frac{1}{\mathcal{MU}_i}$. 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 generates higher dollar-equivalent benefits for agents with a higher marginal utility of income.

Figure D-2 shows utilitarian welfare (black solid line), the money-metric component (orange long-dashed line) and the redistribution component (green dotted 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, which offers additional school choice to neighborhoods with a share of ED families above 39 percent. As reported in the main text, the money-metric component is negative, corresponding to a loss of about $1,030 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 as we target higher-income neighborhoods—which are also located closer to receiving schools. For those policies, the efficiency and redistribution components are less far apart, with smaller overall utilitarian welfare losses.

D.2.3 Additional Details on the Cost of the School Choice Policy

According to the North Carolina Department of Public Instruction, total expenditures for school transportation in the WCPSS amounted to $57,658,097 in the 2014-15 school year, for 75,511 students transported daily, that is, $764 per transported student per year ($801 in 2017 dollars). Transportation data are available only for all grades combined, so this figure averages over elementary, middle and high school students.[42] In the baseline equilibrium of the model, 84.1 percent of all children, in public and private schools, attend a school served by district transportation (their base school or an option school with transportation).

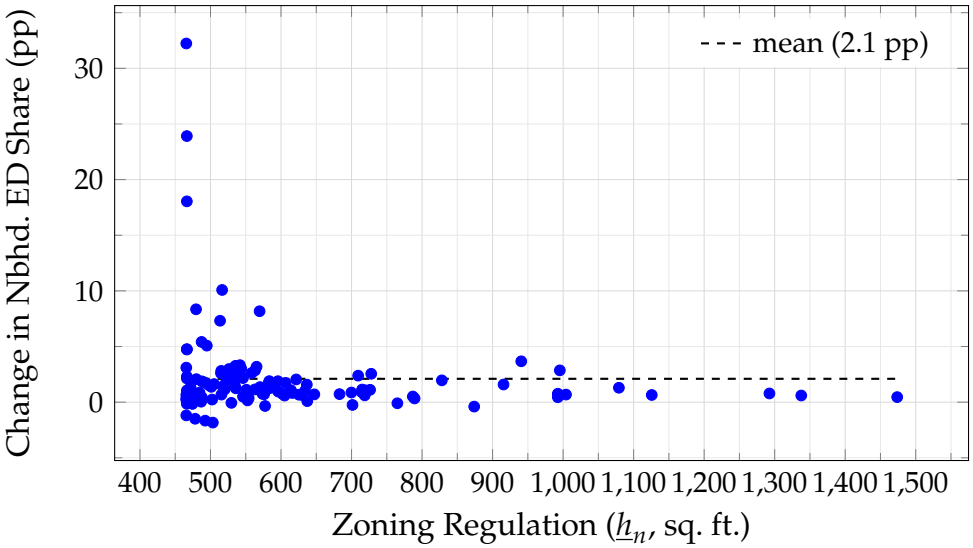
The school choice expansion increases transported miles per child, averaged over all children, by 19 percent. Since baseline transportation spending per child equals the expenditure per transported student times the transported share, the additional cost is $801 $\times$ 0.841 $\times$ 0.190 = $128 per child per year. This cost amounts to 0.13 percent of average family income.

D.3 Additional Results on the Voucher Policy

Figure D-3 plots, for each receiving neighborhood, the change in the share of resident ED families induced by the voucher policy against the neighborhood's level of zoning regulation. As discussed in Section 5.2, voucher recipients relocate predominantly to receiving neighborhoods with less stringent minimum lot-size requirements.

[42] Source: https://www.ncbussafety.org/documents/2014-15_nc_pupiltransportationdata.pdf, p. 3 (students transported and total miles) and p. 11 (expenditures). With 17,091,229 route miles, the report also implies a cost of $3.37 per bus-mile ($3.54 in 2017 dollars). Since transported students travel on average 2.92 miles to school in the baseline equilibrium of the model, twice a day, five days a week, 36 weeks per year, the $801 per student corresponds to $0.76 per child-mile, about one fifth of the cost per bus-mile.

Figure D-3: Zoning and Change in Neighborhood Composition under the Voucher Policy



The figure shows how receiving neighborhoods' shares of ED families change (in percentage points) as the voucher policy is implemented, as a function of their level of zoning regulation, $\underline{h}_n$ (blue dots). Zoning regulation is defined as the minimum housing size (in square feet) allowed in a neighborhood, obtained from minimum lot-size requirements (see Section 3.1). The dashed line represents the average change in shares of ED families in the receiving neighborhoods.

