The Willingness to Pay for School Quality, Neighborhood Attributes, and Later Life Outcomes*

Sadegh Eshaghnia James J. Heckman Goya Razavi

May 15, 2022

Draft: Please Do Not Circulate

*This research was supported by a Rockwool Foundation grant, an anonymous donor to CEHD, and the Luxembourg National Research Fund (FNR grant # 13572740). Our research partner Rasmus Landersø made numerous helpful comments on a previous draft.

Abstract

Education in Denmark is freely available to all. Yet, despite equalized school expenditures, we observe substantial differences in school quality across areas due to the sorting of individuals across neighborhoods. This paper evaluates the willingness to pay for school quality and neighborhood socio-demographics and proposes a new methodology to do so. We use contiguous housing clusters, comprising of 250 households (nested within 1000-households clusters), which are homogeneous neighborhoods. Using within-cluster variation alleviates the potential problem of sorting that often plagues the willingness to pay in the literature. We estimate a willingness to pay of about 2.6% for houses associated with a school whose quality is one standard deviation above the mean. This estimate survives a variety of sensitivity tests. Using a similar strategy, we estimate the valuation of neighborhood attributes. Finally, using rich longitudinal data, we find that attending better schools significantly impacts later life outcomes, increasing college education and wages while reducing criminality and teenage pregnancy.

JEL Codes: H0, H4, H7, I2, R0, R2, R3

Keywords: Hedonic valuation, amenities, residential sorting, peer effects, educational attainment, labor market outcomes.

S.M. Sadegh Eshaghnia
Department of Economics
and Center for the Economics
of Human Development
University of Chicago
1126 East 59th Street
Chicago, IL 60637
eshaghnia@uchicago.edu

James J. Heckman
Department of Economics
and Center for the Economics
of Human Development
University of Chicago
1126 East 59th Street
Chicago, IL 60637
jjh@uchicago.edu

Goya Razavi
Harris School of Public Policy
University of Chicago
1307 East 60th Street
Chicago, IL 60637
razavi@uchicago.edu

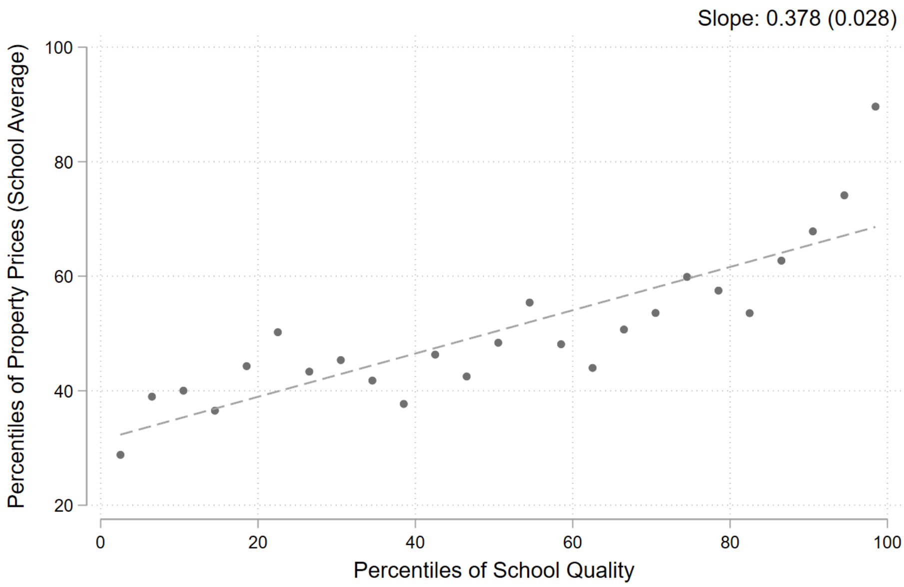
1 Introduction

This paper develops a novel empirical strategy for estimating the marginal willingness to pay (WTP) for schools in Denmark, where school assignment is residence-based and public schools are free. The Scandinavian welfare state is often touted as an exemplary system for reducing inequalities and equalizing opportunities, inter alia, by providing universal high quality education system that is free for all. Yet, there is growing evidence that such equalization, enshrined in the law, is undone in practice, through the sorting of households as well as teachers across neighborhoods (Eshaghnia, 2020; Gensowski *et al.*, 2021; Heckman & Landersø, 2021). Places of residence can affect social mobility (see, for instance, Bénabou, 1993, 1996; Durlauf, 1996). By estimating the WTP for school quality, this paper investigates one possible mechanism behind notable inequities in Denmark (Eshaghnia *et al.*, 2021; Landersø & Heckman, 2017), despite the strong egalitarian welfare system.

