People- or Place-Based Policies to Tackle Disadvantage? Evidence from Matched Family-School-Neighborhood Data*

Lucienne N.Y. Disch†
University of Pennsylvania

JOB MARKET PAPER

[click here for latest version]

November 28, 2024

Abstract

Many studies examine the impact of families, schools, and neighborhoods on children's outcomes, but there is little comprehensive quantification of their interactive effects. This paper investigates how family-school-neighborhood combinations influence human capital accumulation, as measured by test score gains. I propose a framework that accounts for family sorting into neighborhoods and schools, as well as potential nonlinear interactions among the three factors. To do so, I build on Bonhomme, Lamadon, and Manresa (2019), extending their clustering approach to an educational setting. I estimate the model using matched family-school-neighborhood data from North Carolina and decompose the distribution of test score gains into match-specific sources. The institutional setting, where multiple residential areas are assigned to the same school and multiple schools serve the same area, allows me to disentangle neighborhood effects from school value-added. My identification strategy leverages test score variation from children who move and/or change schools. The empirical findings highlight the crucial role of the family and reveal significant positive complementarities that are especially pronounced in environments with relatively high test score gains, particularly benefiting children at the lower end of the test score distribution. I examine the potential impacts of two potential educational policies on test scores – improving school quality and reallocating children across different schools and neighborhoods – and find that neighborhood effects outweigh school effects, as school value-added is shaped by location.

*I am indebted to Francesco Agostinelli, Jere Behrman, Margaux Luflade, and Petra Todd for many fruitful discussions and suggestions. For very helpful comments, I thank Stephane Bonhomme, Aline Bütikofer, Ana Costa-Ramón, Olivier De Groote, Chao Fu, Joseph Huang, Rasmus Lentz, Marko Mlikota, Remo Nyffenegger, Chris Taber, Matt Wiswall, and seminar and conference participants at the Institute of Labor Economics (IZA), the JHU-Penn-NYU Applied Micro conference, the Dynamic Structural Econometrics conference of the Econometric Society (Dynamic Structural Models: Policy Evaluation and Heterogeneity Measurement), the Global Labor Organization, the SEA Meeting, the Swiss Society of Economics and Statistics, and the University of Pennsylvania. All errors are my own.

†Department of Economics, University of Pennsylvania. Email: brunnerl@sas.upenn.edu.

1 Introduction

The ongoing debate over whether and what kind of people- or place-based approaches are most effective in helping children from disadvantaged families underscores the complexity of quantifying the factors that impact children's outcomes.[1] The question of whether support should target distressed individuals or specific locations is challenging, because socioeconomic disadvantage is often spatially concentrated. Designing effective targeted policies requires understanding the sources of inequality in outcomes, such as test scores, which are highly predictive of future (labor market) success. Yet, much remains unknown about the factors that contribute to human capital formation. What are the contributions of family, school, and neighborhood to a child's educational outcomes? To what extent are spatial disparities in outcomes driven by families sorting into specific neighborhoods and schools versus the direct causal influence of these environments? Is there heterogeneity in neighborhood effects and school value-added, and are there important complementarities?

To answer these questions, I develop an empirical framework à la Bonhomme, Lamadon, and Manresa (2019) for my matched family-school-neighborhood network structure that captures three-sided unobserved heterogeneity and includes test score gains as the outcome. The framework accounts for nonrandom family sorting across neighborhoods and schools and allows for unrestricted interactions among families, schools, and neighborhoods. Assuming discrete heterogeneity, the approach enables identification of the distributions of test score gains for children from various types of families in various types of neighborhoods and schools. These match-specific distributions provide insights into the nature and extent of interactive effects among family, school, and neighborhood contexts on children's test score gains, as well as the heterogeneity in neighborhood treatment effects and school value-added: the same type of neighborhood or school could impact children differently, and the same type of child could experience different test score gains depending on where they live and go to school. Identification of these match-specific treatment effects comes from the variation induced when children from different types of families change schools and/or move to a new place of residence. The idea is to focus on children who undergo the same type of transition and probabilistically assign each child to a latent subpopulation (family type) based on their test score gains and movement patterns. By doing this for all possible transitions in the fully connected network, this approach uncovers match-specific distributions of test score gains, providing insights into the determinants of variation in test scores. Children from the same family type undergoing the same type of transition are assumed to share identical preferences and face the same constraints.

[1]This debate goes back to Winnick (1966)'s seminal work.

I estimate the model using data from North Carolina, USA. The institutional context is such that some neighborhoods are served by several schools, and some schools draw students from multiple neighborhoods, leading to various combinations of neighborhood-school pairings. This setting allows me to disentangle neighborhood effects from school effects. North Carolina administers statewide standardized tests in mathematics and reading to all children in grades $3-8$. Using administrative microdata from the North Carolina Education Research Data Center (NCERDC), I construct a panel with three observation periods for each child attending public school, capturing key information such as test scores, school identifiers, and the census block groups (neighborhoods) where they reside.

The estimation procedure builds on Bonhomme, Lamadon, and Manresa (2019)'s matched worker-firm two-step estimator, extending it in two essential ways to suit my educational context. First, I incorporate a third dimension to account for family-school-neighborhood effects, capturing the interplay across these environments. Second, I focus on test score gains rather than levels as the measurement variable, requiring an additional period for each child compared to their static model. This approach mitigates the initial condition problem in test score analysis, allowing a more precise understanding of the role of neighborhood context and school value-added for human capital accumulation. Both the approach of adding a third layer and the focus on changes rather than levels can be applied to other settings where an additional dimension, such as geography, is relevant for worker-firm or other type-specific relationships among economic agents, or where analyzing growth in the outcome variable is essential.

I estimate the model in three steps. First, using three years of consecutive test score levels, I calculate two test score gains for each child, with each isolating the current value-added net of prior performance. In the second step, I use weighted distributional k-means algorithms to separately classify neighborhoods and schools into types, effectively reducing the dimensionality of their heterogeneity. This classification is based on the distribution of test score gains among stayers in the first period. For tractability, I then group all possible permutations of neighborhood and school type pairs into distinct *places*. In the third step, I estimate the actual model by maximum likelihood, conditional on these places. This last step mirrors Bonhomme, Lamadon, and Manresa (2019)'s second estimation step, as the frameworks align in that I model latent family types as discrete random effects correlated with the discrete place "fixed effects". This setup allows me to capture unobserved family differences specific to each place, without assuming uniformity across all families, thereby allowing for potential heterogeneity in family-place interactions.

Model estimates shed light on key determinants of variation in test scores. The parameter estimates from the pre-clustering step, based on three types of neighborhoods and schools, reveal that median test score gains are $17 - 19$ percentage points (or $40 - 46\%$) higher in the top-performing neighborhoods and schools compared to the lowest-performing ones. In low-performing neighborhoods, 40% of children attend low-performing schools, while only 21% have access to high-performing schools. Conversely, in high-performing neighborhoods, only 14% of children are in low-performing schools, whereas 47% attend high-performing ones. Results indicate a clear positive effect on school performance from attending a higher-quality school (measured by mean test score gains), even when controlling for neighborhood type. This highlights the significant role of school quality in shaping outcomes. Although, neighborhood and school contexts appear to interact, within schools of the same performance level, mean test score gains increase monotonically with neighborhood performance level. This suggests that neighborhood context can amplify or limit school effectiveness, as schools of a given type produce different outcomes depending on the performance level of the neighborhood in which their students are located.

Estimation results from the third and core step, which further distinguishes between family types, reveal match-specific distributions of test score gains. Based on three types of families, I find limited sorting based on gains, but strong evidence of family heterogeneity, as well as significant heterogeneity in neighborhood treatment effects and school value-added. Results show that family influence is the most significant driver of test score gains. I also find that the same neighborhood and school can have different impacts on different types of children: conditional on neighborhood type or school setting, mean test score gains for high-achieving children are roughly 10% higher compared to their peers with lower gains. Additionally, I find that children benefit most when they live in neighborhoods and attend schools where peers achieve relatively high test score gains. Among families of the same type, the difference in their children's average test score gains between the bottom and top environments is approximately 5%. Particularly children at the lower end of the test score distribution benefit most from attending schools where peers have relatively high average test score gains, allowing them to substantially catch up to higher performance levels and offset some of the disadvantages associated with family background. This finding underscores an important complementarity between children at the lower end of the test score distribution and schools with high test score gains. Sensitivity checks, including leveraging children changing schools after exogenous redrawing of school boundaries – that I identify using geospatial maps – confirm the robustness of these results.

In terms of mobility, results indicate that children with relatively low test score gains generally experience test score losses in the year following a transition, particularly when moving from a high-gain to a low-gain environment (-50%). In contrast, children with relatively high test score gains tend to experience losses only when transitioning from a high-gain to a low-gain environment. Otherwise they show gains, with the largest improvements occurring when they transition from a low-gain to a high-gain environment (+20%).

I use the estimated model to analyze two types of interventions: a place-based policy targeting the improvement of school quality for schools with relatively low test score gains and a child-focused reallocation across schools and neighborhoods. I simulate the first policy holding neighborhoods constant and imposing the distribution of test score gains from high-gain schools onto low-gain schools. Although all children would benefit from improved school quality – on average achieving approximately 0.33% higher gains – those at the lower end of the test score distribution would experience the greatest improvements. However, since the policy benefits the entire distribution, it only modestly reduces the overall dispersion of test score gains. I then analyze counterfactual scenarios to quantify the extent of sorting and interaction effects among families, schools, and neighborhoods, assessing how these factors contribute to disparities in educational outcomes, particularly when access to high-quality schools is restricted to specific neighborhoods. Simulations of random reallocation of children across neighborhoods and schools reveal that neighborhood context plays a pivotal role by shaping school value-added – either amplifying or constraining schools' effectiveness. While neighborhood reallocation shifts the overall environmental quality, amplifying the combined impact of both settings, school-only reallocation lacks a transformative effect due to the absence of complementary neighborhood diversity.

Overall, the findings highlight the critical role of family in academic achievement, suggesting that targeted, person-centered policies could be an effective strategy for improvement. Such interventions might include tutoring programs for children at the lower end of the test score distribution or housing mobility programs that provide vouchers and support to help disadvantaged families move to better neighborhoods. Place-based policies aimed at improving neighborhood environments could also significantly impact educational outcomes, offering a complementary approach to fostering equity and performance.

Related literature and contribution. This paper is related to an extensive literature studying the effects of neighborhoods on children's outcomes (Bergman et al. (2024); Chetty and Hendren (2018a); Chyn (2018); Chetty et al. (2016); Ludwig et al. (2013), among many others).[2] A key insight in this literature is that it remains uncertain to what extent children's outcomes are shaped by their family environments compared

[2] Chyn and Katz (2021) provide a more detailed overview of the literature on neighborhood effects. Graham (2018) and Durlauf (2004) review the literature from a more methodological and technical point of view.

to their schools and residential areas, which makes it difficult to infer reasonable policy implications (see, e.g., Chetty and Hendren (2018b); Chyn and Katz (2021)).[3] Therefore, it is important to account for the fundamental drivers of variations in children's outcomes. In the last years, the literature has made substantial progress in explaining why and how outcomes, such as academic performance, differ among children by zeroing in on the family[4], the school[5], the neighborhood[6], or a combination of two of these drivers.[7]

A difficult challenge is that neighborhoods are background contexts against which families affect children. Families not only have direct influence on children, but also indirect by choosing the neighborhood they live in, which in turn, determines children's peers from schools and communities with influential social impact. Plus, because schools are often tightly linked to the place of residence with limited or no school choice, it is typically not feasible to separate the school from the neighborhood effect. Thus, academic performance is the result of many different blurred and nested dimensions, that existing identification strategies struggle to disentangle.

This paper contributes to this literature by proposing a new strategy that can shed light on the factors contributing to human capital production by considering family-, school-, and neighborhood-effects in a unified framework. Whenever a family or their child interacts with a school and a neighborhood, family-school-neighborhood matches (or networks) arise. A key feature of such networks is the presence of unobserved heterogeneity (Bonhomme, 2020). From a methodological perspective, my approach to study the sources of variation in test scores is closely related to the identification and estimation of models with latent heterogeneity.[8] There exists a growing empirical literature going back to the seminal work of Abowd et al. (1999), who combine matched worker-firm data with a two-way fixed effects model. Several papers extend this approach to investigate teacher or school value-added (e.g., Chetty et al. (2014)), neighborhood effects (e.g., Chetty and Hendren (2018a); Card et al. (2023)), or sorting

[3]This debate goes back to Winnick (1966)'s seminal work. For a more recent overview, see McCann (2023), for example. A difficult problem in this context is a comprehensive quantification, allowing to identify the separate and match-specific contributions of family, school, and neighborhood to variations in test scores, which is related to Manski (1993)'s reflection problem.

[4]E.g., Cunha and Heckman (2007); Cunha and Heckman (2008); Doepke et al. (2019); Agostinelli and Wiswall (2024), among many others.

[5]E.g., Angrist et al. (2024); Shan and Zölitz (2024); Goldsteyn et al. (2021); Hanushek (2005); Rivkin et al. (2005), among many others.

[6]E.g., Agostinelli et al. (2024); Eckert and Kleineberg (2024); Chyn and Daruich (2021); Chetty et al. (2016); Chyn (2018), among many others.

[7]E.g., Todd and Wolpin (2003); Laliberté (2021); Chetty and Hendren (2018b), Abdulkadiroğlu et al. (2020), among many others.

[8]For example, see Arcidiacono and Jones (2003) or Arcidiacono and Miller (2011) for early examples adapting the expectation–maximization algorithm to integrate unobserved heterogeneity into conditional probability estimators.

of children across schools (Kramarz et al., 2015).

Performance of fixed effects estimators depends on the number of existing connections between the agents in the network – in sparsely connected networks, fixed effects estimators are often susceptible to the incidental parameter bias (Andrews et al. (2008); Jochmans and Weidner (2019)). While from a neighborhood's and school's perspective, there are many links to children, so their fixed effects can be reliably estimated, child fixed effects, however, would be estimated based on few observations only. Thus, I condition on the neighborhood and school side of the network – treating them as fixed effects – and then treat family heterogeneity differently by adopting a one-sided correlated random effects approach à la Bonhomme et al. (2019).[9] This correlated random effects specification accounts for the fact that the effect of a neighborhood or school on the child could vary from family to family due to differences (heterogeneity) among families, and preserves the tractability of single-agent models.

This procedure further allows me to take the family-school-neighborhood network as given. Network exogeneity makes family-school-neighborhood connections independent of potential test score outcomes, *conditional* on unobserved family heterogeneity – and schools and neighborhoods that then act as observed covariates (Bonhomme, 2020).[10]

Another advantage of Bonhomme et al. (2019)'s framework with no functional form assumptions is that it can account for potential nonlinearities (Durlauf (2004); Brock and Durlauf (2001)).[11] The key idea of this paper is to extend Bonhomme et al. (2019)'s approach to the three-way matched family-school-neighborhood network structure and test score production function. However, while this approach allows unrestricted interaction effects between family, school, and neighborhood, it comes at the cost of requiring point identification, that is, one needs to observe transitions of all children from and to all neighborhood-school combinations possible ("connecting cycles" or "graph connectivity"). Since there is limited upward mobility in the data, and sparsely connected networks can lead to weak connectivity (Bonhomme, 2020), I reduce the dimension of all three sides and proceed with an estimation approach based on discrete unobserved heterogeneity. Because some families are very similar, and so are some schools and some neighborhoods, I do not need to distinguish between highly similar entities and

[9] This approach is related to mixed membership models (Airoldi et al. (2008); Blei et al. (2003)).

[10] A correlated random effects specification (correlated with observables) on all three sides would also make sense, but would require to explicitly model network formation, which is substantially more challenging with unobserved heterogeneity (Bonhomme, 2020). Bonhomme et al. (2019) adopt a two-sided random effects approach in their supplement. However, since endogenizing link formation requires modeling the joint distribution of potential outcomes and links, it often does not work very well due to numerical challenges. Specifically, with two or more random effects, the likelihood becomes an intractable sum over all these terms. While variational approximations may offer a solution, this remains beyond the current research frontier (Bonhomme, 2020).

[11] The important role of complementarities and its relation to sorting goes back to Becker (1973)'s theory of marriage, where matching is positively assortative if types are complements.

instead cluster them into groups. This way, two families of the same type who live in the same type of neighborhood, and send their child to the same type of school, are assumed to have identical preferences and constraints.