D.4 Cross-county Migration

Throughout the analysis in this paper, we have taken the distributions of family types, $\phi_{w,a}$, and of the other households' types, ϕ^o_w, as given in each year. We then performed our baseline counterfactual exercises keeping such distributions fixed at their values in a specific school year, 2015–16. In this section, we extend the model to incorporate the choice of a county of residence, effectively allowing for endogenous migration responses to our policy counterfactuals. We briefly introduce the equations describing such choice and the strategy we adopt to estimate the outside values and the migration elasticity. The counterfactual welfare outcomes in this modified environment are reported together with the other robustness checks in Section 5.3. We preview that introducing migration does not change our conclusions.

Theory. The economy is made of the 100 counties of North Carolina, indexed by c. Let $U_{c;w,a}$ be the expected utility of families of type (w,a) from living in county c, where the expectation is taken with respect to the realizations of idiosyncratic preference shocks over neighborhoods and schools. Let $U_{c;w} = \mathbb{E}_{\phi_{a|w}}[U_{c;w,a}]$ be the expected utility of families with income w, where the expectation is taken with respect to the conditional distribution of children's skills $\phi_{a|w}$. We assume that families choose their county of residence

before observing their children's skills. Ideally, we would like to observe the distribution of families across income *and* children's skills across all counties. This information is not easily available so we restrict families' county choice to depend on income only. As we discuss at the end of this section, relaxing this restriction is unlikely to affect our conclusions. Families with income w solve

$$\max_c U_{c;w} + \sigma_{C;w}\varepsilon_c,$$

where ε_c is an Extreme Value Type 1 idiosyncratic preference shock with income-specific dispersion parameter $\sigma_{C;w}$. The probability that a family with income w chooses county c is

$$\pi_{c;w} = \frac{\exp\left(\frac{U_{c;w}}{\sigma_{C;w}}\right)}{\sum_{\tilde{c}} \exp\left(\frac{U_{\tilde{c};w}}{\sigma_{C;w}}\right)}.$$

The problem of the rest of households is analogous: it is described by values $U^o_{c;w}$ and the dispersion parameter $\sigma^o_{C;w}$, and it pins down county choice probabilities $\pi^o_{c;w}$.

Data and Measurement. The choice set includes Wake County and the other 99 counties in North Carolina. To perform counterfactuals in the economy with migration, we need to recover the values $U_{c;w}$ and $U^o_{c;w}$ for all counties—excluding Wake County, whose values can be immediately obtained from our baseline equilibrium—and the values of $\sigma_{C;w}$ and $\sigma^o_{C;w}$.

Conditional on the value of $\sigma_{C;w}$, the vector $U_{c;w}$ is obtained from the empirical cross-sectional shares, $\pi_{c;w}$, available from the ACS. We are left to estimate the value of $\sigma_{C;w}$, which amounts to recovering the elasticity of county choice with respect to the value of residing in a given county. To do so, we simulate a five-percent increase in productivity (proxied by a five-percent increase in each income w) in Wake County only, and find the value of $\sigma_{C;w}$ that generates, in equilibrium, a ten-percent increase in the share of families with baseline income w that choose Wake County, $\Delta \ln \pi_{\text{Wake};w} = 0.1$. That is, we choose an implied migration elasticity of two. Since our baseline results rule out migration, we explore the robustness of our results using a migration elasticity above most estimates in the literature (Kennan and Walker, 2011; Monte et al., 2018; Notowidigdo, 2020). With the same strategy, we estimate $U^o_{c;w}$ and $\sigma^o_{C;w}$ for the rest of households. The dispersion parameters for idiosyncratic preference shocks over counties are reported in Table D-1.

Counterfactuals. The presence of migration makes the population in Wake County, $\phi'_{w,a}$,

Table D-1: Estimated Dispersion Parameters for Preference Shocks over Counties

Idiosync. County Preferences	$\sigma_{C;w_1}$	$\sigma_{C;w_2}$	$\sigma_{C;w_3}$	$\sigma_{C;w_4}$	$\sigma_{C;w_5}$	$\sigma_{C;w_6}$	$\sigma_{C;w_7}$	$\sigma_{C;w_8}$	$\sigma_{C;w_9}$	$\sigma_{C;w_{10}}$
Families	1.70	1.61	1.58	1.57	1.55	1.51	1.40	1.34	1.25	1.25
Rest of Households	2.01	1.64	1.59	1.55	1.53	1.50	1.47	1.45	1.40	1.40

The table shows estimates of the dispersion parameters of idiosyncratic preference shocks over counties, $\sigma_{C;w}$ for families and $\sigma^o_{C;w}$ for the rest of households, by income group.