Tiebout originated the discussion of sorting and local public good provision. The essence of the problem raised by sorting in a variety of markets is analyzed in Tinbergen (1956) with later applications by Rosen (1974). In the analysis of this paper, households sort across neighborhoods based on heterogeneous preferences for vectors of neighborhood characteristics. This leads to a correlation of neighborhood and individual characteristics which constitute neighborhood attributes. Not accounting for unobserved neighborhood attributes can result in biased estimates of individual valuations of neighborhood amenities in hedonic price regressions.

Figure 1 shows a positive relationship between percentiles of school quality in Denmark (as measured by average test scores) on percentiles of property values of the biological parents, averaged at the school level. Of course, sorting into neighborhoods would lead to correlate school quality and other neighborhood amenities. Our goal in this paper is to disentangle the role played by school quality on house prices, above and beyond other neighborhood attributes.

Figure 1: House Prices and School Quality



Notes: The figure shows a binned scatterplot, with a linear fit, of percentiles of parental property values averaged at the school level on percentiles of school quality. School quality is measured by taking the average test scores of students attending a given school. The sample contains property values for parents who have a child attending 9th grade in a public school between 2002 and 2006.

To deal with the sorting of households across neighborhoods, we analyze sorting on parental traits and preferences in narrowly-defined neighborhoods. We capitalize on variation in house prices and school quality within very small neighborhoods (median size of 0.3 square miles) arising from discontinuous school catchment areas within these clusters. Given the small geographic area spanned by these clusters, we claim to control for unobserved neighborhood amenities by creating homogeneity within clusters. Within such small neighborhoods, access to neighborhood amenities (except for the assigned school) does not vary, allowing us to identify the effect of changes in school quality on house prices by looking at variation within homogeneous neighborhoods.

To motivate our general approach, we begin with a descriptive analysis of the variation within and across different neighborhood levels in Denmark. Denmark has well-functioning private housing markets and labor markets. Our evidence for Denmark likely applies to other Western economies with functioning markets. We present evidence that, by forming smaller geographic units, we get more homogeneous clusters of individuals. Adding geographic fixed effects considerably impacts the coefficients in a standard hedonic price regression. This strategy allows to control for unobserved neighborhood attributes, which are correlated with observable neighborhood attributes through the sorting process. This is illustrated, for example, by the fact that adding these fixed effects flips the sign of some coefficients from standard OLS regressions.

We have access to administrative data covering the full Danish population. This rich data allows us to link every individual to their physical address, at the street level. For each housing unit, we are able to observe a number of attributes including the assigned school district, type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. This allows us to pin down the WTP for school quality by controlling for characteristics that may be correlated with school quality. We also capture rich individual and household characteristics allowing us to control for the key drivers of the sorting process. Finally, school quality measures are constructed based on administrative data, using students grades and teacher employment and academic records, allowing us to isolate the different features of school quality contributing to its capitalization into house prices.

Our strategy requires to satisfy the following identifying assumption. Unobserved housing and neighborhood characteristics are locally independent of each other and with observed characteristics within the geographic clusters we use as fixed effects.

We estimate that households are willing to pay around 2.6% (ranging from 2% to 3.5%) of house prices for a one standard deviation increase in school quality, measured by average school grades. This is broadly in line with estimates found in other countries with greater inequality in income and wealth than Denmark, including Australia, France, the UK and the US (see Black & Machin, 2011 for a review). Households positively value neighborhood attributes such as being in a neighborhood with a larger fraction of more educated individuals and with higher average income.

Our results on the willingness to pay for school quality are robust to various specifications. In particular, we further assess the potential sorting bias that may plague our estimates of mean preference due to individuals sorting differentially across neighborhood driven by heterogeneous preferences. We address this issue by using an approximation to a formal polychotomous neighborhood choice model, developed by Dahl (2002).