An important restriction in Bonhomme et al. (2019)'s two-step estimator for earnings is that the geographical location of workers and firms does not play a role. Lentz et al. (2023) extend their method to further allow for different drivers of sorting, and Mann (2024) builds on it to account for potential spatial differences. Mann (2024)'s algorithm requires multiple periods of longitudinal data over a worker's lifetime and a sufficiently large number of jobs per worker as well as a large number of job matches per firm. But, in an educational context, where location can play a vital role, and with test score gains of school-aged children as the measurement variable, there are only relatively short panels available. My approach accounts for the fact that locations can shape school value-added, by bringing Bonhomme et al. (2019)'s approach to an educational setting. To achieve this, I extend their estimation in two important ways. First, my approach is based on longitudinal data with at least three consecutive periods in order to incorporate a value-added test score model. I account for changes in test scores – rather than levels, as it is typically done in the labor literature where the outcome variables are log-earnings. This allows for human capital accumulation. The second main distinction is the addition of a third layer. The rationale for the three dimensions is similar to the labor literature (c.f., Mann (2024)): two schools of the same category could, for example, face different patterns of student transitions if they are located in distinct neighborhoods. Similarly, the two school's test score distributions could be very different if they or their students are located in distinct neighborhoods even though they have the same distribution of family types. My approach is general and can be applied to other settings where growth in the outcome, rather than levels, is of interest, or where a third dimension such as geography might be relevant for worker-firm or other match-specific relations of economic agents.

In this paper, I find that not every neighborhood improves or enhances a child's test scores equally. This is relevant as other attempts addressing the family-neighborhood decomposition problem often incorporate the effects of parents and neighborhoods as additive components in a linear way – ruling out any sort of interaction between family and neighborhood attributes (Chetty and Hendren (2018b); Chetty et al. (2014)). My findings of significant treatment heterogeneity at the neighborhood level are consistent with, e.g., Wodtke et al. (2016)'s finding that family and other social connections can be mutually reinforcing. My analysis reveals new and specific evidence on who and where to target. To create effective – and where necessary, targeted – policy interventions, it is crucial to account for heterogeneous neighborhood effects in addition to sorting, and to untangle neighborhood effects from school value-added.

This paper proceeds as follows. In the next section, I discuss the framework of analysis and elaborate on model identification. In Section 3, I provide institutional background information on North Carolina's public school system and network structure, give an overview of the data, and lay out details on the construction of the panel data frame. Section 4 gives details on the three-step estimation procedure. Section 5 presents the results and Section 6 shows the counterfactual analyses. Section 7 concludes.

2 Framework of Analysis

In this section, I build an empirical framework for the cognitive achievement of children from different types of families, schools, and neighborhoods, which allows for interaction effects between unobserved family-, school-, and neighborhood-heterogeneity without assuming a functional form. The framework aims to uncover heterogeneous effects of schools and neighborhoods on children's test scores, taking into account non-random sorting by families. Owing to institutional features in my setting, some neighborhoods are served by multiple schools, and some schools draw students from multiple neighborhoods, resulting in various combinations of neighborhood-school pairings. The primary source of identification for these match-specific effects stems from the variation introduced when children from different types of families change schools and/or move to new residences.

The set of neighborhoods is partitioned into classes and the set of schools is partitioned into categories, based on a preliminary cluster analysis. Because identification in the model depends on movements between schools and neighborhoods – and upward mobility occurs infrequently – there is a risk of bias arising from limited mobility (Andrews et al., 2008). By pre-classifying neighborhoods into classes and schools into categories, the model becomes more parsimonious and there is sufficient variation across these groups to capture meaningful heterogeneity. Conditional on these pre-classified groups, I use a correlated random effects approach to identify unobserved heterogeneity among children and classify them into family types. As a result, unobserved heterogeneity of families, schools, and neighborhoods is at the type-, category-, and class-level, respectively. While neighborhood classes and school categories are treated as given within the model, family types are modeled as draws from a discrete distribution. By treating family heterogeneity differently, I can take the network of family-school-neighborhood links as given, eliminating the need to specify a probabilistic structure for the mobility patterns of families between neighborhoods and schools (Bonhomme, 2020).

The objectives of the empirical framework are twofold. First, the purpose is to recover the distributions of test score gains across children of various family types in neighborhoods of different classes and schools of different categories. These match-specific distributions are informative about heterogeneous effects of neighborhoods and schools on children's test scores: the same type of child could experience different test score gains depending on where they live and go to school. Second, the framework enables the identification of the shares of various family types within schools and neighborhoods of the same category and class, respectively.

My procedure builds on Bonhomme et al. (2019)'s matched worker-firm two-step estimator and brings it to my family-school-neighborhood context by extending it in two crucial ways. First, I add a third dimension to account for family-school-neighborhood effects. Second, I use test score gains rather than levels as the measurement variable. Although this requires three observation periods for each child – an additional period compared to their static model – it helps mitigate the initial condition problem in test score analysis by better accounting for human capital accumulation over time, which is crucial for understanding the impact of schools and residential locations on student outcomes. This strategy differs from the traditional approach in the labor literature, which typically focuses on log-earnings levels because wages inherently reflect the cumulative effect of past human capital investments.

Both the approach of adding a third layer and the focus on gains are general and can be applied to other settings where an additional dimension, such as geography, is relevant for worker-firm or other type-specific relationships among economic agents, or where analyzing growth in the outcome variable, rather than levels, is essential.

In what follows, I first explain the setting, timing, and key assumptions that form the foundation of the analysis. I then discuss the two empirical equations, which are the core of the econometric framework, and under what conditions the parameters in the equations are identified.

2.1 Environment

Setting and Notation. I consider a setting in which there are J neighborhoods, C schools, and N families who each have only one child i. Correspondingly, j_{it} denotes the neighborhood where child i lives at time t, and c_{it} represents the school that child i attends at time t. Child i receives *test score gains* ν_{it} at time t. Test score gains reflect current test score performance that is not explained by lagged test scores. They are estimated in Section 4.1.

Families, schools, and neighborhoods are heterogeneous. Heterogeneity across neighborhoods is characterized by their *class*: $k_{it} = k(j_{it}) \in \{1, ..., K\}$ is the class of neighborhood j_{it}, where $K < J$. Heterogeneity across schools is characterized by their *category*:

$s_{it} = s(c_{it}) \in \{1, ..., S\}$ is the category of school c_{it}, where $S < C$. Heterogeneity across families, including their child, is characterized by their *type*: $\alpha_i \in \{1, ..., L\}$ represents the type of both child i and their family.

Heterogeneity of neighborhoods, schools, and families is assumed to be time-invariant over the considered time period. The categorization of schools does not depend on location. However, neighborhoods of different classes may vary in terms of the proportion of schools from each category that are available. In every period t, child i is associated with a place $p_{it} = (k_{it}(j_{it}), s_{it}(c_{it}))$ based on the class k_{it} of their neighborhood j_{it} and category s_{it} of their school c_{it}. I estimate both neighborhood classes and school categories in Section 4.2. Child types are estimated in Section 4.3.

Time is discrete, and I consider two consecutive years of observations: $t \in \{1, 2\}$. Between periods 1 and 2, some children move ($j_{i1} \neq j_{i2}$) and/or transition to a different school ($c_{i1} \neq c_{i2}$). Mobility between a neighborhood in period 1 and another neighborhood in period 2 ($j_{i1} = j$ to $j_{i2} = j'$) and/or mobility between a school in period 1 and another school in period 2 ($c_{i1} = c$ to $c_{i2} = c'$) is denoted as $m_{i1} = 1$. Mobility also implies a change of place (from $p_{i1} = p$ to $p_{i2} = p'$): $p \to p'$. The new place can either be of the same (neighborhood class, school category)-pair or not: $p = p'$ or $p \neq p'$, respectively.

Timing and Main Assumptions. In period 1, the type of child i, α_i, is drawn from a distribution that depends on the neighborhood class k_{i1} and the school category s_{i1} where they live and go to school, respectively. Then, the child draws test score gains ν_{i1} from a distribution that depends on α_i, k_{i1}, s_{i1}.

At the end of period 1, the child transitions to another neighborhood and/or school ($m_{i1} \in \{1, 0\}$) with a probability that may depend on their type α_i, current neighborhood class k_{i1}, and current school category s_{i1}. The probability that the class of the neighborhood where the child moves to is $k_{i2} = k'$ and category of the school where the child transitions to is $s_{i2} = s'$ may also depend on α_i, k_{i1}, s_{i1}.

In period 2, for transitioners ($m_{i1} = 1$), the child's test score gains ν_{i2} after changing places ($p \to p'$) are drawn from a distribution that depends on α_i, k_{i2}, s_{i2}. For stayers ($m_{i1} = 0$), ν_{i2} is drawn from an unrestricted distribution that may depend on $\alpha_i, k_{i1} = k_{i2}, s_{i1} = s_{i2}$.

As in Bonhomme et al. (2019), two formal "network exogeneity" assumptions follow.

Assumption 1 (Mobility Determinants).
$(m_{i1}, p_{i2}) \perp\!\!\!\perp (\nu_{i1}) \mid (\alpha_i, p_{i1})$

Assumption 2 (Conditional Serial Independence of Test Score Gains for Transitioners).
$(\nu_{i2}) \perp\!\!\!\perp (\nu_{i1}, p_{i1}) \mid (\alpha_i, p_{i2}, m_{i1} = 1)$

In words, *Assumption 1* states that, conditional on the type of the child (α_i) and the (neighborhood class, school category)-place in period 1 (p_{i1}), mobility (m_{i1}) and the (neighborhood class, school category)-place in period 2 (p_{i2}) are independent of test score gains in period 1 (ν_{i1}). Put differently, conditional on knowing family type (α_i), school category and neighborhood class in period 1 (p_{i1}), knowing how much was effectively learned in period 1, the test score gain (ν_{i1}) does not provide additional information regarding whether (m_{i1}) and where (p_{i2}) the child will transition to in period 2. So, mobility does not directly depend on (ν_{i1}), given (α_i, p_{i1}).

Assumption 2 states that test score gains in period 2 (ν_{i2}) are independent of test score gains (ν_{i1}) and (neighborhood class, school category)-place (p_{i1}) in period 1 conditional on the type of the child (α_i), the new place (p_{i2}), and mobility ($m_{i1} = 1$). In other words, once the child's type and new place are known and a transition happened, then their test score gains and place in period 1 do not further predict test score gains in period 2. Thus, (ν_{i2}) after a transition is not allowed to depend on the child's previous place or test score gains, conditional on the type and the new place.

The conditional independence assumptions have statistical implications, as I will explain in the next subsection. First, I want to illustrate the assumptions of my model with a simple example.

Example. Consider the following dynamic test score regression:

$$Y_{it} = \psi_t Y_{i,t-1} + a_t(p_{it}) + b_t(p_{it})\alpha_i + \varepsilon_{it},$$

where $\mathbb{E}(\varepsilon_{it} \mid \alpha_i, p_{it}, Y_{i,t-1}) = 0$. In this regression model, Y_{it} is the test score of child i at time t, and $Y_{i,t-1}$ is the lagged test score from the previous period. These past test scores of one lag can be interpreted as a sufficient statistic for everything that happened up until the previous period. Further, p_{it} is the place, where $a_t(p_{it})$ is the place effect. The term $b_t(p_{it})\alpha_i$ represents the interaction between the child's type α_i and the place p_{it}. While it is possible to separate the neighborhood and school effects, as well as their interactions, for the sake of clarity, these elements are incorporated together as places. The residual term ε_{it} captures any unexplained variation in test scores after accounting for these factors, and is assumed to be uncorrelated with them. One may further include a vector of child specific controls (such as grade and year (cohort) of test taken), but I am abstracting from such observable characteristics for conciseness.

A simpler version than the above regression model would be $Y_{it} = \psi_t Y_{i,t-1} + \nu_{it}$. Under the assumption that $\psi_t = 1$ – meaning that past test scores fully persist into the current period without any decay or amplification – and that lagged test scores summarize all relevant past learning, making them a sufficient statistic for predicting current performance, it follows that $\nu_{it} := Y_{it} - Y_{i,t-1}$ is an auxiliary variable representing the first difference in test scores. This difference reflects the incremental change in performance, effectively isolating the influence of the current period's environment and circumstances on the student's progress. While the above regression includes test score levels Y_{it} as outcome, the general model in the next subsection will have first differences ν_{it} as measurement variables. The general model posits that the factors influencing the change in test scores, excluding the impact of prior scores ($Y_{i,t-1}$), denoted as ν_{it}, can be further explained by the match of a child's family background (α_i) with their residential neighborhood (k_{it}) and school environment (s_{it}) in period t.

However, while the above test score regression model assumes additive separability, I will now introduce a more general model with no functional form assumptions, except that family types are finite and discrete. The model captures heterogeneous and potential nonlinear relationships between different types of families, schools, and neighborhoods. This provides more nuanced insights into how these match-specific interactions shape educational outcomes, as test score gains may be nonmonotonic in neighborhood-school productivity. In my empirical implementation, I will use a stratified finite mixture specification à la Bonhomme et al. (2019).

The framework is represented by a finite mixture model to address unobserved heterogeneity among children. The model assumes that the overall distribution of test score gains is generated from a combination of distributions (or components), each corresponding to a different latent child-type group. The model "mixes" these distributions probabilistically based on group memberships that are not directly observed. This approach serves both to understand the overall structure of the data by identifying latent groups and to provide insights at the individual level by estimating each child's probability of belonging to these groups. Consequently, the mixture model explains test score gains as arising from a mixture of underlying subpopulations (latent child groups), where each group has a distinct set of parameters. The aim is to estimate both the group memberships (reflecting latent child heterogeneity) and the parameters governing each group's distribution.

In addition to the unobserved child-type heterogeneity, the model incorporates observed covariates – specifically, neighborhood classes and school categories – which are treated as fixed effects. Because these neighborhood classes and school categories are predetermined, including them as fixed effects allows the model to control for the influence of neighborhood and school environments when identifying the latent child types from the data (Bonhomme et al., 2019).

2.2 Empirical Model

I now provide conditions for the identification of distributions of test score gains for all family types (α), school categories (s), and neighborhood classes (k), as well as family type distributions for all places $p_{it} = (k_{it}, s_{it})$ with a more general model à la Bonhomme et al. (2019) with no functional form assumptions.

The model underlays *Assumptions 1-2* and takes ν_{it} and $p_{it} = (k_{it}, s_{it})$ as given. Because place is a function of both the neighborhood class and the school category, I am left with a bipartite network connecting children to places in every period. In Section 3, I will show how to estimate test score gains ν_{it} as well as category membership $s(c)$ for each school c and class membership $k(j)$ for each neighborhood j. As a result of this pre-grouping, both school categories and neighborhood classes are treated as discrete fixed effects in the model. In contrast, family types are modeled as discrete correlated random effects, correlated with places (school categories and neighborhood classes).

Let $\alpha_i \in \{1, ..., L\}$ denote the latent child types, where L is known and taken as given. Let $F_{p\alpha}(\nu_1)$ denote the cumulative distribution function (cdf) of test score gains in period 1 in place p for child type α. Let $G_{p'\alpha}(\nu_2)$ denote the cdf of test score gains in period 2 for place p' and type α for children transitioning to a new place between periods 1 and 2 ($m_{i1} = 1$, i.e., $p \rightarrow p'$). Let $\pi_{p\rightarrow p'}(\alpha)$ denote the probability distribution of α_i for children who transition between a place p and place p'. Let $q_p(\alpha)$ denote the cross-sectional distribution of α_i for all children in place p in period 1.