endogenous and given by

$$\phi'_{w,a} = \phi_{w,a} \frac{\pi'_{\text{Wake};w}}{\pi_{\text{Wake};w}},$$

where

$$\pi'_{c;w} = \frac{\exp\left(\frac{U'_{c;w}}{\sigma_{C;w}}\right)}{\sum_{\tilde{c}} \exp\left(\frac{U'_{\tilde{c};w}}{\sigma_{C;w}}\right)},$$

and all the variables indexed by the prime superscript refer to post-policy outcomes. For simplicity, we assume that all values outside Wake County are unaffected by the policies, $U'_{c;w} = U_{c;w}$ for $c \neq$ Wake.[43] The analogous equilibrium conditions apply to ϕ'^{o}_{w}, $\pi'^{o}_{c;w}$, and $U'^{o}_{c;w}$. Section 5.3 in the main body of the paper describes the main results.

In this extension, due to data availability, we assumed that only income is observed at the time of county choice. We experimented with alternative formulations that allow county choice to vary by ability as well. One way to achieve that is to assume, for the sole purpose of measurement, that each county in North Carolina has the same relative shares of children with different abilities, conditional on income, as Wake County. Together with the migration elasticity adopted in this section, one can then recover ex-ante values by income and abilities in all counties in the choice set. The counterfactual analysis then allows families to observe both income and ability before choosing their county. We performed such an exercise and obtained similar but larger welfare losses, primarily because high-ability non-ED children are the most likely to leave the county.

[43] Formally, the overall change in value for a family with income w is given by the sum of three terms,

$$\sum_c \left[(U'_{c;w} - U_{c;w})\pi_{c;w} + U'_{c;w}(\pi'_{c;w} - \pi_{c;w}) + \sigma_{C;w}(\ln(\pi_{c;w})\pi_{c;w} - \ln(\pi'_{c;w})\pi'_{c;w}) \right].$$

Our school choice and voucher policies are enacted exclusively in Wake County, the only county that contributes to the first term. The second and third terms, due to the flows of households across counties, are very small relative to the first.

D.5 Results from Robustness Exercises

In this section we provide the complete set of results associated with the discussion in Section 5.3. Tables D-2 and D-3 are analogous to Table 1 in the paper, and report the parameter estimates of the two robustness exercises that require the re-estimation of the model: the specification with preferences over neighborhood income composition and the specification with school fixed effects. Figure D-4 maps the clusters of schools used in the latter. Clusters are obtained by k-means on school locations, taking the best of 10,000 runs of Lloyd's algorithm from k-means++ starting points. The number of clusters is close to the rule of thumb $\sqrt{n/2} \approx 7.5$ for the 111 public schools in our sample, and results using 6 and 10 clusters are similar.

Figures D-5 and D-6 show the robustness of the welfare effects of the school choice expansion and housing voucher policies. Table D-4 is analogous to Table 4 in the paper, and shows outcomes of receiving families under the school choice expansion policy across robustness exercises. Table D-5 is analogous to Table 5 in the paper, and shows outcomes of receiving families under the housing voucher policy across robustness exercises.

D.5.1 Parameter Estimates

Table D-2: Estimated Parameters: Preferences over Neighborhood Composition

School Peer Quality	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	0.75	0.72	0.73	0.80	0.94	1.03	1.05	1.02	1.18	1.61
Non-ED Families	0.47	0.69	0.77	0.94	1.01	1.10	1.22	1.44	1.77	2.57

Commuting Cost	κ_{NT}	κ_T
ED Families	0.14	0.06
Non-ED Families	0.09	0.04

School Value-Added	θ
All Families	0.07

Idiosyncratic Preferences	σ_S	σ_P	σ_N
All Families	0.27	0.67	0.93

Neighborhood Composition	η
All Families	0.04

The table reports the parameter estimates of the specification with preferences over neighborhood income composition, Equation (5.2). The parameter η is set, not estimated; all other parameters are re-estimated targeting the same moments as in the baseline specification.