Building on this framework, we further study whether the households' WTP for school quality translates into better later life outcomes for the children who attend these schools. Early studies of school quality, beginning with Coleman (1968), find no association between school inputs and student achievement on standardized tests (see Hanushek *et al.* (1996), Betts (1995), Betts (1996), and Heyns (1997) for a few surveys of early studies). Some papers such as Johnson & Stafford (1973) and Card & Krueger (1992) use variation across states or across age cohorts within states to identify the effects of school input measures on the rate of return on schooling and find economically significant effects. Heckman *et al.* (1995) show that misspecifications in Johnson & Stafford (1973) and Card & Krueger (1992) are partially responsible for their results and call into question the strength of the evidence for associations between school quality and earnings based on aggregate data. Some more recent studies use microdata and analyze the impact of various school inputs such as teachers' characteristics and salaries, class size, pupil-teacher ratios, expenditures per pupil, school resources, and the composition of peers on student performance (Altonji & Dunn (1996); Dearden *et al.* (2002a); Rockoff (2004); Rivkin *et al.* (2005); Cullen *et al.* (2005); Hastings & Weinstein (2007); Clark (2010); Chetty *et al.* (2011); Jackson (2013); Deming *et al.* (2014); Bernal *et al.* (2016)).[1] These papers report mixed results regarding the relevance of school inputs for students' performance.

To estimate the impact of school quality on later life outcomes, we estimate an empirical educational production function. Our first strategy relies on variation in school quality within clusters, holding constant a large vector of individual, household and neighborhood

[1] Glewwe *et al.* (2011) review studies about the impact of school expenditures and teacher characteristics on students' years of completed schooling

characteristics. Our second strategy capitalizes on within family variation in school quality, arising from changes in peers and teacher quality across cohorts. Third, we build on within school and across cohort variation to further pin down the effect of better schools on later life outcomes.

Our paper is distinct from the previous studies in the literature in several important ways. First, we link school quality to various long-run outcomes of children, a unique feature lacking from previous studies in the literature. These outcomes include various income definitions, college completion, criminal behavior and teenage birth, amongst others. Second, we take into account neighborhood quality and a rich vector of parental characteristics, which helps us better capture the impact of various factors correlated with school quality. Third, unlike other papers in the literature that focus on a particular school district or a local policy change, we use administrative data on the whole population of an entire country in which most children attend public schools freely available to all.

Our findings are consistent with the results of Chetty *et al.* (2014) who use a value-added approach and find that students assigned to high-value-added teachers are more likely to attend college and earn higher salaries as adults. This paper, however, considers a broader array of inputs and examines both teachers' characteristics and the composition of peers at schools and neighborhoods of residence.

Looking at both the WTP for neighborhood attributes and their effect on later life outcomes, this paper speaks to the important role of sorting across neighborhoods in Denmark. We provide evidence that this purposeful neighborhood sorting leads to important differential outcomes later in life. In fact, the combination of sorting together with the residence-based assignment rules of households to schools is associated with later life outcomes, increasing income by around 4%, for a one standard deviation increase in school quality attended. We also show that students attending schools one standard deviation above the mean are, on average, 2.1% more likely to attend college, conditional on neighborhood quality and parental characteristics. Similarly, we show that better schools are associated with reduced criminality and teenage pregnancy. This provides evidence that household valuation of school quality, as seen through its capitalization in house prices, is potentially driven by parents realizing its important effect on later life outcomes as they value it.

By combining our estimates for the WTP for school quality with the results of our later life outcome analysis, we estimate the internal rate of return of investing in children through living in a more expensive neighborhood with higher school quality. We find internal rates of return ranging from 3.7% to 8%, which is in range of the internal rates of return estimated by García *et al.* (2021) for the early childhood education program of Perry Preschool Project in the US.