Then, the likelihood of a child receiving test score gains ν_{i1} and ν_{i2} less than or equal to specific values ν_1 and ν_2 in periods 1 and 2, respectively, can be expressed with a joint conditional probability. In particular, the bivariate distribution of test score gains conditional on observing children transitioning from place p in period 1 to place p' in period 2, can be written as:

$$Pr(\nu_{i1} \leq \nu_1, \nu_{i2} \leq \nu_2 | p \rightarrow p') = \sum_{\alpha=1}^{L} \pi_{p\rightarrow p'}(\alpha) \cdot Pr(\nu_{i1} \leq \nu_1, \nu_{i2} \leq \nu_2 | p \rightarrow p', \alpha_i = l),$$

by the law of total probability. The equation represents the structure of a finite mixture model, where the observed test score gains (ν_{i1}, ν_{i2}) on the left-hand side are the endogenous variables, conditional on the transition from p to p', which is treated as given. This conditional bivariate distribution of test score gains is a data object that can be expressed as an expectation. The expectation is formulated as a sum over the types, where each term represents the product of two components: the probability of the type given the transition and the probability of the observed test scores given the transition and the type. Thus, the right-hand side includes unobserved heterogeneity, which is captured by summing over all possible latent groups α. The weights $\pi_{p \to p'}(\alpha)$ represent probabilities (or mixing proportions) associated with every latent group α. The probability of being of a certain type is left completely unrestricted in the model, with the only assumption that types are finite and discrete. This allows the types to vary flexibly across predetermined places.[12] Given *Assumptions 1-2*, the distribution can further be rewritten as a mixture of independent distributions:

$$Pr(\nu_{i1} \leq \nu_1, \nu_{i2} \leq \nu_2 | p \to p') = \sum_{\alpha=1}^{L} F_{p\alpha}(\nu_1) \cdot G_{p'\alpha}(\nu_2) \cdot \pi_{p \to p'}(\alpha) \qquad (1)$$

Equation 1's finite mixture model representation holds because the framework underlays the previously discussed *Assumption 1*, that (ν_{i1}) is independent of (m_{i1}, p_{i2}) conditional on (α_i, p_{i1}), and *Assumption 2*, that (ν_{i2}) is independent of (ν_{i1}, p_{i1}) conditional on $(\alpha_i, p_{i2}, m_{i1} = 1)$.[13]

Equation 1 is the first key empirical equation of this framework and is primarily intended to identify the treatment heterogeneity of neighborhoods and schools for different types of families among transitioning children. The idea is that it reveals matched α-p-specific test score gains distributions, i.e. specific distributions $F_{p\alpha}$ for period 1 and

[12]The idea is that the right-hand side of the model transforms the child's latent type, along with the (neighborhood class, school category)-places, which are treated as exogenous variables, into test score outcomes. The left-hand side of the mixture model represent outcomes, while children within the same type undergoing the same transition are assumed to have identical preferences and constraints. In particular, while neighborhood and school assignments are fixed and certain ("hard-clustered"), the assignment of children to latent family type groups in this finite mixture model is probabilistic ("soft-clustered"). For example, if the population is assumed to consist of three different family types, $\alpha \in \{1, 2, 3\}$, each child is assigned a probability of belonging to latent type 1, 2, or 3, with these probabilities summing to 1 for each child.

[13]The statistical implications of *Assumptions 1-2* enhance identification by allowing lower-dimensional objects on the right-hand side. See Bonhomme et al. (2019) for further details.

$G_{p'\alpha}$ for period 2 for different α and p in period 1 and different α and p' in period 2, respectively.

From *Equation 1*, we can derive the second key equation. The marginal distribution for ν_{i1} can be expressed as a mixture of the conditional distributions $F_{p\alpha}(\nu_1)$ weighted by $q_p(\alpha)$, which represents the distribution of child types in place p:

$$Pr(\nu_{i1} \leq \nu_1 | p) = \sum_{\alpha=1}^{L} F_{p\alpha}(\nu_1) \cdot q_p(\alpha) \qquad (2)$$

Note that *Equation 2* focuses on the first period only, before a potential transition, and considers the entire cross-sectional sample, including the non-transitioning children. The finite mixture representation of *Equation 2* shows a decomposition of the cdf of test score gains in period 1. Its main role is to identify the extent of sorting of various types of families into different types of neighborhoods and schools. In particular, $q_p(\alpha)$ intends to reveal the share of all family types for every place separately.

With *Equation 1* and *Equation 2*, the empirical model includes two stratified finite mixture models that build the key empirical equations of this framework. I will estimate these equations separately in Chapter 4. Next, I provide conditions under which all parameters appearing in *Equations 1-2* are identified. The list of parameters to identify is $F_{p\alpha}$ and $G_{p'\alpha}$ for all (α, p, p'), $\pi_{p\rightarrow p'}(\alpha)$ for all α for all (p, p')-pairs for which there are transitions $p \rightarrow p'$, and $q_p(\alpha)$ for the cross-section in period 1.

2.3 Model Identification

The overarching objective of the empirical framework is to decompose the overall distribution of test score gains into heterogeneous subpopulations (latent child-type groups) that capture specific family-school-neighborhood match characteristics. Key source of identification is given by children moving and/or switching to a different school. By fully exploiting test score gains around a transition, heterogeneous neighborhood-school effects on a child's test score gains can be identified and consistently estimated, provided the following conditions of *Assumption 3* are satisfied.

This assumption is identical to the one presented in Bonhomme et al. (2019), where a more formal version is offered. *Definition 1*, which is introduced first, serves as the foundation for understanding the conditions specified in *Assumption 3*.

Definition 1 (Connecting Cycle).
A *connecting cycle* of length R is a pair of sequences of places $(p_1, \ldots, p_R)$ in period 1, and $(p'_1, \ldots, p'_R)$ in period 2, with $p_{R+1} = p_1$, such that $\pi_{p_r,p'_r}(\alpha) \neq 0$ and $\pi_{p_{r+1},p'_r}(\alpha) \neq 0$ for all $r \in \{1, \ldots, R\}$ and $\alpha \in \{1, \ldots, L\}$.

The condition $p_{R+1} = p_1$ means that after progressing through all R places, the sequence loops back to the starting place, forming a closed cycle. This ensures the structure is circular rather than linear, which is key to defining a connecting cycle. The conditions $\pi_{p_r,p'_r}(\alpha) \neq 0$ and $\pi_{p_{r+1},p'_r}(\alpha) \neq 0$ for all $r \in \{1, \ldots, R\}$ and $\alpha \in \{1, \ldots, L\}$ ensure that there is a non-zero transition or connection between the places p_r and p'_r, as well as between p_{r+1} and p'_r, for all possible values of α. This guarantees that every place in the cycle is strongly connected, both within its period and across periods.

Assumption 3 (Graph Connectivity).
Any two (neighborhood class, school category)-places p and p' belong to a connecting cycle.

Assumption 3 includes two additional components. First, the distribution of family types across different (neighborhood class, school category)-places must be asymmetric. In other words, there must be non-random sorting of families into places; otherwise, identical family types would cancel out, hindering the identification of effects. Second, the match-specific distributions must be linearly independent to ensure they lead to the identification of unique parameters.

A crucial condition for identification, as stated with *Assumption 3*, is the presence of connecting cycles ("graph connectivity").[14] For example, with three possible neighborhood classes (A, B, C) as well as three possible school categories (1, 2, 3) in the sample population, one gets nine distinct (neighborhood class, school category)-places, $p = (k, s)$ and $p' = (k', s')$, for both the period before and the one after the transition, respectively, where all nodes are required to be connected (see Figure 1).[15] This full heterogeneity requirement at the (neighborhood class, school category)-place level enables the identification of the heterogeneous effects that schools of different categories interacting with neighborhoods of different classes have on the child's outcomes.

[14] See Bonhomme (2020) or Bonhomme et al. (2019) for more details.

[15] Every (neighborhood class, school category)-pair needs to contain transitions of all types of children. The presence of a connecting cycle requires that there is a positive proportion of every child type among transitioning children from $A1 \rightarrow A'1'$, $A1 \rightarrow A'2'$, ..., $C3 \rightarrow C'3'$, respectively.

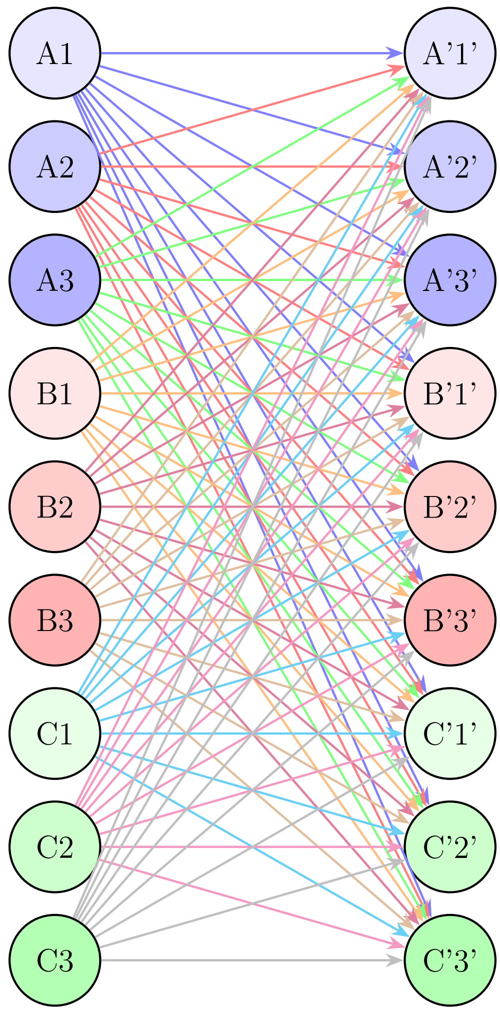


Figure 1: Graph Connectivity

This graph illustrates the concept of connecting cycles ("graph connectivity"), where every possible combination of (neighborhood class, school category)-pair p *in period* 1 *needs transitions from and to every other (neighborhood class, school category)-pair* p' *in period* 2.

It is important to note that the categorization of a school is independent of the neighborhoods that the school is associated with. In practice, due to institutional constraints, neighborhoods of certain classes may not contain schools of all categories, and some transitions between nodes might be empirically absent. In such cases, *Assumption 3*, which requires graph connectivity, would be violated. Even when the graph is fully connected, a scarcity of certain types of transitions between specific nodes can lead to only partial or "weak" identification of the model or parameters of interest, potentially resulting in biased estimates (Bonhomme, 2020). If point-identification is not the primary objective, the requirements for cycles can be relaxed; and documenting such limitations remains valuable from an applied perspective. However, given the nonlinear context of my analysis, imposing additional structure might be inappropriate, which is why I aim for point identification to capture the heterogeneity in neighborhood effects and school value added. Therefore, there is a trade-off between using fewer family types, school categories, and neighborhood classes – which ensures graph connectivity and facilitates the identification of heterogeneous neighborhood and school effects on test scores – and using more subpopulations, that provide finer granularity but may limit the ability to observe certain effects, thus resulting in weak identification.

Taking Stock. Under the structure of the model, maintaining *Assumptions 1-3*, and given observations on school categories, neighborhood classes, and test score gains, one can uniquely recover both the match-specific distributions of test score gains, as well as the family type shares. This provides a non-parametric identification result for the model under discrete family heterogeneity. In particular, for a chosen number of finite latent child types α: (i) $F_{p\alpha}$ and $G_{p'\alpha}$ are identified for all (α, p, p'); (ii) $\pi_{p\to p'}(\alpha)$ is identified for all α for all (p, p')-pairs for which there are transitions $p \to p'$; (iii) type proportions $q_p(\alpha)$ are identified for the cross-section in period 1.[16]

3 Institutional Context and Data

The identification strategy in this paper is motivated by the following observations. Distinguishing the school effect from the neighborhood effect in explaining variation in academic performance can be key, as targeted policy implementations hinge on knowing who to target and where to intervene. This motivates the focus on North Carolina for my empirical analysis. The state's extraordinary public school system provides two institutional conveniences. First, multiple residential areas are assigned to the same school and numerous schools can be linked to the same residential area. Second, North Carolina executes statewide standardized test that are administered annually to children enrolled in public and charter schools in the K-12 system. In addition to this unique neighborhood-school network structure and consistent statewide testing, the North Carolina Education Research Data Center (NCERDC) provides comprehensive microdata, allowing for detailed observation of various longitudinal characteristics of all school-aged children. The idea is to construct a panel data set containing children that are followed over time, keeping track of their test scores, and where they live and go to school. Key variation comes from children who change schools and/or move to

[16]See Bonhomme et al. (2019) for a formal theorem.

a different neighborhood during panel coverage, which can be used to identify potential changes in school and neighborhood attributes around the transitioning event that could impact standardized test scores differently. Before discussing the construction of the matched family-school-neighborhood panel and detailed data sources, I provide additional institutional background on North Carolina's neighborhood-school network structure and statewide standardized testing.

3.1 Institutional Background

North Carolina's schools system has a peculiar structure in that it consists of a valuable network structure connecting neighborhoods and schools in different combinations, and it administers standardize end-of-grade tests for all children enrolled in a public school.

Neighborhood-School Network Structure. In North Carolina, multiple neighborhoods share the same school, and a single neighborhood is often associated with multiple schools. Figure 2 illustrates this with an example. This institutional setting generates a suitable network structure with various neighborhood-school pairings. In addition, while children can both move and transition to a new school, children can also just move to a different neighborhood but stay in the same school, or change schools but stay in the same neighborhood. Utilizing all these variations leads to different neighborhood-school combinations that could impact test scores differently.

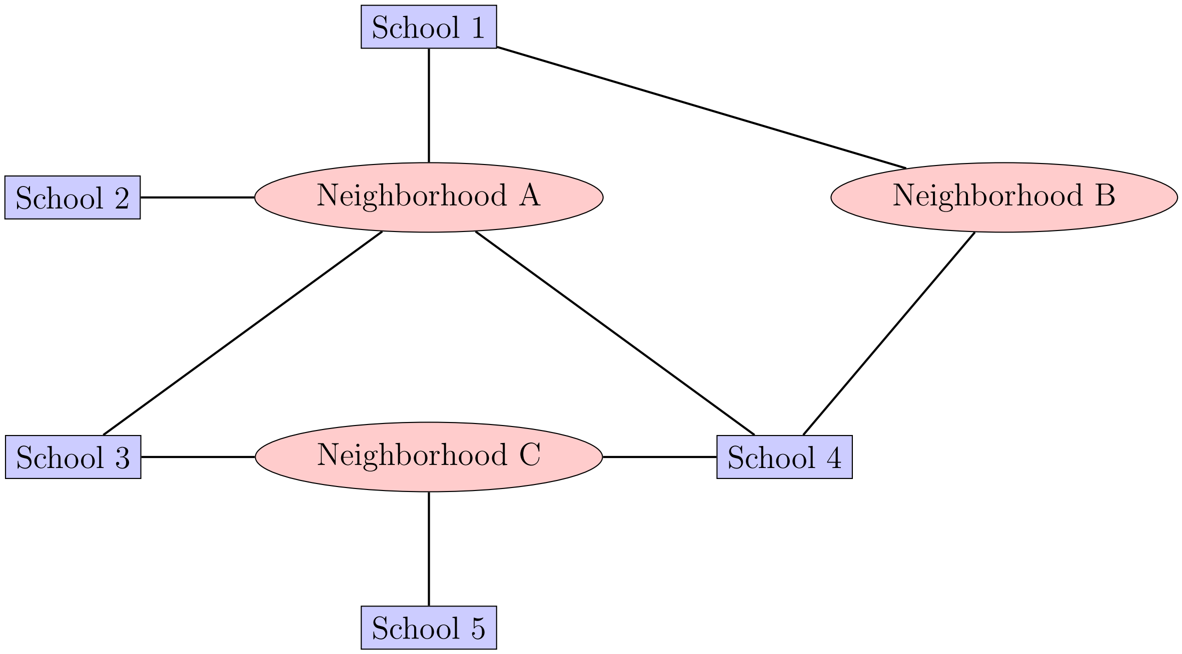


Figure 2: Example of the Neighborhood-School Network Structure

Standardized Tests. All children in North Carolina's public schools from grade three to eight are required to participate in statewide standardized tests. These assessments are dependent on grade level and can be used to measure the extent to which children understand and can apply what they are learning in school. Thanks to its standardized nature, the results allow the comparison of a child's scores with those of other students across all public schools in North Carolina, which can help to find out inefficiencies and potential need for action. The most comprehensive assessments are the end-of-grade (EOG) tests in both mathematics and reading that students take in their elementary and middle school years. I focus on the EOG since it is tested annually and thus allows me to follow children from grades three through eight, leveraging up to six consecutive years of test score performance observations for all grades during this period. More details follow in the subsequent subsection, where I describe the data and their sources.

3.2 Construction of the Family-School-Neighborhood Panel

The matched family-school-neighborhood dataset is built by leveraging detailed longitudinal microdata from the NCERDC and neighborhood-level demographic data from the American Community Survey (ACS) of the U.S. Census Bureau.

Data Sources. My data comes primarily from the NCERDC in North Carolina, which maintains one of the largest databases on a state's public schools, school districts, teachers, and students in the United States. The student-level data include children's performance in annual EOG tests in Mathematics and Reading from grades three to eight, their school identifiers, and the 2010 census block group (neighborhood) they live in. I further match the child identifiers with information regarding individual characteristics, such as economically disadvantaged status, as well as school characteristics, including teacher background information and classroom-related details, which are also provided by the NCERDC. The neighborhood identifiers, stemming from 2010 census block group assignments based on children's residential addresses documented in the NCERDC, are further linked with Census block group-level information, such as population counts, average income and education, and racial composition, from the American Community Survey (ACS) five-year estimates ($2013-2017$).