Table D-3: Estimated Parameters: School Fixed Effects

School Peer Quality	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	0.82	0.79	0.80	0.88	1.04	1.14	1.16	1.13	1.31	1.75
Non-ED Families	0.57	0.83	0.92	1.12	1.19	1.30	1.43	1.67	2.02	2.85

Commuting Cost	κ_{NT}	κ_T
ED Families	0.16	0.07
Non-ED Families	0.11	0.05

School Value-Added	θ
All Families	0.07

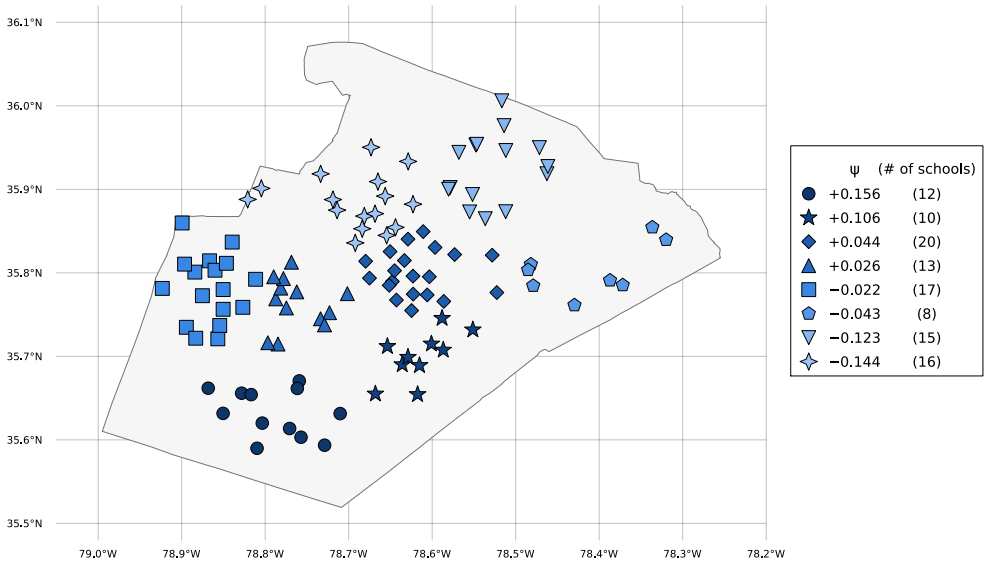
Idiosyncratic Preferences	σ_S	σ_P	σ_N
All Families	0.31	0.75	0.93

School FE Loading	$\lambda_{\text{non-ED}}$
Non-ED Families	2.36

School Fixed Effects	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6	ψ_7	ψ_8
All Families	−0.14	−0.12	−0.04	−0.02	0.03	0.04	0.11	0.16

The table reports the parameter estimates of the specification with school fixed effects $\lambda_g \psi_s$ (Section 5.3). ψ_k is the common value of ψ_s for the schools in cluster k of the 8 clusters of geographically proximate schools, sorted from lowest to highest and normalized to have mean zero across clusters; $\lambda_{\text{non-ED}}$ is their loading for non-ED families, with λ_{ED} normalized to 1. All other parameters are re-estimated targeting the same moments as in the baseline specification.