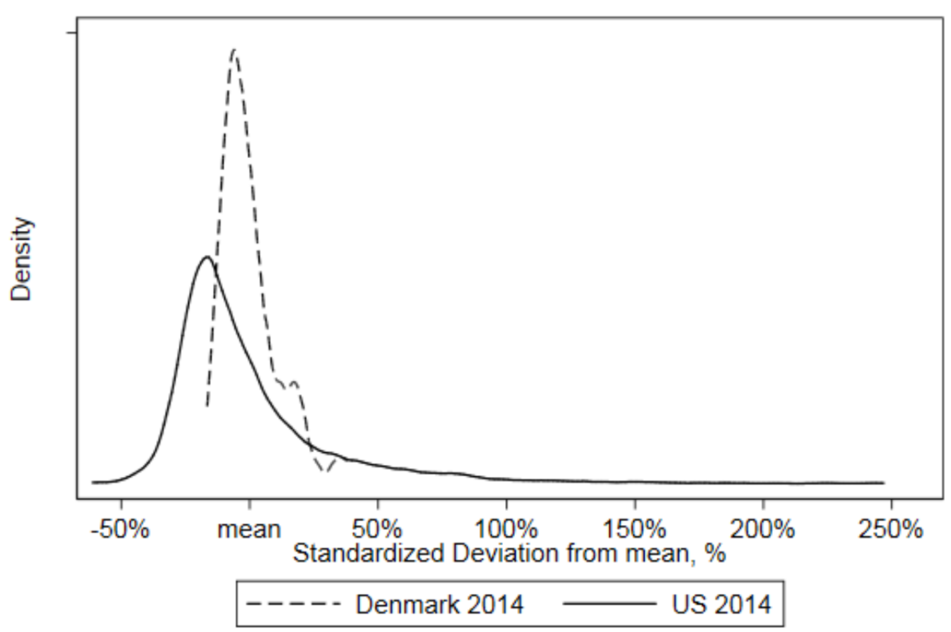
2 Public Schooling System in Denmark

The Danish schooling system is based upon the principle of schooling for all, which is provided at no charge in public schools. Danish municipalities' primary revenue in Denmark (about 70%) comes from local taxes, which vary only minimally across municipalities (OECD, 2016). A system of redistribution across municipalities also exists to correct for differences in tax revenues. Importantly for our purposes, this constrains the amount of variation in per pupil expenditure as can be seen in Figure Gensowski *et al.* (2021)).

Despite an equalized school expenditure and teacher salary distribution, teachers still sort based on the potential quality of the students they may be teaching, as exemplified by Gensowski *et al.* (2021). In terms of our strategy, this means that differences in school quality would not be driven by differences in school expenditure but from sorting of teachers, as well as students. Using a measure of teacher quality in Denmark, we are able to investigate household valuation in these different aspects of school quality, including teacher quality and peers, bearing in mind that any municipality differences in funding or tax rate would be subsumed in our fixed effects.

Access to public schools in Denmark is residence-based. Each housing unit is part of a school district that is assigned to a single school. Parents can defy these school district rules in certain cases, although it is contingent on available capacity in alternative schools. We test the sensitivity of our estimates to such non-compliance in Section 7.2 and find that our coefficient is robust.

Figure 2: Comparing school expenditure between Denmark and US



Notes: The figure shows average per pupil school expenditures in public schools in 2014 relative the to the country average. Source: Gensowski *et al.* (2021)

3 Methodology

This paper develops a framework to recover household preferences for school quality. Our empirical strategy builds on Tinbergen (1956) and Rosen (1974). Their analyses rationalize the interaction between consumers and suppliers in competitive markets with differentiated goods. An equilibrium is reached when supply equals demand at each traded point of quality. The hedonic price function $P(z)$ is defined for $z = (z_1, z_2, \ldots, z_n)$, vector of attributes of the good. In our context, z is comprised of neighborhood public services such as local school quality. The gradient of the hedonic price function with respect to school quality gives the equilibrium differential that allocates individuals across locations. Locations with poor neighborhood public services, such as low school quality, must have (ceteris paribus) lower housing prices, to attract potential buyers. In this framework, at each point on the hedonic price function, the marginal prices of housing characteristics are individual consumer's marginal willingness to pay for that characteristic and will be equal to the individual supplier's marginal cost of producing it for those who supply and purchase it.

A long line of research using hedonic demand models building on (Rosen, 1974) include Epple (1987), Ekeland *et al.* (2004), Bajari & Benkard (2005), and Heckman *et al.* (2010). A key issue that the literature has aimed to address is the matching of neighborhood characteristics which arises from household sorting. In this section, we describe our regression framework, to recover estimates of the marginal WTP for school quality, measured by test scores, in the presence of household sorting.