Sampling. I focus on all children between grades $3-8$ in the NCERDC database for whom I observe EOG test scores for at least three consecutive grades and years. I drop all children for whom I do not observe their geocoded addresses and associated 2010 census block group level and I also eliminate children with missing information regarding the identifier of the public school attended across all considered consecutive years. I further drop children from the sample who attend a newly created school in period 1 or a school with an identifier that cannot be found in an earlier period in association with another child. This is because the categorization of schools is done using stayers in period 1. If a mover's period 2 school identifier cannot be matched with a stayer's period 1 school identifier, I am not able to assign a category to that school. In addition, I exclude all neighborhoods with fewer than 10 children residing there, as well as all schools with fewer than 10 children enrolled in my sample. I focus on all children who did neither residentially relocate nor experience a change in schools during their first two observations. All these variables are available yearly for at least three times in a row between 2010 and 2017. Even though I require only three consecutive observations for every child, I can make use of six observations for the same child by pooling the data and treating the last three observations as if they belong to a different child. This is solely to gain power and hopefully increase precision without loss of generality.

Measurement and Data Transformation. Children's EOG test scores in Mathematics and Reading are each standardized by cohort using z-scores and then combined to represent the child's average test score performance for each available year. For simplicity, I assume that every family has exactly one child, so the terms "family" and "child" can be used interchangeably. The family is identified by the child's identifier, which is time-invariant. The school is identified by the school identifier, and the neighborhood is identified by the 2010 census block group identifier. This implies that all residential addresses in the NCERDC that share the same 2010 census block group belong to the same neighborhood. Given the three observations between 2010 and 2017 for every child, I transform each child's year identifiers into years 0, 1, and 2. Year 0 information is primarily used to ensure that the child's school and neighborhood identifiers do not change between years 0 and 1, and it is otherwise only used in step one of my estimation procedure to calculate the first differences in test scores between years 0 and 1, providing a measure of test score gains in period 1 for each child. I create an additional transition dummy variable indicating whether a child moved to a new neighborhood and/or changed schools or not between periods 1 and 2. This leaves me with a panel containing stayers, for whom both neighborhood and school identifiers are constant throughout the observation window, and transitioners, for whom at least one of the identifiers changes between their last two observations.

Final matched family-school-neighborhood dataset. My final three-sided matched family-school-neighborhood dataset contains the following key information for every child i: a sequence of test scores (Y_{i0}, Y_{i1}, Y_{i2}); school identifiers $(c_{i0} = c_{i1}, c_{i2})$; and neighborhood identifiers $(j_{i0} = j_{i1}, j_{i2})$. These identifiers are further matched with various observed characteristics related to the child's background, school, and place of residence. For convenience, the dataset also contains a mobility indicator variable, $m_{i1} \in \{1, 0\}$, indicating whether the child changes their neighborhood and/or school between periods 1 and 2. My final sample includes $5,423$ neighborhoods and $1,690$ schools. There are $31,029$ distinct neighborhood-school combinations. In addition, my sample consists of $336,225$ children across six cohorts who where within grades $3-8$ during the sample period $2010-2017$. Among them, $45,035$ (13.4%) experience mobility. This results in considerable mobility rates: the ratio of transitioning children to staying children is 15.5% in the sample.[17] Details regarding different types of mobility samples are reported in Table 1. The three-sided matched family-school-neighborhood panel created can be mapped to the model and is key for the analysis. This dataset allows for tracking children over time, including where they live, which school they attend, and how they perform on tests. Children who move and/or transition to a different school introduce potential variation in matched neighborhood-school attributes, which could affect their academic performance in different ways.

[17] In comparison, Bonhomme et al. (2019) observe relatively low mobility rates for their workers, with a ratio of job movers to job stayers of 3.3% in their sample.

Table 1: Mobility Summary Statistics

Sample	Number of Children (%)
Total children in the sample	$336,225$ (100.0%)
Stayer	$291,190$ (86.60%)
Neighborhood + School changed	$10,818$ (3.22%)
Neighborhood only changed	$17,862$ (5.31%)
School only changed	$16,355$ (4.86%)
Neighborhood and/or School (at least one) changed	$45,035$ (13.4%)
Neighborhood changed	$28,680$ (8.53%)
School changed	$27,173$ (8.08%)

4 Model Estimation

The conditions necessary for identifying family-school-neighborhood match-specific distributions of test score gains in the context of sorting and interactions, as outlined in Chapter 2, take test score gains and places as given. I use test score gains as the measurement variable, enabling a clearer account of human capital accumulation over time, which is crucial for understanding the impact of schools and residential locations on student outcomes. In addition, the model takes places as given, which are a function of neighborhood classes and school categories, and not single neighborhoods and schools, respectively. Theoretically, identification is not restricted to the group level, and a neighborhood class k_{it} can align with the neighborhood j_{it} and a school category s_{it} can coincide with the school c_{it}. But, most children transition to similar neighborhoods and schools, and some experience downward mobility. Instead, upward mobility transitions (e.g., transitioning from a neighborhood and school with relatively lowest test score gains to an environment with relatively highest test score gains is very rare in the data). However, the goal is to achieve full heterogeneity in the empirical model, observing transitions from all types of families from all types of environments to all types of environments, which requires point identification. To mitigate incidental parameter biases due to limited mobility, I rely on a dimension reduction technique to reduce the dimension of the heterogeneity of neighborhoods and schools. Since some schools and neighborhoods are very similar, I do not need to distinguish between highly similar entities. I extend Bonhomme et al. (2019) strategy and apply a distributional k-means algorithm to reduce the number of parameters by separately clustering neighborhoods into classes as well as schools into categories. This reveals, for each child, the neighborhood class where the child lives as well as the school category of the school that the child attends. I then group these new cluster assignment identifiers into (neighborhood class k, school category s)-places. Technically, this last step is not necessary, but it allows for computational tractability without loss of generality. I now describe my three-step estimator, where I first recover test score gains with a value-added approach, then recover places in a second step, and estimate the distributions of match-specific test score gains in a third step.

4.1 Recovering Test Score Gains with a Value-Added Approach

In my analysis, I utilize three consecutive years of test scores for every child in the data. Children's End-of-Grade (EOG) test scores in Mathematics and Reading are each standardized by cohort (grade and year) using z-scores and then combined to represent the child's average test score performance for each available year: Y_{i0}, Y_{i1}, Y_{i2}. I then run the following simple value-added regression models, which are essentially first-differences in test scores:

$$Y_{i1} = \gamma_1 Y_{i0} + \nu_{i1}, \quad \text{where} \quad \gamma = 1$$

$$Y_{i2} = \rho_2 Y_{i1} + \nu_{i2}, \quad \text{where} \quad \rho = 1$$

This results in $\hat{\nu}_{i2} = Y_{i2} - Y_{i1}$ and $\hat{\nu}_{i1} = Y_{i1} - Y_{i0}$, which represent test score gains. This approach allows me to isolate the new learning achieved by students, independent of their starting ability, and to assess how family, school, and neighborhood factors contribute to this improvement. Thus, estimation of the test score gains consists of running the above two value-added regressions and then isolating the residuals and using them in the key equations of the model (cf., *Equation 1* and *Equation 2*). By measuring the growth in test scores, rather than levels, I mitigate the initial condition problem in my framework of analysis. This further allows me to account for some sort of human capital accumulation – something that is missing when working with levels as in Bonhomme et al. (2019).

Prior to clustering similar neighborhoods into classes as well as similar schools into categories with test score gains of stayers in period 1, I normalize the input data for better performance. K-means algorithms often overvalue negative values whenever the data contains both negative and positive values. Therefore, I use min-max scaling to normalize test score gains. This allows me to transform the numerical test score gains scores into a non-negative range $\in [0, 1]$. The formula for min-max scaling is:

$$\nu_{it} = \frac{\hat{\nu}_{it} - min(\hat{\nu}_t)}{max(\hat{\nu}_t) - min(\hat{\nu}_t)}$$

where $\hat{\nu}_{it}$ is a child's test score gain in the EOG test, $min(\hat{\nu}_t)$ is the lowest score value among all children of the same cohort who took that test in the data, $max(\hat{\nu}_t)$ is the highest performance of that test and cohort, and ν_{it} is the considered child's scaled value of the test between 0 and 1. By subtracting the lowest test score performance from each test score gain and then dividing by the range, the individual test score gains are scaled between 0 and 1, ensuring that the minimum value is mapped to 0 and the maximum value is mapped to 1, while maintaining the relative relationships among the other values.

4.2 Recovering Places with Distributional K-Means

As previously described, I reduce the dimension of neighborhoods and schools by grouping them into classes and categories, respectively. I use a distributional k-means algorithm as the clustering method to separately partition j neighborhood observations into k classes according to their test score growth distribution, and similarly, c school observations into s categories. The clustering is based on stayers only in period 1, and on that account, not based on the outcome variable but a different sample.

The rationale for using changes in test scores to cluster neighborhoods lies in their role as neighborhood-specific outcomes within a model where the sole source of ex-ante neighborhood heterogeneity is the neighborhood class. This theoretical foundation justifies the clustering: in the model, two large neighborhoods of the same class would exhibit identical distributions of test score changes. While this assumption does not hold perfectly in practice, k-means clustering seeks to identify the neighborhood classes that most effectively differentiate the distributions of test score changes across neighborhoods. The same idea holds for schools.

To control for population sizes, I weight every school by the number of children attending that school; and every neighborhood by the number of children living there.

Every neighborhood then belongs to the class with the nearest cluster centroid (mean distribution), and every school then belongs to the category with the nearest cluster mean distribution. I consider the entire distribution of test score changes, albeit at specific discrete points for estimation purposes. The analysis is based on 40 percentiles.

4.2.1 Clustering Neighborhoods into Classes

I consider the distributions of test score changes across neighborhoods and solve the following distributional k-means problem in order to partition the J neighborhoods into K classes, with the clustering process weighted by the number of children in each neighborhood:

$$\min_{V,\mathbf{k}} \sum_{j=1}^{J} w_j \sum_{i=1}^{N} \left\| \hat{F}_j(\hat{\nu}_i) - v_{k_j}(\hat{\nu}_i) \right\|^2,$$

where:

- $\hat{F}_j(\hat{\nu}_i)$ is the value of the $\hat{\nu}_i$-th percentile for the j-th neighborhood, where
  - $j \in \{1, ..., J\}$ represents the neighborhoods
  - $\hat{F} = \{\hat{F}_1, ..., \hat{F}_J\}$ is the set of neighborhood vectors, where each $\hat{F}_j = \{\hat{F}_j(\hat{\nu}_1), ..., \hat{F}_J(\hat{\nu}_N)\}$ represents the distribution of test score changes for neighborhood j across the N percentiles (so each neighborhood j is a d-dimensional vector representing the test score changes at different discrete percentiles, $d = 40$.
- $v_{k_j}(\hat{\nu}_i)$ is the $\hat{\nu}_i$-th percentile of the cluster center assigned to the j-th neighborhood, where
  - $k \in \{1, ..., K\}$ represents the clusters
  - $\mathbf{k} = \{k_1, ..., k_J\}$ is the cluster assignment vector for neighborhoods, where $k_j \in \{1, ..., K\}$
  - $V = \{\mathbf{v_1}, ..., \mathbf{v_K}\}$ is the set of cluster centers, where each $\mathbf{v_k} = [v_k\hat{\nu}_1), ..., v_k(\hat{\nu}_N)]$ represents the average distribution of test score changes for cluster k across the N percentiles
- w_j is the weight for the j-th neighborhood, representing the number of children in the neighborhood

The distributional k-means algorithm is an iterative procedure consisting of the following two steps that come after the initialization.

1. *Cluster Assignment Step.* Update **k** by assigning each neighborhood to the nearest cluster center based on the entire distribution of test score changes:

$$k_j = \arg\min_k \sum_{i=1}^{N} w_j \left\| \hat{F}_j(\hat{\nu}_i) - v_k \hat{\nu}_i) \right\|^2$$

2. *Update Step.* Update cluster centers V by recalculating the cluster centers as the weighted mean of the neighborhoods assigned to each cluster across all N percentiles. The update formula for the $\hat{\nu}_i$-th percentile of the cluster center $\mathbf{v_k}$ is:

$$v_k(\hat{\nu}_i) = \frac{\sum_{j:k_j=k} w_j \hat{F}_j(\hat{\nu}_i)}{\sum_{j:k_j=k} w_j},$$

   where the summation is over all neighborhoods j that are assigned to cluster k (i.e., $k_j = k$). So, for each cluster k, update its center v_k to be the weighted average of all its assigned neighborhoods' test score changes. Larger neighborhoods have more influence through the weights w_j. And then normalize this weighted sum by the total weight of all neighborhoods assigned to cluster k. This ensures that new cluster center v_k represents average test score growth distribution of the neighborhoods in that cluster, weighted by the number of children.

4.2.2 Clustering Schools into Categories

In a similar vein as in the previous subsection, I consider the distributions of test score changes across all schools in the sample, and then minimize the following objective function to cluster the C schools into S categories, weighted by the number of children in every school c:

$$\min_{V,\mathbf{s}} \sum_{c=1}^{C} w_j \sum_{i=1}^{N} \left\| \hat{F}_c(\hat{\nu}_i) - v_{s_j}(\hat{\nu}_i) \right\|^2,$$

where $\hat{F}_c(\hat{\nu}_i)$ is the value of the $\hat{\nu}_i$-th percentile for the c-th school, $v_{k_j}(\hat{\nu}_i)$ is the $\hat{\nu}_i$-th percentile of the cluster center assigned to the c-th school, and w_c is the weight for the c-th school, representing the number of children in the school.

4.2.3 Aggregation into *(Neighborhood Class, School Category)-Places*

I then combine neighborhood classes and school categories so that I have different combinations of (neighborhood class, school category)-pairs. For example, with three neighborhood classes, $k \in \{A, B, C\}$, and three school categories, $s \in \{1, 2, 3\}$, this yields the following possible group combinations resulting in nine distinct places:

$$place_{it}(k, s) = \begin{cases} A1 & \text{if } k_{it} = A,\, s_{it} = 1 \\ A2 & \text{if } k_{it} = A,\, s_{it} = 2 \\ A3 & \text{if } k_{it} = A,\, s_{it} = 3 \\ B1 & \text{if } k_{it} = B,\, s_{it} = 1 \\ B2 & \text{if } k_{it} = B,\, s_{it} = 2 \\ B3 & \text{if } k_{it} = B,\, s_{it} = 3 \\ C1 & \text{if } k_{it} = C,\, s_{it} = 1 \\ C2 & \text{if } k_{it} = C,\, s_{it} = 2 \\ C3 & \text{if } k_{it} = C,\, s_{it} = 3 \end{cases}$$

Because clustering was based on the sample with stayers in the first period only, but key variation comes from children who make a transition, I then impose the place assignment identifiers to all child-time observations in the sample. In the next and last estimation step, I fit the core model using the Expectation Maximization (EM) algorithm to estimate $\alpha - p-$specific distributions of test score gains.

4.3 Stratified Finite Mixture Models

Combining the neighborhood class and the school category into pairs, q_{itjs} denotes the identifier of the (neighborhood j, school c)-pair associated with child i at time t. This results in heterogeneity across pairs of neighborhood and school types, so that $p_{it} = p(q_{it}) \in \{1, ..., P\}$ denotes the (neighborhood class k, school category s)-place.

In the same vein as the identification strategy, I follow Bonhomme et al. (2019) and first estimate test score growth densities using transitioners only, and then family-type proportions in the first period using both transitioners and stayers. Computation of the parameter vectors is executed using the Expectation-Maximization Algorithm (cf., Dempster et al. (1977)).

While the identification result for the models under discrete child heterogeneity is non-parametric, I consider a specification where children belong to L latent types, so that the model is parametric given child, school, and neighborhood heterogeneity. I let both test score growth densities be normal with $(p(k,s),\alpha)$-specific means and variances. That is, means and variances of test score changes are allowed to differ between all combinations of child types, neighborhood classes and school categories.

I illustrate the estimation process for the first stratified finite mixture model equation, *Equation 1*, where α is discrete and p stands for place – representing the (neighborhood class k, school category s)-place.

Mathematical representation of the bivariate stratified finite mixture model aka key Equation (1):

$$P(\hat{\nu}_1, \hat{\nu}_2|\Theta) = \sum_{\alpha=1}^{L} \pi_{\alpha,p\to p'} \mathcal{N}(\hat{\nu}_1|\mu_{\hat{\nu}_1\alpha p}, \sigma^2_{\hat{\nu}_1\alpha p}) \times \mathcal{N}(\hat{\nu}_2|\mu_{\hat{\nu}_2\alpha p'}, \sigma^2_{\hat{\nu}_2\alpha p'}),$$

where $\Theta = (\mu_{\hat{\nu}_1}, \mu_{\hat{\nu}_2}, \sigma^2_{\hat{\nu}_1}, \sigma^2_{\hat{\nu}_2}, \pi, p \to p')$.