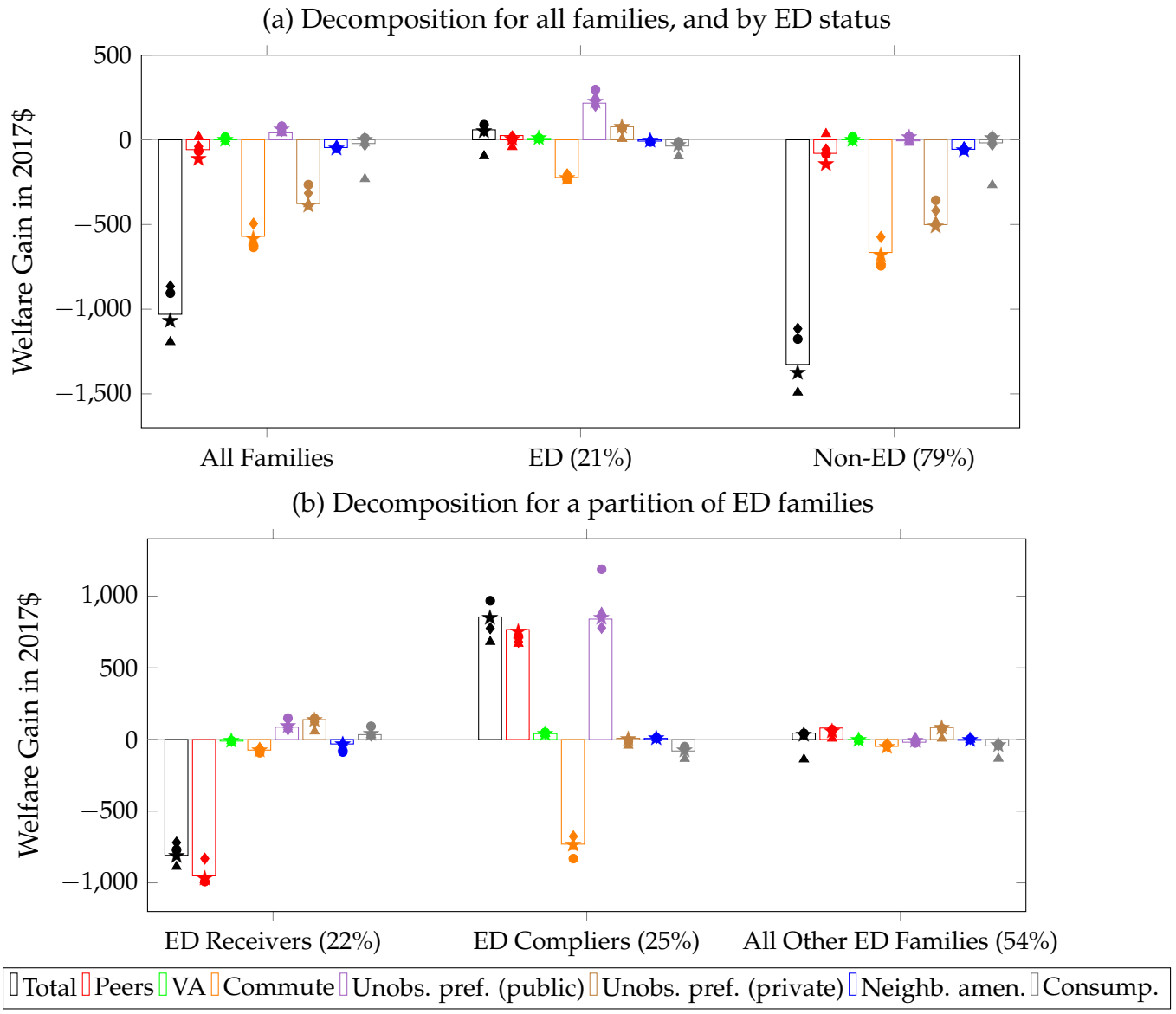
Figure D-4: Geographic School Clusters



The figure shows the 8 clusters of geographically proximate public schools in the specification with school fixed effects (Table D-3). Clusters are obtained by k-means on school locations. Schools in the same cluster share the same marker and color, and the same value of ψ_s.

D.5.2 School Choice Expansion Policy

Figure D-5: Welfare Effects of the School Choice Policy: Robustness



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). The hollow bars depict the baseline values in the paper. The ▲ symbol indicates results under flexible private-school prices P_n; the ★ symbol indicates results when allowing for cross-county migration; the ⧫ symbol indicates results when preferences for neighborhood peers are set to $\eta = 0.04$; and the • symbol indicates results under the specification with public-school fixed effects (ψ_s) (with the welfare component due to the school fixed effects omitted and equal to \$5 for all families, −\$26 for ED families, and \$13 for non-ED families).