3.1 Hedonic Framework

Our main estimating equation relates house prices to a vector of housing and neighborhood characteristics, including school quality. We add a set of cluster fixed effects to control for unobserved neighborhood heterogeneity and estimate the following hedonic price regression:[2]

$$\ln(p_{imk}) = \alpha + \beta S_{mk} + \gamma X_{imk} + \rho_{kt} + \varepsilon_{imk}, \tag{1}$$

where $\ln(p_{imk})$ denotes log property values of individual i who attends school m in cluster k. S_{mk} denotes our measure of school quality for school m in cluster k. We further add a set of housing and neighborhood characteristics denoted by X_{imk}, as well as neighborhood-by-cohort fixed effects, ρ_{kt}. Finally ε_{imk} represents unobserved neighborhood and housing attributes that are assumed to be iid.

[2] We also run a local linear version of this specification in Appendix A.8, which shows that the log-linear specification captures the underlying nonlinear nature of the hedonic price function well.

3.2 Proximity Theorem

Our methodology for recovering the WTP for school quality relies on the Proximity Theorem.

Theorem 1 (Proximity Theorem). *$\hat{\beta} \to \beta$ as $\mathrm{var}(\varepsilon) \to 0$ or $\mathrm{cov}(S, \varepsilon \mid X, \rho) \to 0$, if S has full rank.*

In other words, the least square estimator is consistent if (1) the variance of the disturbance approaches zero or (2) as the probability limit of the correlations between the disturbance and regressors approaches zero, if the matrix of regressors has full rank (Fisher, 1966).

Formally, consider $Y = \beta S + \varepsilon$, where Y is the price of housing and S a is school quality. We acknowledge that the regressors of the hedonic equation might be correlated with the disturbance term, specifically:

$$\mathrm{Cov}[S, \varepsilon \mid X, \rho] \neq 0 \tag{2}$$

Still, thanks to our cluster fixed effect strategy, which captures the neighborhood sorting at a very local level, we assert that variance of the disturbance term in the hedonic equation is small.

The least square estimator for β is as follows: $\hat{\beta} = \beta + \frac{\mathrm{cov}(S,\varepsilon)}{\sigma_S^2}$. While S may be correlated with ε, σ_ε^2 is small by design (where σ^2 denotes the variance). As the size of the neighborhood, N_X shrinks, $N_X < \varepsilon_T$, we have that $\hat{\beta} = \beta + \frac{\mathrm{cov}(S,\varepsilon)}{\sigma_S^2}$, where $\frac{\mathrm{cov}(S,\varepsilon)}{\sigma_S^2} \to 0$.

The Cauchy-Schwarz inequality ($\mathrm{cov}^2(S, \varepsilon) \leq \sigma_S^2 \sigma_\varepsilon^2$) implies that the difference between $\hat{\beta}$ and β, in a probability limit sense, is not greater than $\frac{\sigma_\varepsilon}{\sigma_S}$.[3] Since our empirical strategy sets σ_ε close to zero, our estimate of the willingness parameter is nearly consistent.[4]

This argument suggests a test: as the geographic unit becomes smaller, the R-squared of the hedonic regression should increase. In Section 5.3, we present the results of such a test,

[3] From the Cauchy-Schwarz inequality, we have $\mathrm{cov}(S, \varepsilon) \leq \sqrt{\sigma_S^2 \sigma_\varepsilon^2}$. Thus $\sigma_S^2(\hat{\beta} - \beta) \leq \sqrt{\sigma_S^2 \sigma_\varepsilon^2}$. The result follows.

[4] The full rank assumption for the matrix Z, guarantees that all the components of the vector β are identified.

which confirms our hypothesis regarding the relationship between the size of the geographic unit and the variance of the disturbance.

3.3 Sorting

Households might still sort on either side of school district boundaries to gain access to preferred schools, as we document below (and as reported in the literature, see for instance, Bayer *et al.*, 2007). This issue could bias our estimates so long as parents value neighborhood characteristics more locally than the 0.3 square mile neighborhoods we consider. Our rich data enables us to control for hyper-local neighborhood (0.1 square mile) sociodemographics, which not only allows us to recover household valuation of school quality, but also other hyper-local neighborhood attributes. The fact that our estimates further decrease when adding these hyper-local controls showcases the importance of sorting across neighborhoods in Denmark, in particular to locate in the catchment area of better schools.