Initialization:

$$\begin{aligned}
\mu_{\hat{\nu}_1\alpha p} &= \frac{\sum_{i=1}^{N\alpha p} \hat{\nu}_{i1,\alpha p}}{N_{\alpha p}} \\
\mu_{\hat{\nu}_2\alpha p'} &= \frac{\sum_{i=1}^{N\alpha p'} \hat{\nu}_{i2,\alpha p'}}{N_{\alpha p'}} \\
\sigma^2_{\alpha\hat{\nu}_1 p} &= \frac{\sum_{i=1}^{N_{\alpha p}} (\hat{\nu}_{i1,\alpha p} - \mu_{\hat{\nu}_1,\alpha p})^2}{N_{\alpha p}} \\
\sigma^2_{\alpha\hat{\nu}_2 p'} &= \frac{\sum_{i=1}^{N_{\alpha p'}} (\hat{\nu}_{i2,\alpha p'} - \mu_{\hat{\nu}_2,\alpha p'})^2}{N_{\alpha p'}} \\
\pi_{\alpha,p\to p'} &= \frac{N_{\alpha,p\to p'}}{N}
\end{aligned}$$

Expectation:

$$P(\hat{\nu}_{i1}, \hat{\nu}_{i2} \in \alpha_l|\hat{\nu}_{i1}, \hat{\nu}_{i2}) = \frac{P(\hat{\nu}_{i1|\hat{\nu}_{i1}\in\alpha_l,p\to p'}) \cdot P(\hat{\nu}_{i2}|\hat{\nu}_{i2} \in \alpha_l, p \to p') \cdot P(\alpha_l, p \to p')}{P(\hat{\nu}_{i1}, \hat{\nu}_{i2}|p \to p')}$$

$$\Leftrightarrow \underbrace{\gamma(z_{i\alpha|p\to p'})}_{\text{posterior}} = \frac{\mathcal{N}(\hat{\nu}_{i1}|\mu_{\hat{\nu}_1\alpha_l,p\to p'}, \sigma^2_{\hat{\nu}_1\alpha_l,p\to p'}) \times \mathcal{N}(\hat{\nu}_{i2}|\mu_{\hat{\nu}_{i2}\alpha_l p'}, \sigma^2_{\hat{\nu}_{i2}\alpha_l p'}) \times \pi_{\alpha_l|p\to p'}}{\sum_{\alpha=1}^{L} \mathcal{N}(\hat{\nu}_1|\mu_{\hat{\nu}_1,\alpha,p\to p'}, \sigma^2_{\hat{\nu}_1,\alpha,p\to p'}) \times \mathcal{N}(\hat{\nu}_2|\mu_{\hat{\nu}_{i2},\alpha,p'}, \sigma^2_{\hat{\nu}_{i2},\alpha,p'}) \times \pi_{\alpha|p\to p'}}$$

Maximization:

$$\mu_{\hat{\nu}_1,\alpha,p\to p'} = \frac{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})\times\hat{\nu}_{i1}}{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})}$$

$$\mu_{\hat{\nu}_2,\alpha,p\to p'} = \frac{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})\times\hat{\nu}_{i2}}{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})}$$

$$\sigma^2_{\hat{\nu}_1,\alpha,p\to p'} = \frac{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})\times(\hat{\nu}_{i1}-\mu_{\hat{\nu}_1,\alpha,p\to p'})^2}{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})}$$

$$\sigma^2_{\hat{\nu}_2,\alpha,p\to p'} = \frac{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})\times(\hat{\nu}_{i2}-\mu_{\hat{\nu}_2,\alpha,p\to p'})^2}{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})}$$

$$\pi_{\alpha,p\to p'} = \frac{\sum_{i=1}^{N}\gamma(z_{i\alpha|p\to p'})}{N}$$

Convergence:

$$\sum_{i=1}^{N_p}\sum_{p=1}^{P}\sum_{p'=1}^{P}\mathbb{1}\{\hat{p}_{i1}=p\}\mathbb{1}\{\hat{p}_{i2}=p'\}\ln\left(\sum_{a=1}^{L}\pi_{kk'}(\alpha;\theta_p)f_{p\alpha}(\hat{\nu}_{i1};\theta_f)f^{p}_{p'\alpha}(\hat{\nu}_{i2};\theta_{fp})\right)$$

5 Results

In this section, I report the empirical results from mapping the model to data from North Carolina, USA. First, I document the resulting test score gains that are estimated as first differences. Second, I present the results from the pre-clustering step, where I estimate both neighborhood classes and school categories with distributional k-means algorithms, and then combine them into pairs. Third, I show match-specific estimation results from the stratified finite mixture model.

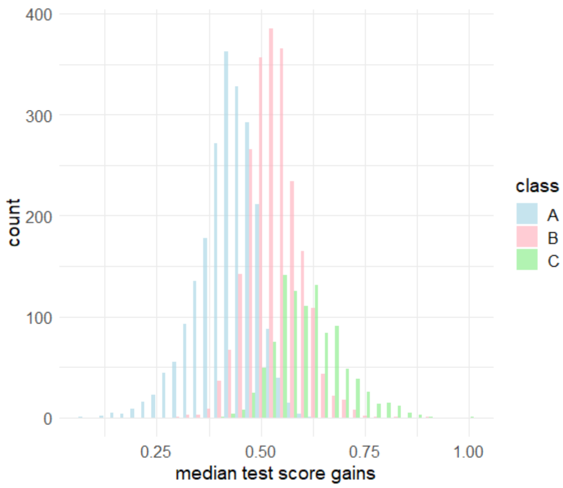
5.1 Value-Added Results

I estimate test score gains for period 1 as first differences between periods 0 and 1, and test score gains for period 2 as first differences between periods 1 and 2. Test score gains then represent test score performance not explained by lagged test scores. Summary statistics for period 1 and period 2 are presented in the top panel and the bottom panel, respectively, of Table 8 in *Appendix A*. The estimates indicate that within North Carolina, those who remain in the same neighborhood and school ("stayers") throughout the three considered periods are, on average, not very different from those who make a transition to a different neighborhood and/or school ("transitioners") between periods 1 and 2. In period 2, test score gains are slightly lower relative to period 1, but for both stayers and transitioners. This implies that, on average, compared to stayers, transitioning children do not experience significant differences in test score gains when moving to a different neighborhood or school. The similarity in gains between stayers and transitioners suggests that factors beyond mobility, such as individual and family characteristics or the specific nature of the transition, may play a more influential role in determining test score performance.

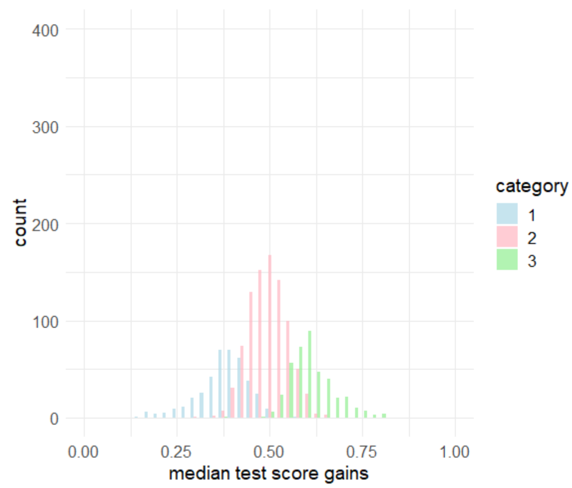
5.2 Clustering Results Based on Distributional K-Means

I estimate both neighborhood classes and school categories separately using a distributional k-means algorithm à la Bonhomme et al. (2019). The classification focuses on period 1 and includes only stayers who do not change their school- or neighborhood-identifiers throughout the three periods $(0, 1, 2)$ considered. I calculate the neighborhoods' cumulative distribution functions (cdfs) of period 1 test score gains on a grid of 40 percentiles of the overall distribution of test score gains. I do the same for schools' cdfs. The k-means algorithm is weighted by neighborhood size and school size, respectively. My baseline number of neighborhood classes is $K = 3$, where $K \in \{A, B, C\}$, and my baseline number of categories is $S = 3$, where $S \in \{1, 2, 3\}$. I then pair all possible neighborhood class and school category permutations into places $P = 9$, where $P \in \{A1, A2, A3, B1, B2, B3, C1, C2, C3\}$.

Classification of neighborhoods and schools. I classify neighborhoods into classes and schools into categories. Figure 3 displays histograms of the median test score gains across the three neighborhood classes and school categories. Both neighborhood classes and school categories are ordered by their mean test score gains within each group.



(a) Neighborhood Classes



(b) School Categories

Figure 3: Histograms of median test score gains by cluster

This figure visualizes results for the classification of neighborhoods into classes (Figure 3a) and schools into categories (Figure 3b). A is the neighborhood class with relatively lowest mean test score gains and C is the neighborhood class with relatively highest mean test score gains; 1 *is the school category with relatively lowest mean test score gains and* 3 *is the school category with relatively highest mean test score gains. The histograms show the distribution of median test score gains across the three neighborhood classes and school categories, using* 40 *bins from* 0 *to* 1*. Each bar represents the number of neighborhoods or schools (vertical axis) within a specific range of median test score gains (horizontal axis) for each class or category. For example, in Figure 3b, the highest bar (red) can be interpreted as representing approximately* 170 *schools assigned to category* 2 *with median test score gains around* 0.5.

The visualizations in Figure 3 indicate that neighborhood classes are notably larger than school categories, reflecting the higher number of neighborhoods compared to schools in the sample. The two classifications also display subtle distribution differences: neighborhood classes exhibit a somewhat more uniform distribution, while school categories show slightly greater heterogeneity across clusters. Within clusters, school categories appear marginally more homogeneous than neighborhood classes. Table 2 provides a detailed summary, presenting the number of neighborhoods within each class and the number of schools in each category, along with test score gains across different quartiles.

Table 2: Cluster Summary Statistics

Class	#Neighborhoods	Q1	Median	Q3
A	$2,177$	0.39	0.43	0.47
B	$2,238$	0.49	0.53	0.57
C	$1,008$	0.55	0.60	0.67

(a) Neighborhood Classes

Category	#Schools	Q1	Median	Q3
1	568	0.37	0.41	0.44
2	745	0.47	0.51	0.54
3	377	0.57	0.60	0.65

(b) School Categories

The summary statistics of stayers in Table 2 reveal that the median test score gains of neighborhood class C and school category 3 are $17 - 19$ percentage points $(40 - 46\%)$ higher than those of neighborhood class A and school category 1, and $7 - 9$ percentage points $(13 - 18\%)$ higher than those of neighborhood class B and school category 2. Next, rather than examining them separately, I analyze neighborhood classes and school categories together, focusing on how these pairings are distributed among the non-transitioning children in my sample.

Pairing neighborhood classes and school categories. Given the resulting three neighborhood classes and three school categories, I pair them into nine possible combinations of (neighborhood class, school category)-places. Figure 4 presents the mean test score gains as well as the sizes for each of them, ordered according to mean test score gains in each (neighborhood class, school category)-place possible.

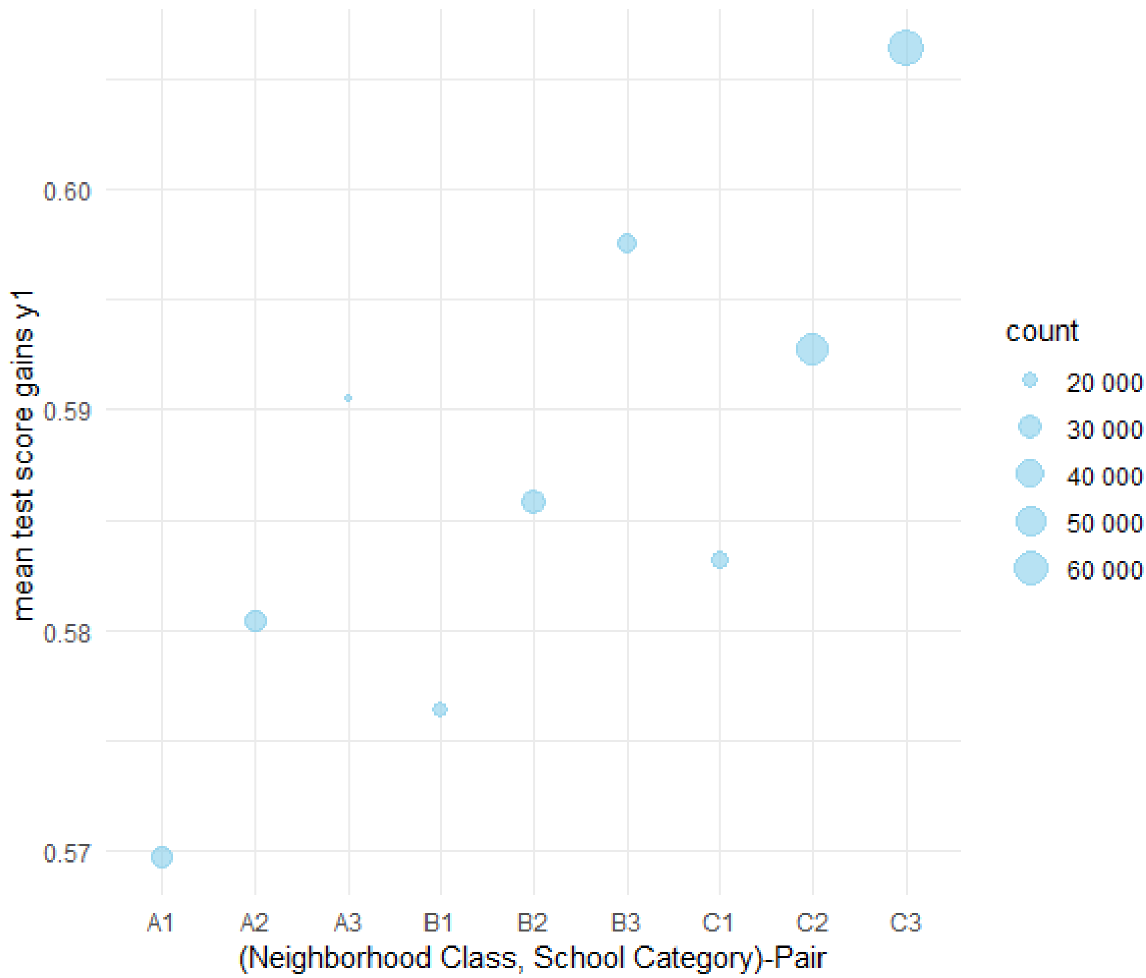


Figure 4: Mean test score gains and size of cluster-pairs

This figure shows the count of non-transitioning children in all neighborhood class and school category pairings, with the vertical axis reflecting their mean test score gains. Larger bubbles indicate higher counts of children within a particular (neighborhood class, school category)-place.

The bubble sizes in Figure 4 indicate that most children tend to sort into one of two extremes – belonging either to neighborhoods and schools with relatively low test score gains (left two bubbles) or those with relatively high test score gains (right two bubbles) – or fall into an intermediate neighborhood paired with an intermediate school (center bubble).[18] In particular, most children reside in a neighborhood of class C and attend

[18]Note that despite having fewer neighborhoods in class C and fewer schools in category 3 relative to other clusters (cf., Table 2), the $C2$ and $C3$ pairings form the largest bubbles (cf., Figure 4), indicating

a school of category 2 or 3 (i.e., $C2$ and $C3$). However, a significant number also live in neighborhoods and attend schools with relatively lower test score gains ($A1$ or $A2$). Table 4 provides a more detailed summary statistics with numbers, and additional visualizations are reported in Figure 9 in *Appendix B2*.

The distribution indicates a positive relationship between specific neighborhood-school combinations and student academic performance, with higher test score gains observed in higher-performing neighborhood and school pairings. Furthermore, in neighborhoods with relatively lowest test score gains (A), few children attend a school with relatively high test score gains (3). In contrast, in neighborhoods with relatively high test score gains (C), most children attend a school with relatively high test score gains (2 or 3), suggesting that access to higher-category schools may be more limited in lower-performing neighborhoods. This aligns with Agostinelli et al. (2024)'s observation that schools with higher-skilled peer groups are predominantly located in high-income neighborhoods and often lie outside the choice set for lower-income communities.

Conditional on the class of the neighborhood, results show a clear positive effect of attending a school of a relatively higher category (i.e., with relatively higher mean test score gains) on a child's performance. This suggests that school quality, measured by mean test score gains, plays a critical role in shaping student outcomes, even when controlling for the neighborhood in which the child resides. However, Figure 4 demonstrates the importance of accounting for both neighborhoods and schools: mean test score gains for the same school category are monotonically increasing with the neighborhood class. This indicates that the interaction between neighborhoods and schools plays a crucial rule. Even within the same school category, the neighborhood context seems to influence school effectiveness. A school of a certain category in a neighborhood of a particular class produces different test score gains than a school of that same category in a neighborhood of a different class. Thus, for the same school category, the neighborhood context can amplify or limit school effectiveness.[19]

higher population densities in high-gain neighborhoods and schools.