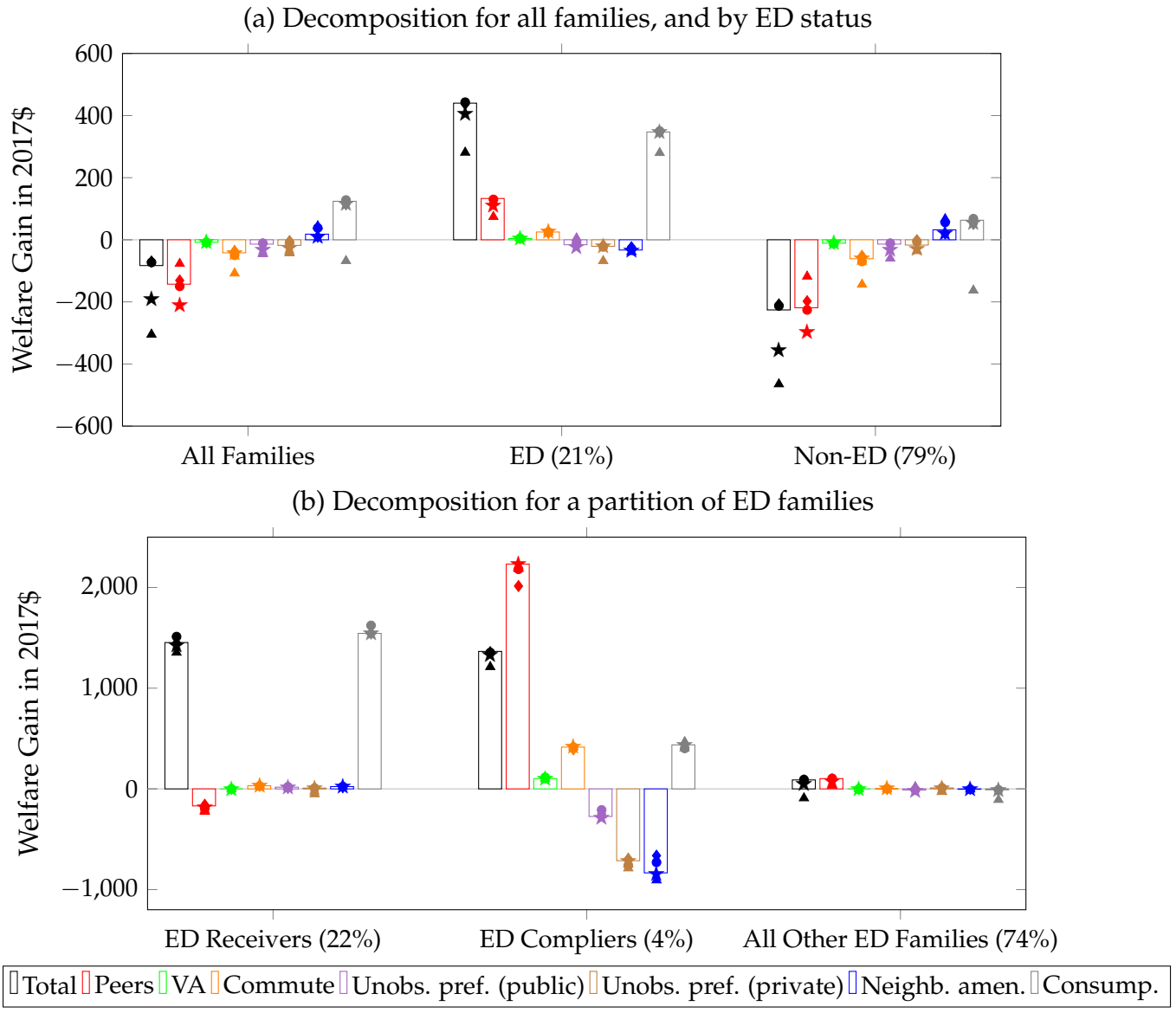
Table D-4: Outcomes of the School Choice Policy for Receiving Families: Robustness

	ED Receivers	Non-ED Receivers
Flexible private-school prices P_n		
Change in average school peers (% std)	−41.3	−13.9
Change in share attending private school (pp)	0.9	4.6
Share locating outside receiving nbhds (%)	12.5	8.7
Cross-county migration		
Change in average school peers (% std)	−40.8	−13.8
Change in share attending private school (pp)	1.9	4.9
Share locating outside receiving nbhds (%)	13.4	8.6
Preferences over neighborhood composition		
Change in average school peers (% std)	−38.3	−12.3
Change in share attending private school (pp)	1.7	4.6
Share locating outside receiving nbhds (%)	12.0	7.8
Public-school fixed effects		
Change in average school peers (% std)	−40.0	−13.0
Change in share attending private school (pp)	1.6	4.4
Share locating outside receiving nbhds (%)	13.0	8.6

The table shows, across robustness exercises, 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.

D.5.3 Housing Voucher Policy

Figure D-6: Welfare Effects of the Voucher Policy: Robustness



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). The hollow bars depict the baseline values in the paper. The ▲ symbol indicates results under flexible private-school prices P_n; the ★ symbol indicates results when allowing for cross-county migration; the ♦ symbol indicates results when preferences for neighborhood peers are set to $\eta = 0.04$; and the • symbol indicates results under the specification with public-school fixed effects (ψ_s) (with the welfare component due to the school fixed effects omitted and equal to −\$10 for all families, −\$3 for ED families, and −\$12 for non-ED families).

Table D-5: Outcomes of the Voucher Policy for Receiving Families: Robustness

	ED Receivers	Non-ED Receivers
Flexible private-school prices P_n		
Change in average school peers (% std)	−11.4	−9.5
Change in share attending private school (pp)	−0.5	1.1
Share locating outside receiving nbhds (%)	0.2	2.7
Cross-county migration		
Change in average school peers (% std)	−9.3	−8.0
Change in share attending private school (pp)	0.1	1.1
Share locating outside receiving nbhds (%)	0.3	2.1
Preferences over neighborhood composition		
Change in average school peers (% std)	−8.7	−7.7
Change in share attending private school (pp)	0.1	1.0
Share locating outside receiving nbhds (%)	0.3	2.2
Public-school fixed effects		
Change in average school peers (% std)	−9.4	−7.9
Change in share attending private school (pp)	0.1	1.0
Share locating outside receiving nbhds (%)	0.3	2.5