Sorting of households is an important concern in hedonic models, in particular for recovering mean preferences of heterogeneous agents. For instance, if individuals who value highly school quality live in areas with better schools, the marginal WTP may reflect the preferences of this sub-population. With heterogeneous tastes, the marginal WTP recovered may not align with the average marginal WTP. In Appendix A.3, we present a framework developed in Bayer *et al.* (2007) that maps estimates of hedonic model to mean preference estimates from a sorting model, under preference heterogeneity. This analysis shows that for attributes which vary continuously throughout the country (supplied at various levels across neighborhoods), hedonic models do recover mean preferences.

In Section 8, we further provide a test for whether our estimates are biased due to households self-selecting into neighborhoods, on the basis of taste dispersion, using a methodology taken from Dahl (2002) (which approximates the Heckman (1979) correction model in the multidimensional choice context).

3.4 Boundary Discontinuity Design

Our method is distinct from Black (1999) who used boundary discontinuity (BDD) to address this endogeneity issue. Black's approach uses boundary fixed effects in order to compare houses that are near but on opposite sides of school catchment areas' borders. The identifying assumption for this approach is that unobserved amenities vary continuously at the border while school characteristics are determined by attendance zones, and are discontinuous at boundaries. Estimates using this approach are typically five times lower than cross-sectional estimates (see Bayer *et al.*, 2007; Gibbons *et al.*, 2013). We find a similar discrepancy when running a standard hedonic price regression compared to our cluster fixed effects approach.

BDD approaches have been used exploiting changes in boundaries (see for instance Bogart & Cromwell, 2000 and Ries & Somerville, 2010). These studies both lack information on neighborhood quality and school composition. Moreover, variation can be coming from houses that are geographically distant from eachother, albeit being close to the boundary (and thus lie in different types of neighborhoods). A further important drawback to methods that use temporal shocks to school quality to derive WTP estimates is exposited by Kuminoff & Pope (2014). Their work shows that these studies need to assume (and do so without providing evidence) that the price function is constant over several years (sometimes decades). Our strategy does not require such an assumption and we show that the hedonic price schedule is not time-constant. Finally, BDD designs rely on variation at the boundary, which without further evidence, may not be representative of the broader population.

In Section 7.2, we present a strategy akin to a BDD. We show that our results are robust.

4 Data

This paper relies on administrative data from Statistics Denmark, which provides data for the whole population. We focus on five cohorts of 9th graders, who attended 9th grade between

2002 and 2006 and whose biological parents are homeowners. Below we describe in more details the different key variables used in the analysis.

4.1 School Quality Measures

A key aspect to any study assessing the WTP for school quality is to define meaningful measures of school quality. Different types of measures of school quality have been used in the literature, including output-based, input-based, and value-added measures.

Value-added measures require tracking of students' performance over time and are thus difficult to construct. Moreover, Brasington (1999), Downes & Zabel (2002) and Brasington & Haurin (2006) find little support for such measures being capitalized into house prices. Input-based measures, such as per-pupil spending agree (Hanushek, 1986, 1997). This has led to the more prevalent use of output-based measures, which are our main measures of school quality.

We construct our output-based school quality measure by averaging over students' grades in exams that are taken in their last year of compulsory education. These students take national exams in a wide range of subjects and complete them, for the majority, at age 16. Average grades at the school level, broken down by subjects, are available publicly to parents. This makes it a potentially important signal of school quality.

We also use a measure of teacher quality based on Gensowski *et al.* (2021). Using administrative records, all employees in teaching positions in schools between 2009 and 2016 are matched to (1) their academic records from high school (grades in Danish and Mathematics exams) and university as well as (2) employment records to identify unemployment spells. Children's GPA are then regressed on these teacher's characteristics. A national rank of school quality is then generated using linear regression.

4.2 Property Prices

Our outcome variable is defined as the value of the property owned by the biological parents. We use governmental valuations of property prices.[5] It is measured at the start of the school year in which the final exam is passed, taken on average at age 16. We take the natural logarithm of house prices in the empirical analysis. Moreover, we drop all outlying housing values below the first or above the 99th percentile (keeping observation with house prices above \$44,000 and below \$2.5 million (2010 USD)).[6] The average number of years spent in the house is, in our sample, 11 years. About 40% of our sample never moves after the child is born. About 30% move only once and would end up living on average 8 years in the house. Thus, most people spend their whole school-age time (10 years) in the same house.