[19] Even though neighborhood classes came out slightly more homogeneous across classes compared to school categories (cf., Figure 3), the same school category can lead to different outcomes depending on their related neighborhood class. And recall that the categorization of schools was done independently of their geographic location (i.e., neighborhood characteristics).

Table 4: Places Summary Statistics

Place	Count	Mean	SD	Q1	Median	Q3
A1	27,639	0.57	0.06	0.53	0.57	0.61
A2	27,733	0.58	0.06	0.54	0.58	0.62
A3	14,411	0.59	0.06	0.55	0.59	0.63
B1	19,578	0.58	0.06	0.54	0.58	0.62
B2	33,557	0.59	0.06	0.55	0.59	0.63
B3	23,641	0.60	0.06	0.56	0.60	0.64
C1	20,241	0.58	0.06	0.54	0.58	0.62
C2	54,981	0.59	0.06	0.55	0.59	0.63
C3	67,291	0.61	0.06	0.57	0.61	0.65

Table 4 highlights that in neighborhoods of relatively lowest class, only 21% of children attend a school of highest category, and 40% attend a school of lowest category. In neighborhoods of relatively highest class, only 14% of children attend a school of relatively lowest category and 47% attend a school of relatively highest category.

Taking the assignments of neighborhoods into classes and schools into categories as given and fixed ("hard-clustered"), I then impose the corresponding (neighborhood class, school category)-place assignments that result from the pre-classification with stayers on the transitioning children, by matching their neighborhood and school identifiers. Conditional on transitioning children's mobility pattern (their neighborhood classes and school categories), the latent family types can then be derived probabilistically with a stratified finite mixture model ("soft-clustering"), revealing matched family-school-neighborhood estimates. Next, I turn to these key match-specific estimation results.

5.3 Match-Specific Results Based on Estimated Parameters

Conditional on the hard-clustered neighborhood classes and school categories, I estimate the latent family types probabilistically using a Gaussian finite mixture model. The resulting family-school-neighborhood match-specific estimates reveal the distribution of test score gains for all (neighborhood class, school category, family type)-combinations for every period.

My baseline estimates are based on a finite mixture model with $L = 3$ types of children. Based on the sample of children who make a transition between periods 1 and 2, I estimate distributions of test score gains for all type-specific family-school-neighborhood matches. Based on the cross-section in period 1, I also estimate the share of family types in all (neighborhood class, school category)-places.

Heterogeneity in neighborhood treatment effects and school value-added. In Figure 5, I plot estimates of the mean test score gains for each type-specific family-school-neighborhood match. I order the nine (neighborhood class, school category)-places on the x-axis by mean test score gains. On the y-axis, I report the estimates of mean test score gains for all $L = 3$ family types (colored lines). Family types are also ordered by mean test score gains, so that α_1 (red) represents the type with lowest test score gains relative to all family types, and α_3 (green) represents the type with highest test score gains relative to all family types.

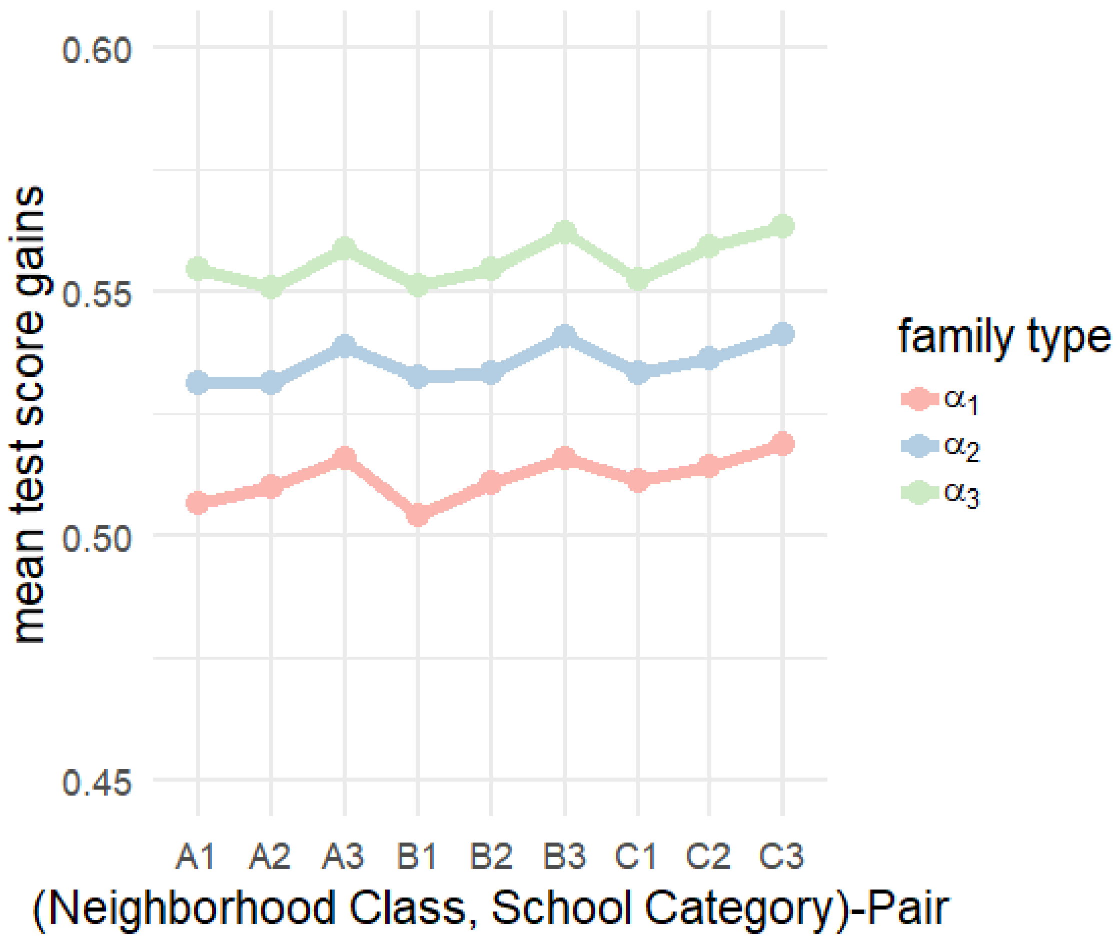
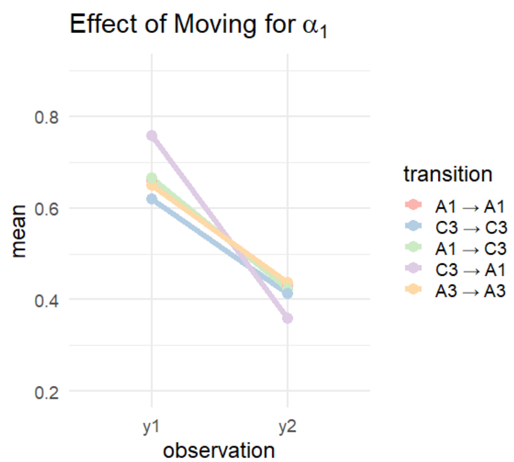


Figure 5: Family-school-neighborhood match-specific effects

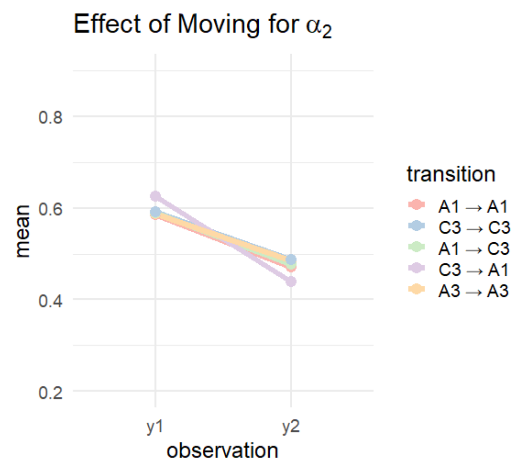
The results show family-school-neighborhood match-specific effects pointing to heterogeneity in neighborhood treatment effects and school value-added. The plot reveals strong evidence of family heterogeneity and significant variation in test score gains across school categories and across neighborhood classes. Notably, children seem to benefit most from attending schools with high average test score gains (3), especially those children with lower average test score gains themselves (α_1). Estimates suggests an important complementarity between a school with high test score gains and children at the lower end of the test score distribution. If a child of family type α_1 attends a school of category 3, then they can almost catch up to the next type, α_2, in terms of test score gains, making up most of the disadvantages coming from the family. The same holds for a child of type α_2. These findings indicate that the observed variation in test score gains across neighborhoods of different classes and schools of different categories is partly attributable to differences in test score gains for a given child type.

The interaction between neighborhoods and schools plays a key role. For example, school category 3 in neighborhood class A produces lower mean test score gains than it does in neighborhood class C, for all family types. Since schools of the same category tend to perform differently depending on the neighborhood they serve, this indicates a contextual complementarity where the neighborhood environment can either enhance or limit a schools' effectiveness. Thus, to maximize gains, policies may need to consider both school quality, as measured by mean test score gains, and neighborhood context together, as their interaction appears crucial for optimizing educational outcomes.

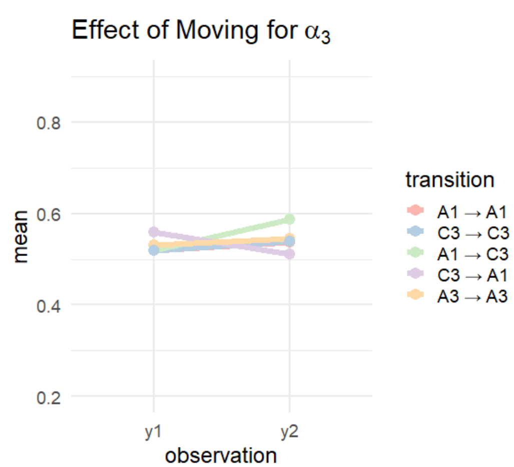
Heterogeneity in the Effect of Moving. To investigate potential heterogeneity in the impact of mobility on children's test score gains, I focus on the three possible child types separately. I then obtain the mean test score gains for the period before the move ($y1$) and the period after the move ($y2$) for different transitions. Figure 6 displays three plots, with the top left (a) showing the effect of moving for a child of type α_1, which represents children with the relatively lowest test score gains. The top right plot (b) shows the effect of moving for a child of type α_2. The bottom plot (c) shows the effect of moving for a child of type α_3.



(a)



(b)



(c)

Figure 6: Effect of Moving for Different Child Types α_i

This figure shows three plots, with the top left (a) showing the effect of moving for a child of type α_1, which represents children with the relatively lowest test score gains. The top right plot (b) shows the effect of moving for a child of type α_2. The bottom plot (c) shows the effect of moving for a child of type α_3. The letters and numbers in the legend, indicating the type of transition, are also in increasing order according to average test score gains. For example, A1 represents the (neighborhood class, school category)-place with relatively lowest test score gains and C3 the one with relatively highest test score gains. A1 $\rightarrow$ A1 indicates that the transition is from a neighborhood of class A and a school of category 1 to a another neighborhood of class A and/or school of category 1.

Results in Figure 6 suggest heterogeneity in the effect of moving. For a child with relatively lowest overall test score gains (α_1), shown in the top left plot, a transition consistently results in a sharp decline in their gains across all transitions, suggesting that this group is particularly sensitive to changes in their environment. Children of type α_2, shown in the top right plot, show a more moderate decline in test score gains after a transition, indicating a slightly more resilient response to environmental changes. Interestingly, in both graphs on the top, children who transition from the place with relatively highest test score gains to the place with relatively lowest test score gains ($C3 \rightarrow A1$), experience the largest decrease in test score gains. Finally, children of the type with relatively highest test score gains overall, α_3, experience the least disruption, with some transitions even showing a small increase in test score gains after moving. Especially children who transition from the bottom place to the top place in terms of test score gains ($A1 \rightarrow C3$), experience the largest gain in test scores. While Figure 5 reveals positive interaction effects in high-gain environments, Figure 6 suggests that mobility policies should be approached with caution. Given the heterogeneous effects, such policies must be carefully targeted to support vulnerable lower-performing children (α_1) while enabling higher-performing children (α_3) to benefit from improved environments.

Shares of family types (sorting). To get a sense of how different types of families self-select themselves into different types of neighborhoods and schools, in Figure 7, I plot the distribution of family types in the cross-section in period 1, before a potential transition.

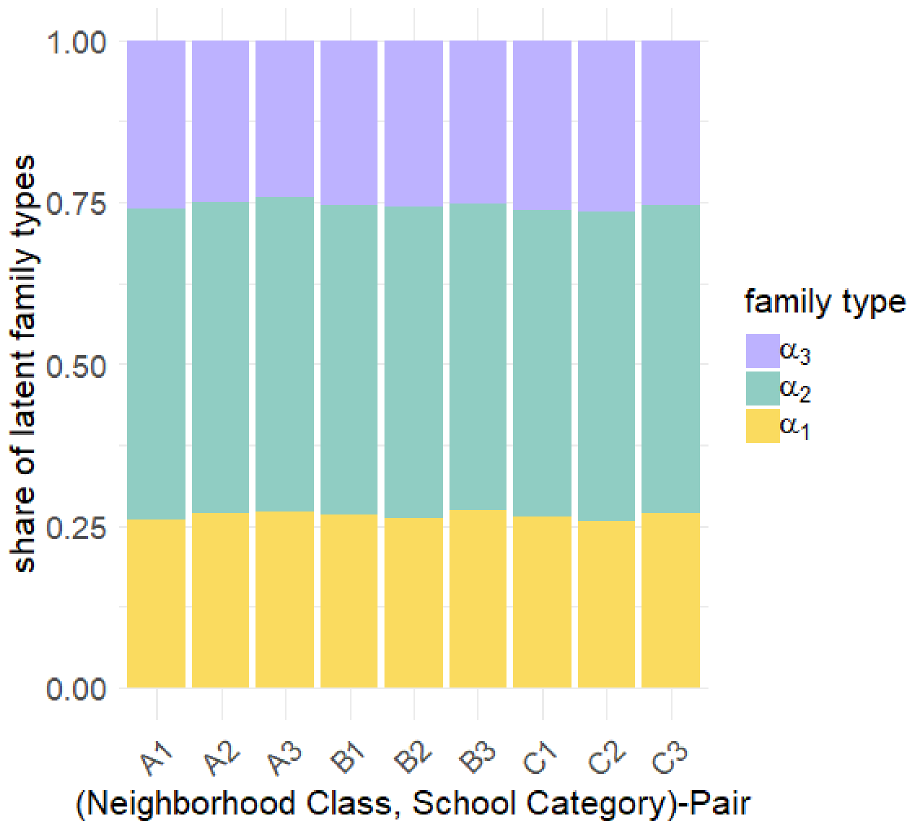


Figure 7: Share of latent family types (sorting)

Estimates in Figure 7 suggest that there is essentially no sorting of family types into environments. The share of children achieving different test score gains is constant and comparable across different (neighborhood class, school category)-places. Even though Figure 4 highlighted that most people sort into the extremes ($A1, A2, C2, C3$) or the very middle ($B2$), this implies that the share of latent child types is comparable across different types of neighborhoods and schools. There are two notable features. First, there might not be huge variation in gains because, for this sorting figure, I am looking at period 1 and only consider those who were already in the exact same environment in period 0 ("stayers"). Second, note that these are estimates from a soft-clustering approach, meaning that children do not get assigned to types with certainty, but only with a probability.[20] However, when doing the same exercise based on test score *levels*, I do find significant sorting patterns (Cf., *Appendix B1*).

[20]For example, suppose there are two children in a place and two latent types are possible. One child has a 70%-probability of being assigned to type 1 and the other child has 30%-probability of being assigned to type 1, then that place would consist of 50% type 1 and 50% type 2 children.

6 Counterfactual Analyses

My model estimates shed light on key determinants of variation in test scores. Conditional on how families sort into neighborhoods and schools, estimates uncover the causal effects of different types of neighborhoods and schools on different types of children's test score gains. Results suggest both heterogeneous neighborhood treatment effects and heterogeneous school value-added. The same type of child experiences higher test score gains if they reside in a neighborhood and attend a school with relatively high average test score gains. This suggests that neighborhoods and schools with high average test score gains positively complement children's learning, particularly improving the performance of those at the lower end of the test score distribution.