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

D.6 Equilibrium Multiplicity

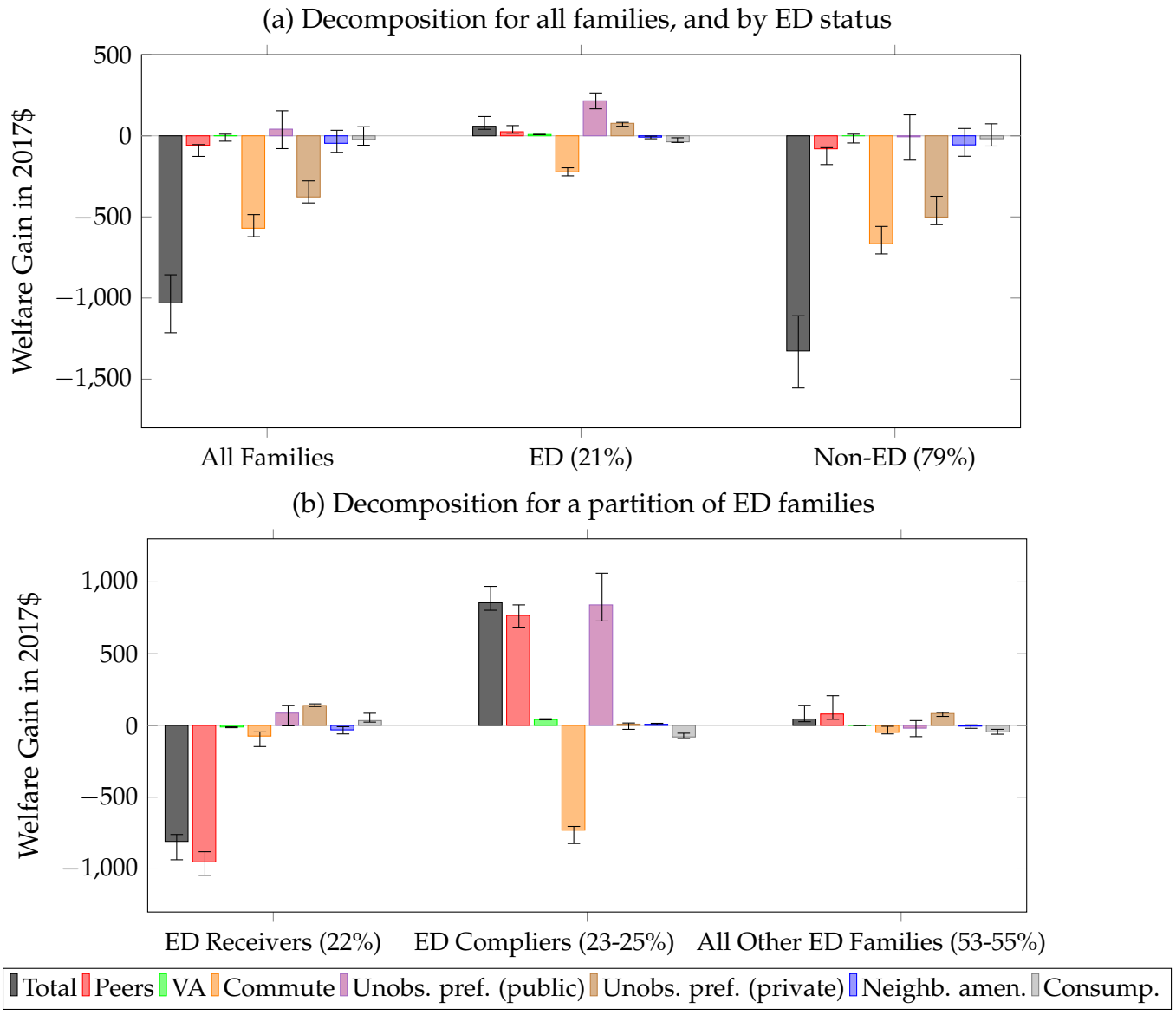
To solve for the counterfactual equilibrium under each policy, we start from a guess of prices, school peer composition, admission probabilities, and the transfer rate. We compute individual choices and then update those equilibrium variables until convergence. In Sections 5.1 and 5.2, our initial guess is the value of those variables in the baseline equilibrium.

We also perform 100 runs of each policy that only differ with respect to the initial guess. In each run, we draw the guess as a random permutation of the corresponding object in the baseline economy. For example, the guess for house prices is a random permutation of baseline prices across neighborhoods. The exercise is deliberately conservative: each run starts from an allocation that bears no connection to the baseline, hence it can reach equilibria that our spatially targeted policies would be unlikely to settle on.