4.3 Neighborhood Concepts

Throughout the paper, we focus on various neighborhood concepts. Table 1 provides a short description of these neighborhoods as well as their numbers in our estimating sample.

In the remainder of this paper, we refer to small clusters (median size of 0.1 square miles) as small neighborhoods or hyper-local neighborhoods, interchangeably. Large clusters (median size of 0.3 square miles) are also referred to as neighborhoods or clusters. Hence our fixed effects are at the "neighborhood" or "cluster" level.

In Appendix A.4, we present a schematic graphical depiction of the relationship between school districts and large clusters and clarify the variation we use visually.

4.3.1 Clusters as Neighborhoods

In constructing the clusters, we build on the methodology implemented by Damm & Schultz-Nielsen (2008), from 1985 to 2004, which satisfies the following relevant criteria: (1) Clusters

[5]Government valuation is computed based on sales of other housing units in the relevant market and adjusted for specific characteristics of the property (such as its square footage).

[6]An exchange rate of 6.7 DKK per US dollar is used to obtain the dollar values.

Table 1: NEIGHBORHOOD CONCEPTS

Neighborhood Concept	#	Description
Region	5	Below federal level. Supersedes municipalities.
Municipality	269	Each municipality lies within a specific region.
School District	1011	Each school district contains one single school in Denmark.
Parish	1309	A neighborhood formed around a specific church.
Large Cluster	1708	See below for details on their construction. Also referred to as "Cluster" or "Neighborhood"
Small Cluster	6204	See below for details on their construction. Also referred to as "Hyper-local" or "Small Neighborhood"

Note: Number and definition of different neighborhood concepts in Denmark.

correspond to geographical areas within which an individual has social contact[7]; and (2) should be unaltered over time.[8]

Following these rules Damm & Schultz-Nielsen (2008) construct clusters on the basis of 431,233 hectare cells (100m x 100m) which exhaust Denmark's surface. They then aggregate these cells until the confidentiality requirements are met in terms of the number of households per cluster.[9] The clustering is defined based on housing type and ownership information.[10]

Moreover, visible features and geographical barriers such as lakes, forests or major roads were used in guiding the different boundaries between clusters (which was not always possible in less dense areas). This is an important feature for our research design since it ensures

[7]In practice, this implies that two neighbors separated by physical barriers such as water, large roads or forests, would not be included in the same cluster.

[8]These clusters define geographic areas which do not vary over time. Their composition varies as individuals move in and out of the cluster.

[9]Cluster sizes need to have at least 150 households for analyses of residential segregation and a minimum of 600 households for descriptive purposes, as required by Statistics Denmark.

[10]Housing type in the register data is divided into four categories: farmhouse or detached house; townhouse or small block of flats; large block of flats; second home or other house. Ownership information is also broken down into four categories, namely private ownership, privately owned rental, publicly owned rental and private cooperative housing. In the calculation of which hectare cell is most similar, the latter is given a weight of 70%, while the former 30% in forming homogeneous clusters.

that within cluster differences in house prices are not driven by these barriers. [11]

Table 2: Summary Statistics of Clusters for 2004

	Median	Mean	Std.Dev.
Small Neighborhoods			
Households 2004	245	272.7	115.0
Persons 2004	526	592.2	285.5
Size (in hectares)	22	47.5	64.46
Large Neighborhoods			
Households 2004	985	1079.7	396.2
Persons 2004	2090	2344.5	1039.0
Size (in hectares)	88	187.9	236.3
# of small neighborhoods	4	4.0	1.3

Note: Summary statistics of the composition of clusters, looking at both small and large neighborhoods. Source: Damm & Schultz-Nielsen (2008)

4.4 Control variables

To complement our data on school quality and property values, we use Denmark's rich administrative data to control for a wide range of characteristics at a hyper-local neighborhood level. Given households' propensity to sort across neighborhoods, even within the clusters we consider, these variables allow us to reduce any potential bias arising from sorting. More specifically, we use variables pertaining to the household, including income[12], education level[13], crime[14] as well as information on family structure.[15] We aggregate these measures both at the small cluster level (used in Section 6.2 where we provide estimates of the capitalization of neighborhoods into house prices) and at the school level (used in Section 6.3 where we provide estimates of the capitalization of school-level attributes into house

[11] Further details on the construction of these clusters and expansion to different years are discussed in Appendix A.5.