In this section, I further explore the presence and extent of heterogeneous neighborhood treatment effects and school value-added, as well as sorting, with two different types of policies. The first one is a place-based policy and investigates the distributional impact of an improvement in school quality, holding neighborhoods fixed. The second exercise consists of counterfactual scenarios to address inequality of both access and quality of neighborhoods and schools, by randomly reallocating children to schools and/or neighborhoods. I aim to quantify the extent of sorting and the importance of interaction effects between families, schools, and neighborhoods, and how these factors contribute to differences in educational outcomes. While these random reallocation exercises do not directly simulate people- or place-based policies, they highlight sources of variation in test score gains, and can inform the effectiveness of policies involving mobility. For example, a common people-based policy is to offer housing vouchers to low-income families, allowing them to move to high-quality neighborhoods and enroll their children in high-performing schools (e.g., Agostinelli et al. (2024), Chetty et al. (2016); Chyn (2018)). Another popular people-based policy is busing, that is driving children, who were enrolled in a school with low average test scores, to a different neighborhood on a daily basis, so that they can attend a school with high average test scores (e.g., Agostinelli et al. (2024); Angrist et al. (2024)).

6.1 School Quality Improvement

I investigate the distributional impact of a place-based policy to address inequality of educational access and quality depending on where children reside in North Carolina. By fixing neighborhoods, I conduct a counterfactual scenario with an improvement in school quality of those schools with relatively low test score gains. The increase in school quality is simulated by imposing the test score gains distribution of schools of a higher category on schools of category 1, which are those with relatively lowest test score gains. The following Table 6 shows results if, ceteris paribus, school quality was improved by imposing the test score gains distributions of schools of category 2 and schools of category 3, respectively, on schools of category 1.

Table 6: Improvement in school quality

Statistic	***a)*** $\boldsymbol{\Delta_{cat1 \rightarrow cat2}}$	***b)*** $\boldsymbol{\Delta_{cat1 \rightarrow cat3}}$
Mean	0.23%	0.48%
1%-Quantile	0.35%	0.56%
5%-Quantile	0.36%	0.55%
10%-Quantile	0.29%	0.50%
20%-Quantile	0.29%	0.51%
30%-Quantile	0.27%	0.49%
40%-Quantile	0.24%	0.45%
Median	0.21%	0.45%
60%-Quantile	0.18%	0.45%
70%-Quantile	0.20%	0.46%
80%-Quantile	0.17%	0.44%
90%-Quantile	0.17%	0.48%
95%-Quantile	0.16%	0.48%
99%-Quantile	0.13%	0.50%
Variance	-0.007	-0.003

Note: Differences in test score gains between the simulated datasets with upgraded school quality and the original dataset, holding neighborhoods fixed (variances is $\times 100$*). The middle column of the table (a) shows how the entire test score gains distribution would change if there is a moderate improvement of school quality, so that schools of category* 1 *become schools of category* 2*,* $\Delta_{cat1 \rightarrow cat2}$*. The last column (b) shows these outcomes if school quality is even higher so that schools of category* 1 *are replaced with schools of category* 3*,* $\Delta_{cat1 \rightarrow cat3}$*.*

The positive differences in statistics between the simulated counterfactual scenarios with upgraded school quality and the original school environment in Table 6 quantify the extent by which test score gains increase when upgrading the category of schools of category 1. While the entire distribution would gain from an improvement in school quality, it would be most beneficial for those at the lower end of the test score distribution. In particular, gains are monotonically increasing in school quality. Compared to the original scenario, performance improves substantially more across the entire distribution when test score gains of schools of category 3 are imposed on schools of category

1. Gains of the upper part of the test score distribution improve more than twice as much when schools of category 1 become schools of category 3 as opposed to becoming schools of category 2.

From a policy perspective, the potential benefits of improving the school quality of lower-performing schools have important implications, as this may help reducing the gap between lower and higher achievers. Because gains are relatively higher for those at the lower end of the distribution, this suggests that the positive effects of upgrading school quality are more pronounced for relatively lower-performing students. However, the near-zero variance suggests that improvements in school quality alone may be insufficient to close the test score gap or meaningfully reduce disparities, particularly if ignoring in which neighborhoods these schools' students are located.[21] Accounting for the interaction between school quality and neighborhood conditions could offer a more comprehensive understanding of how these factors jointly affect student performance.[22]

6.2 Reallocation of Children to Schools and/or Neighborhoods

I now aim to quantify the extent of sorting and examine the interaction effects between families, schools, and neighborhoods, assessing how these factors contribute to differences in educational outcomes. The results of the model estimates in Figure 5 suggest positive interaction effects for children living in neighborhoods and attending schools with high average test scores. I now use the estimated model to simulate counterfactual environments, shedding light on the extent and consequences of sorting into neighborhoods and schools, as well as the interdependencies between them. This analysis will explore how interactions between neighborhoods and schools may influence the results, particularly when access to certain schools is restricted to specific neighborhoods.

I assess the interdependencies between neighborhoods and schools and the contribution of sorting to the distribution of test score gains in three ways. I begin by analyzing how the test score gains distribution would change if there was a random reallocation of children to both neighborhoods and schools. This will reveal how the distribution of test score gains would change if there was a random reallocation of the classes of

[21] Although, I want to note that the overall effect of an improvement in school quality depends on the proportion of children in category 1 compared to the total population. If category 1 represents a smaller fraction of the total, then the aggregate effect might not be that large. The initial distribution of test score gains across all categories also plays a role. If the other categories already have high test score gains, the relative improvement from upgrading category 1-schools might seem less dramatic in the context of the entire distribution. Thus, policymakers may want to consider the broader distribution and target interventions where they can achieve the most significant aggregate impact.

[22] I also conduct a simulation where, instead of improving school quality, I improve the neighborhood quality, taking school quality as given. Details are reported in Table 9 in *Appendix D*.

the neighborhoods where children live as well as the categories of the schools they attend. I then investigate how the test score gains distribution would change if there was a random reallocation of children to neighborhoods only, holding school assignments fixed. Lastly, I analyze how the test score gains distribution would change if there was a random reallocation of children to schools only, holding neighborhood assignments fixed. Table 6 presents differences in test score gains between the simulated datasets with random reallocation and the original dataset.[23]

Table 7: Random Reallocation Exercises

Statistic	*a*) **random both**	*b*) **random nbhood**	*c*) **random school**
Mean	0.29% (0.001)	0.23% (0.017)	−0.02% (0.009)
1%−Quantile	0.05% (0.004)	0.09% (0.016)	−0.17% (0.010)
5%−Quantile	0.15% (0.002)	0.17% (0.016)	−0.06% (0.009)
10%−Quantile	0.18% (0.002)	0.21% (0.016)	−0.05% (0.009)
20%−Quantile	0.21% (0.001)	0.21% (0.016)	−0.06% (0.009)
30%−Quantile	0.22% (0.001)	0.22% (0.017)	−0.05% (0.009)
40%−Quantile	0.28% (0.001)	0.23% (0.017)	−0.04% (0.009)
Median	0.30% (0.001)	0.24% (0.017)	−0.04% (0.009)
60%−Quantile	0.32% (0.001)	0.26% (0.017)	−0.01% (0.009)
70%−Quantile	0.35% (0.001)	0.28% (0.018)	0.02% (0.009)
80%−Quantile	0.38% (0.001)	0.26% (0.018)	0.01% (0.009)
90%−Quantile	0.41% (0.002)	0.25% (0.019)	0.02% (0.009)
95%−Quantile	0.37% (0.002)	0.18% (0.019)	0.02% (0.008)
99%−Quantile	0.39% (0.005)	0.14% (0.020)	0.08% (0.007)
Variance	0.010 (0.000)	0.002 (0.000)	0.005 (0.000)

Note: The table shows differences in means, quantiles, and variances of test score gains between the simulated datasets with random reallocation and the original dataset. The results are obtained using 100 simulations (non-parametric bootstrap standard errors (×100) computed from these simulations are in parentheses).The first column (a) shows how the test score gains distribution would change if there was a random reallocation of children to both neighborhoods and schools. The middle column of the table (b) shows how test score gains would change if there was a random reallocation of children to neighborhoods only; and the last column (c) shows how test score gains would change if there was a random reallocation of children to schools only.

In the first counterfactual exercise in Table 6 (a), I find positive impacts. The entire distribution would benefit from a random reallocation of children to both neighborhoods and schools, especially those at the upper end of the test score distribution. The positive

[23]Note that I do not allow for equilibrium effects in these counterfactuals (test score gains functions for child types, school categories, and neighborhood classes are not affected).

effect of the reallocation on mean test score gains (0.29%) suggests that neighborhood and school treatment heterogeneity seem to have a statistically significant impact on average test score gains. Similarly as in the labor literature (Bonhomme et al. (2019)), the presence of these complementarities, although small, would not be consistent with an additively separable economy between children and places (Graham et al. (2014)). Given the strongest impact on the upper quantiles, it seems that complementarities are especially present for children at the upper end of the test score gains distribution. Moreover, the slight increase in the variance of test score gains suggests that random reallocation would lead to more dispersed outcomes, as children at the upper end of the distribution tend to benefit more.

The second counterfactual exercise, where I fix schools and randomly reallocate children to neighborhoods only (b), reveals similar results to the first simulation. Since there is limited sorting based on test score gains (cf., Figure 7), which is not true for levels (cf., Figure 8b), simulations show that the entire distribution would benefit from a random reallocation of children to neighborhoods (with or without randomizing schools), and test score gains would be slightly more dispersed. When randomizing neighborhoods (either with or without randomizing schools), the outcomes seem largely driven by neighborhood changes. This suggests that, in the context of North Carolina, neighborhood environments play a significant role in the variation of test score gains. A combined reallocation does not necessarily add much more benefit than neighborhood reallocation alone because neighborhood context already shapes school effectiveness to a large degree. Neighborhoods seem to fundamentally alter the context in which schools work, affecting student outcomes, even when the school category remains unchanged. This unified perspective is crucial for understanding and maximizing the impact of any educational interventions.

However, the third counterfactual exercise, where I fix neighborhoods and randomly reallocate children to schools only (c) shows a different pattern. While those at the bottom of the test score gains distribution would be hurt, those at the upper part of the distribution would slightly benefit. Test score gains would also be slightly more dispersed, as reflected in the positive variance. Although in general, effects are relatively small in the third exercise. Random reallocation of schools only – holding neighborhoods fixed – has the smallest effect because, with neighborhoods held constant, children do not experience the broader contextual change that neighborhoods provide. Without the neighborhood shift, children's academic experiences remain largely shaped by their constant neighborhood environment. The negligible impact of reallocation on mean test score gains (-0.02%) aligns with Kinsler (2016)'s conclusion that teacher complementarities across grades are minimal. His finding is based on a value-added model incorporating interactive effects between teachers in North Carolina's primary schools.

Overall, the three reallocation effects demonstrate that neighborhood context is foundational and can either unlock or limit the benefits that different school categories provide. Since neighborhoods and schools are typically aligned in terms of quality, changing neighborhoods introduces a shift in the entire environmental package. The results indicate that reallocation of neighborhoods effectively reallocates the overall quality of both environments, which is why neighborhood changes have a stronger impact in my findings. Reallocating schools alone, without changing neighborhoods, does not offer the same level of transformative impact because it misses the complementary boost that neighborhood diversity brings. Neighborhood effects appear stronger because they create the baseline conditions that influence how effective a school can be. Analyzing schools and neighborhoods as unified influence allows for a better understanding of the combined effect of both environments on a child. Since schools do not operate in isolation, their effectiveness is partly determined by the neighborhoods they serve. Ignoring this could obscure the true impact each has, especially since neighborhood effects can intensify or weaken school effectiveness, depending on the match between the two.

Place-based policies aimed at improving neighborhood environments could have a significant impact on educational outcomes. This could include investments in community resources, safety, and overall living conditions. But also people-based policies such as housing mobility programs, that provide vouchers and support for disadvantaged families to move to better neighborhoods, might be effective in enhancing educational equity and performance.

7 Conclusion

Many factors shape a child's outcomes, and extensive research has examined the effects of family, school, neighborhood, or combinations of these influences. This paper develops a new approach to consider all three sources in a unified way to explore their combined impact on variation in test score gains. The results highlight the central role of family in academic performance, independent of environmental factors. However, a high-gain environment offers additional benefits: conditional on residence, school quality has a significant impact on academic progress, while the value schools contribute remains closely linked to their geographic location. My findings have important implications for understanding the drivers of test score disparities and for identifying effective strategies to narrow these gaps. Results suggest that targeted, people-based policies could be effective in supporting children from disadvantaged backgrounds or those with lower test score gains. By focusing resources on boosting gains for children at the lower end of the distribution, such policies can create crucial opportunities for upward mobility.

My analysis opens avenues for several meaningful extensions. First, I abstract from level differences in initial performance, which can influence the extent of achievable gains and may impact the estimation of family types (Carneiro and Heckman, 2003). Second, my framework could be enriched with different sources of sorting of families into neighborhoods and schools (Lentz et al., 2023), or time-varying family types (Lentz et al., 2024), as this might affect the estimation of the child's type when movements are related to social disruptions such as parental separation (Disch, 2024). Third, I am abstracting from spillovers and equilibrium effects in my random reallocation exercises by assuming that test score gains functions, for all child types and places, are not affected by the reallocation. While this may allow for a focused analysis of place effects on test scores, in order to capture the full complexity of real-world dynamics, future research may want to build upon these initial findings with more comprehensive models that include equilibrium effects (Agostinelli et al. (2024); Eckert and Kleineberg (2024)). Fourth, instead of reducing the heterogeneity of neighborhoods and schools through pre-clustering, I believe integrating it with family types into the same likelihood function is a natural, albeit challenging, way to enhance my analysis (Lentz et al., 2023). Fifth, it would also be interesting to look at more long-term effects such as outcomes in log-earnings (Chetty et al., 2016). Lastly, since parents care about school characteristics when presented with a choice set (Abdulkadiroğlu et al. (2020); Agostinelli et al. (2024)), a deeper investigation into the distinct contributions of neighborhoods and schools to educational outcomes emerges as a critical and policy-relevant research direction. I consider these all to be promising directions for future research.

References

Abdulkadiroğlu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020). Do parents value school effectiveness? *American Economic Review 110*(5), 1502–39.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Agostinelli, F., M. Doepke, G. Sorrenti, and F. Zilibotti (2024). It takes a village: The economics of parenting with neighborhood and peer effects. *Journal of Political Economy, forthcoming*.

Agostinelli, F., M. Luflade, and P. Martellini (2024). On the spatial determinants of educational access. *NBER Working Paper, No. 32246*.

Agostinelli, F. and M. Wiswall (2024). Estimating the technology of children's skill formation. *Journal of Political Economy forthcoming.*

Airoldi, E. M., D. M. Blei, S. E. Fienberg, and E. P. Xing (2008). Mixed membership stochastic blockmodels. *Journal of Machine Learning Research 9*, 1981–2014.

Andrews, M., L. Gill, T. Schank, and R. Upward (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society 171*(3), 673–697.

Angrist, J., G. Gray-Lobe, C. Idoux, and P. A. Pathak (2024). Still worth the trip? school busing effects in boston and new york. *Working Paper*.

Angrist, J., P. Hull, P. A. Pathak, and C. R. Walters (2024). Race and the mismeasure of school quality. *American Economic Review: Insights 6*(1), 20–37.

Arcidiacono, P. and J. B. Jones (2003). Finite mixture distributions, sequential likelihood, and the em algorithm. *Econometrica 71*(3), 933–946.

Arcidiacono, P. and R. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *American Economic Review 79*(6), 823—-1867.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy 81*, 813–846.

Bergman, P., R. Chetty, S. DeLuca, N. Hendren, L. F. Katz, and C. Palmer (2024). Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. *American Economic Review, forthcoming*.

Blei, D., A. Ng, and M. Jordan (2003). Latent dirichlet allocation. *Journal of Machine Learning Research 3*, 993–1022.

Bonhomme, S. (2020). Econometric analysis of bipartite networks. *The Econometric Analysis of Network Data, edited by Bryan Graham and Aureo de Paula, Elsevier Science 1*, 83–121.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*(3), 699–739.

Brock, W. A. and S. N. Durlauf (2001). Discrete choice with social interactions. *The Review of Economic Studies 68*(2), 235–260.

Card, D., J. Rothstein, and M. Yi (2023). Location, location, location. *NBER Working Paper, No. 31587*.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. *Inequality in America: What Role for Human Capital Policies (edited by James Heckman and Alan Krueger), MIT Press 23*(3), 1–10.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Chetty, R. and N. Hendren (2018a). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R. and N. Hendren (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *Quarterly Journal of Economics 133*(3), 1163–1228.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review 106*(4), 855–902.