We find equilibrium outcomes to be remarkably stable. For each school, we compute the standard deviation of log average peer quality across runs. We find an enrollment-weighted mean and median of six and one percent in the school choice (eight and two percent in the voucher) counterfactual. For house prices, the means and medians are never above 0.2 percent. We also report the entire range of changes in welfare and its

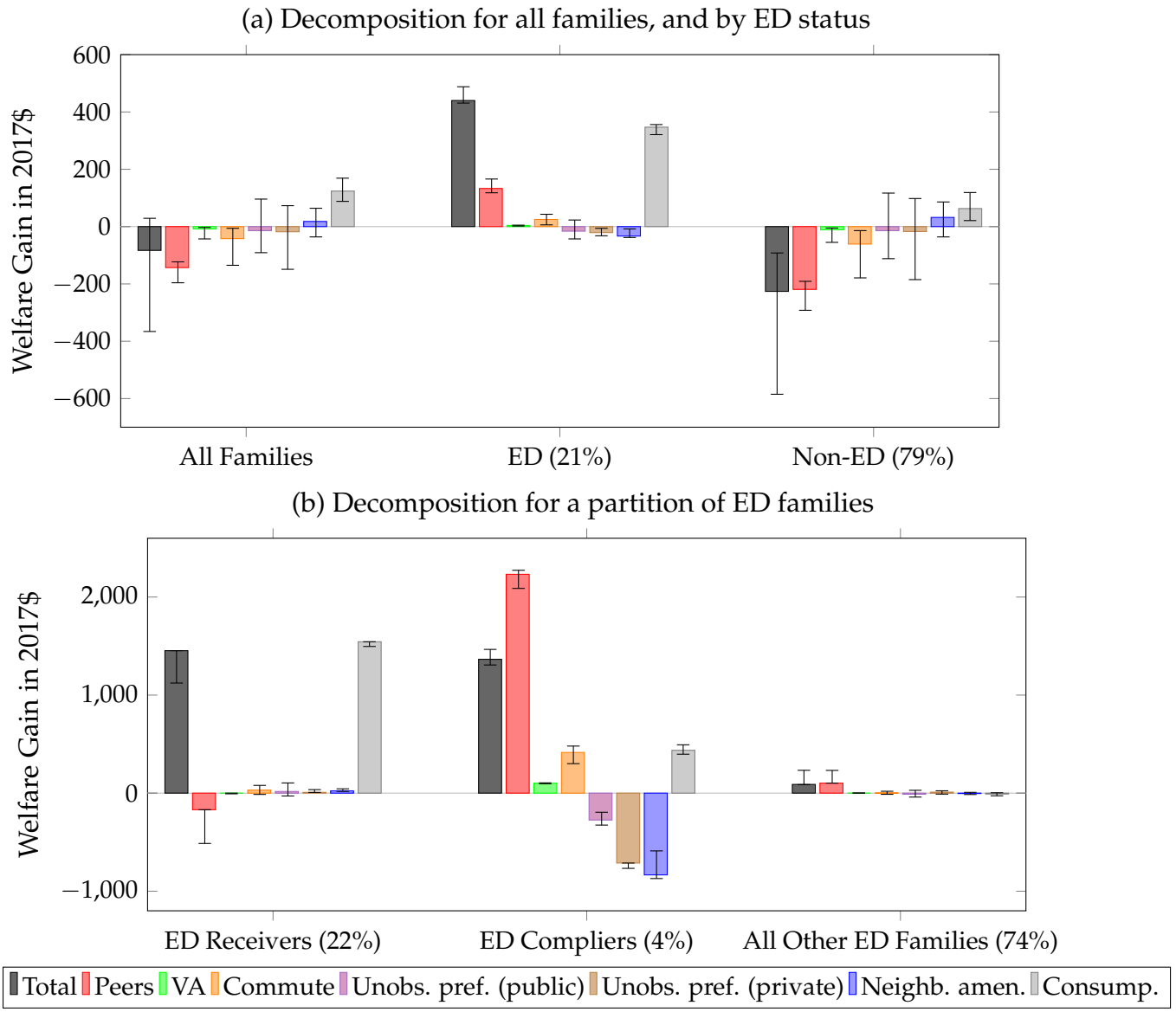
components across all 100 runs in Figures D-7 and D-8, together with our baseline results. Our main conclusions are effectively unchanged across all equilibria.

Figure D-7: 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 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 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). The formal decomposition of welfare into its components is shown in Equation (D-1).

Figure D-8: Decomposition of Welfare Effects of the Voucher Policy Across Equilibria



The figure is analogous to Figure 5 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). The formal decomposition of welfare into its components is shown in Equation (D-1).