Chyn, E. (2018). Moved to opportunity: The long-run effects of public housing demolition on children. *American Economic Review 108*(10), 3028–56.

Chyn, E. and D. Daruich (2021). Equilibrium effects of neighborhood interventions. *Working Paper*.

Chyn, E. and L. F. Katz (2021). Neighborhoods matter: Assessing the evidence for place effects. *Journal of Economic Perspectives 35*(4), 197–222.

Cunha, F. and J. J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *The Journal of Human Resources 43*(4), 738–782.

Dempster, A., N. Laird, and D. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society 39*(1), 1–38.

Disch, L. (2024). Effects of family disruption on child development: The moderating role of residential relocation. *Working Paper*.

Doepke, M., G. Sorrenti, and F. Zilibotti (2019). The economics of parenting. *Annual Review of Economics 11*, 55–84.

Durlauf, S. (2004). Neighborhood effects. *Handbook of Regional and Urban Economics 4*, 2173–2242.

Eckert, F. and T. Kleineberg (2024). The geography of opportunity: Education, work, and intergenerational mobility across us counties. *Working Paper*.

Goldsteyn, B., A. Non, and U. Zölitz (2021). The impact of peer personality on academic achievement. *Journal of Political Economy 129*(4), 1052–1099.

Graham, B. S. (2018). Identifying and estimating neighborhood effects. *Journal of Economic Literature 56*(2), 450–500.

Graham, B. S., G. Imbens, and G. Ridder (2014). Complementaritiy and aggregate implications of assortative matching: A nonparametric analysis. *Quantitative Economics 5*, 29–66.

Hanushek, E. (2005). Teachers, schools, and academic performance. *Econometrica 73*(2), 417–458.

Jochmans, K. and M. Weidner (2019). Fixed-effect regressions on network data. *Econometrica 87*(5), 1543–1560.

Kinsler, J. (2016). Teacher complementarities in test score production: Evidence from primary school. *Journal of Labor Economics 34*(1), 1–274.

Kramarz, F., S. Machin, and A. Ouazad (2015). Using compulsory mobility to identify school quality and peer effects. *Oxford Bulletin of Economics and Statistics 77*(4), 566–587.

Laliberté, J.-W. (2021). Long-term contextual effects in education: Schools and neighborhoods. *American Economic Journal: Economic Policy 13*(2), 336–77.

Lentz, R., L. Hong, and J.-M. Robin (2024). Determinants of worker wages and mobility with dynamic worker types. *Working Paper*.

Lentz, R., S. Piyapromdee, and J.-M. Robin (2023). The anatomy of sorting – evidence from danish data. *Econometrica 91*(6), 2409–2455.

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2013). Long-term neighborhood effects on low-income families: Evidence from moving to opportunity. *American Economic Review 103*(3), 226–31.

Mann, L. (2024). Spatial sorting and the rise of geographic inequality. *Working Paper*.

Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies 60*(3), 531–542.

McCann, P. (2023). How have place-based policies evolved to date and what are they for now? *OECD-EC High-Level Workshop Series: Place-Based Policies for the Future*, 1–58.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain (2005). Teachers, schools, and academic achievement. *Econometrica 73*(2), 417–458.

Shan, X. and U. Zölitz (2024). Peers affect personality development. *Review of Economics and Statistics (forthcoming)*.

Todd, P. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal 113*(485), F3–F33.

Winnick, L. (1966). Place prosperity vs people prosperity: Welfare considerations in the geographic redistribution of economic activity. *Essays in Honor of the Sixty-Fifth Birthday of Leo Grebler, Real Estate Research Program, University of California, Los Angeles*, 263–283.

Wodtke, G. T., F. Elwert, and D. J. Harding (2016). Neighborhood effect heterogeneity by family income and development period. *American Journal of Sociology 121*(4), 1168–1222.

Appendices

Appendix A: Value-Added Results

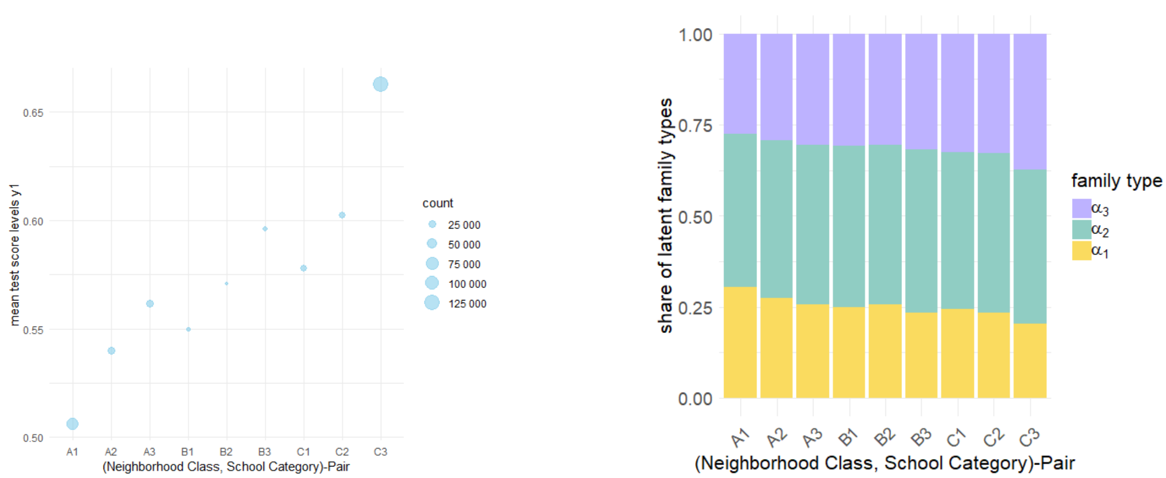
Table 8: Summary statistics of test score gains in period 1 and period 2

Period 1	**mean**	**median**	**sd**	**min**	**max**
All	0.59	0.59	0.06	0.00	1.00
Stayers	0.59	0.59	0.06	0.00	0.96
Transitioners	0.59	0.59	0.06	0.13	1.00

Period 2	**mean**	**median**	**sd**	**min**	**max**
All	0.48	0.48	0.06	0.00	1.00
Stayers	0.48	0.48	0.06	0.00	1.00
Transitioners	0.48	0.48	0.06	0.17	0.82

Appendix B: Additional Estimation Results

Appendix B1: Sorting Based on Test Score Levels

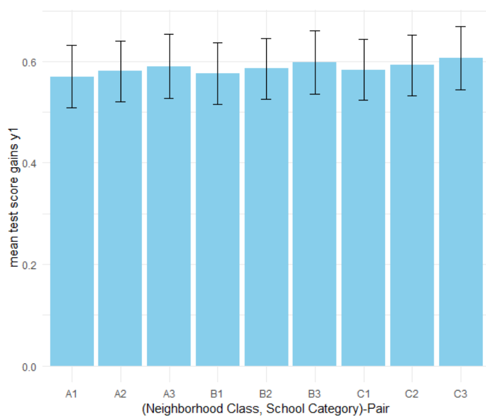


(a) Size of Clusterpairs (Levels)

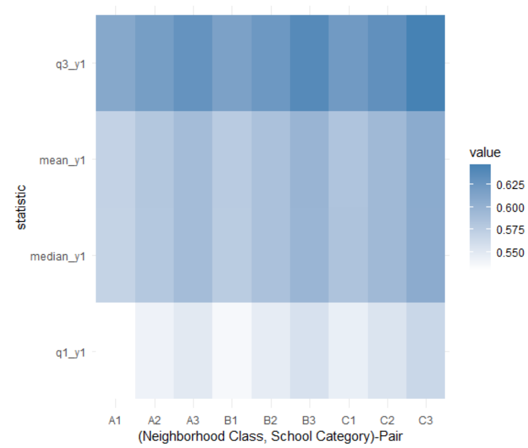
(b) Sorting (Levels)

Figure 8: Estimation based on test score levels

Appendix B2: Additional Cluster-Pair Results



(a) Mean and Standard Deviation

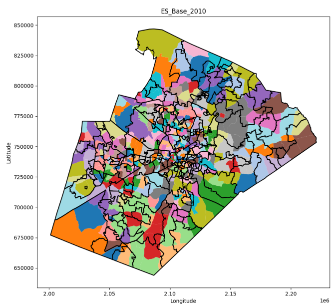


(b) Heatmap

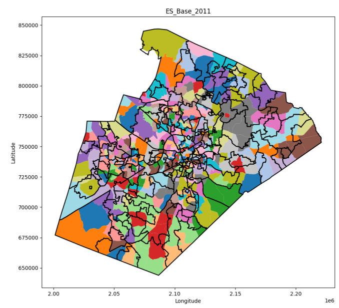
Figure 9: (Neighborhood Class, School Category)-Pairs

Appendix C: Maps Illustrating Changes in School Zones

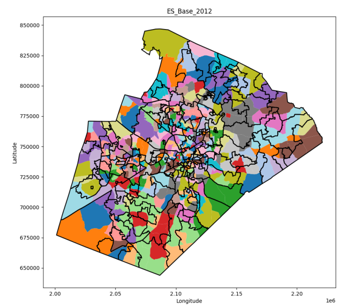
Appendix C1: Rezoning of Elementary Schools



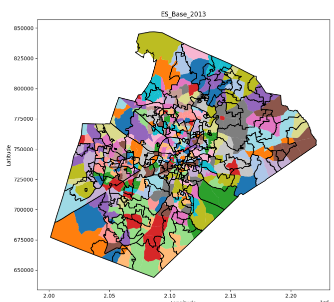
(a) ES Base 2010



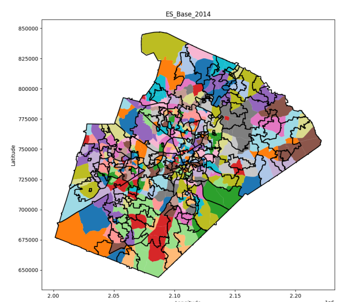
(b) ES Base 2011



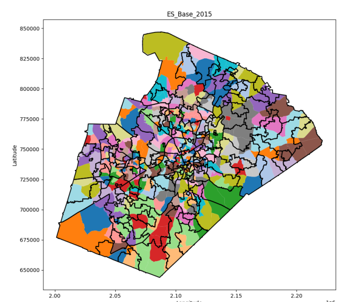
(c) ES Base 2012



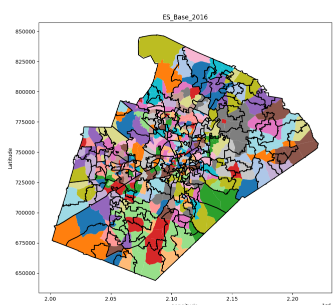
(d) ES Base 2013



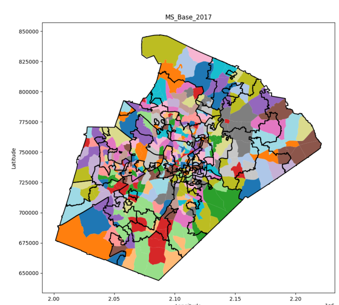
(e) ES Base 2014



(f) ES Base 2015



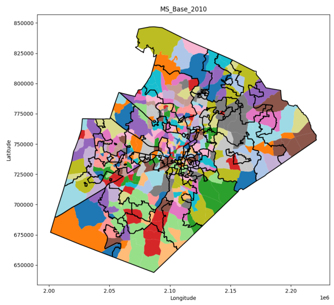
(g) ES Base 2016



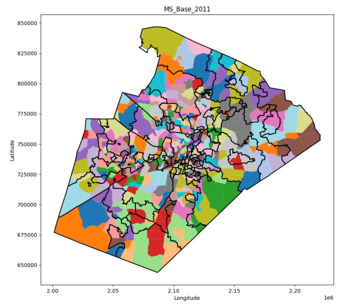
(h) ES Base 2017

Figure 10: Base Schools Maps for Elementary Schools (ES)

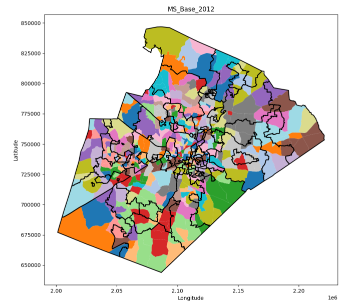
Appendix C2: Rezoning of Middle Schools



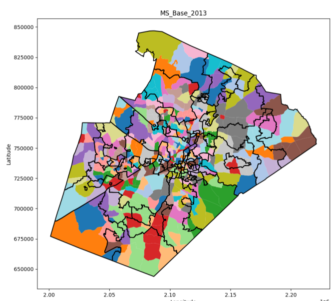
(a) MS Base 2010



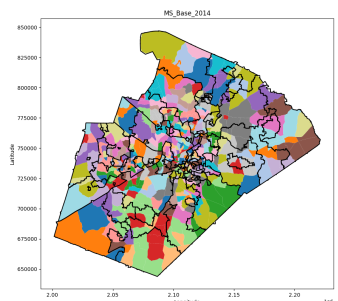
(b) MS Base 2011



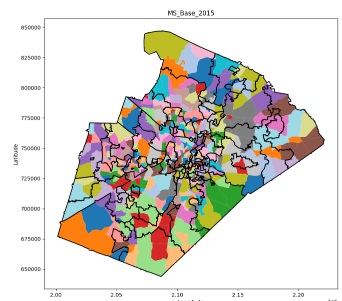
(c) MS Base 2012



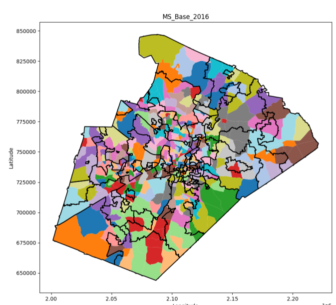
(d) MS Base 2013



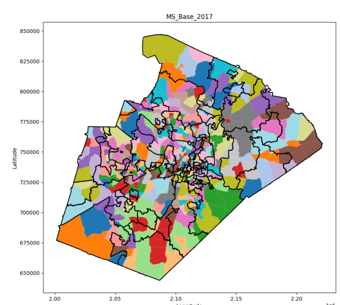
(e) MS Base 2014



(f) MS Base 2015



(g) MS Base 2016



(h) MS Base 2017

Figure 11: Base Schools Maps for Middle Schools (MS)

Appendix D: Additional Counterfactual Results

Neighborhood Quality Improvement. In Table 6, I reported simulation results from a an improvement in school quality. Here, I conduct a counterfactual simulation where, instead of improving school quality, I improve the neighborhood quality, taking school quality as given. In a similar vein, by fixing school categories, I conduct a counterfactual scenario with an improvement in neighborhood quality of those neighborhoods with relatively low test score gains. The increase in neighborhood quality is simulated by imposing the test score gains distribution of neighborhoods of a higher class on neighborhoods of class A, which are those with relatively lowest test score gains. The following Table 9 shows results if, ceteris paribus, neighborhood quality was improved by imposing the test score gains distributions of neighborhoods of class B and neighborhoods of class C, respectively, on neighborhoods of class A. Results are very similar to Table 6, but slightly weaker. The slightly larger impact of replacing lower school categories with higher ones, compared to neighborhood class improvements, may be attributed to the direct effect that school quality has on children's learning growth, whereas neighborhood improvements primarily offer supportive, contextual enhancements. This suggests that improvements in school quality have a stronger impact on learning outcomes.

Table 9: Improvement in Neighborhood quality

Statistic	***a)*** $\boldsymbol{\Delta_{classA \to classB}}$	***b)*** $\boldsymbol{\Delta_{classA \to classC}}$
Mean	0.13%	0.29%
1%-Quantile	0.18%	0.41%
5%-Quantile	0.16%	0.36%
10%-Quantile	0.16%	0.33%
20%-Quantile	0.16%	0.34%
30%-Quantile	0.15%	0.34%
40%-Quantile	0.15%	0.32%
Median	0.13%	0.29%
60%-Quantile	0.11%	0.28%
70%-Quantile	0.13%	0.30%
80%-Quantile	0.12%	0.25%
90%-Quantile	0.08%	0.21%
95%-Quantile	0.05%	0.16%
99%-Quantile	0.05%	0.12%
Variance	−0.000	−0.010

Note: Differences in test score gains between the simulated datasets with upgraded neighborhood quality and the original dataset, holding schools fixed (variances is $\times 100$*). The middle column of the table (a) shows how the entire test score gains distribution would change if there is a moderate improvement of neighborhood quality, so that neighborhoods of class A become neighborhoods of class B,* $\Delta_{classA \to classB}$*. The last column (b) shows these outcomes if neighborhood quality is even higher so that neighborhoods of class A are replaced with neighborhoods of class C,* $\Delta_{classA \to classC}$*.*