[12] We use gross household income excluding transfers.

[13] We use years of completed education. When computing hyper-local neighborhood education level, we first compute the maximum number of years of education at the household level. We then aggregate at the small cluster level.

[14] We use an indicator for whether an individual as committed a crime or not in a given year.

[15] We include a measure of marital status and intact family structure. For the latter, a family structure is considered intact if during the first 18 years of a child being born, both parents are present.

Table 3: Summary Statistics

		Mean	St. Dev.
Neighborhood Level			
	HH Gross Income (Excl. Tr.)	52,745	20,200
	HH Max. Years of Schooling	12.5	.886
	Married Household (%)	.769	.209
	Not Intact Household (%)	.506	.232
	Foreigner (%)	.050	.083
	Non-Western Foreigner (%)	.028	.062
	Crime (%)	.024	.010
School Level			
	HH Gross Income (Excl. Tr.)	95,721	23,423
	HH Max. Years of Schooling	13.0	.827
	Married Household (%)	.768	.097
	Not Intact Household (%)	.514	.109
	Foreigner (%)	.068	.066
	Non-Western Foreigner (%)	.030	.049
Housing Attributes			
	Age of Building (years)	51.0	37.5
	Living Area (sqm)	149.8	47.7
	Number of Floors	1.14	.56
	Number of Apartments	4.19	14.2
	Number of Rooms	5.08	1.45
	Number of Toilets	1.60	.59
	Number of Bathrooms	1.31	.49
	House Price (2010 USD)	255,253	218,115

Note: This table reports summary statistics on the key variables used in the analyses. Our sample comprises of all Danes who complete 9th grade in years 2002–2006. We focus on homeowners, since we can observe their housing prices. Incomes are converted from DKK to 2010 USD. An exchange rate of 6.7 DKK per US dollar is used to obtain the dollar values.

prices).[16] Moreover, we include a host of housing characteristics, including the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property.

This paper relies on administrative data available for the whole population of Denmark as provided by Statistics Denmark. We present summary statistics in Table 3.

[16] We report the correlation between the different neighborhood attributes in Appendix A.2.

5 Neighborhood Composition

5.1 Neighborhood Homogeneity

Before moving to our hedonic framework, this section aims to give a better grasp of our ability to control for housing and neighborhood heterogeneity when controlling for cluster fixed effects.

To analyze the spatial decomposition of inequality in housing types and characteristics across neighborhoods in Denmark, we use the Theil's T Index. Figure A.4 of Appendix A.6 shows the decomposition across neighborhoods by different units of neighborhood, i.e., municipality, parish, large cluster, and small cluster levels. Panel (a) focuses on the number of apartments in each property. The results of Panel (a) suggest that while at municipality-level only about 35% of the inequality can be contributed to between-neighborhood component, the share of between-neighborhood component increases to about 90% when we analyze the inequality across small clusters.[17] Panel (b) considers the number of floors of each property. The results suggest that the share of within neighborhood inequality decreases from about 55% to less than 15% when focusing on the cluster level rather than the municipality level. These results reassure that the housing types in our narrowly-defined neighborhood units do not vary and the variation in house prices is not driven by differences in the housing structure.

Panel (c) and (d) of Figure A.4 present the results for other housing characteristics, i.e., age of the building and the living area. Similar to the results in Panels (a) and (b), the share of between neighborhood inequality increases by a factor of 3 to 4 when focusing on the cluster level rather than the municipality level. Our rich set of housing characteristics in our hedonic regression control for such differences that may affect the house price.[18,19]

[17] Appendix A.6 discusses how we use the Theil's T Index to compute within- and between- neighborhood inequality.

[18] We also analyze the spatial decomposition of income inequality across neighborhoods in Denmark using the Theil's T Index in Appendix A.6.

[19] We also analyze the segregation intensity over the income distribution by neighborhood unit in Appendix A.6

5.2 Neighborhood Sorting

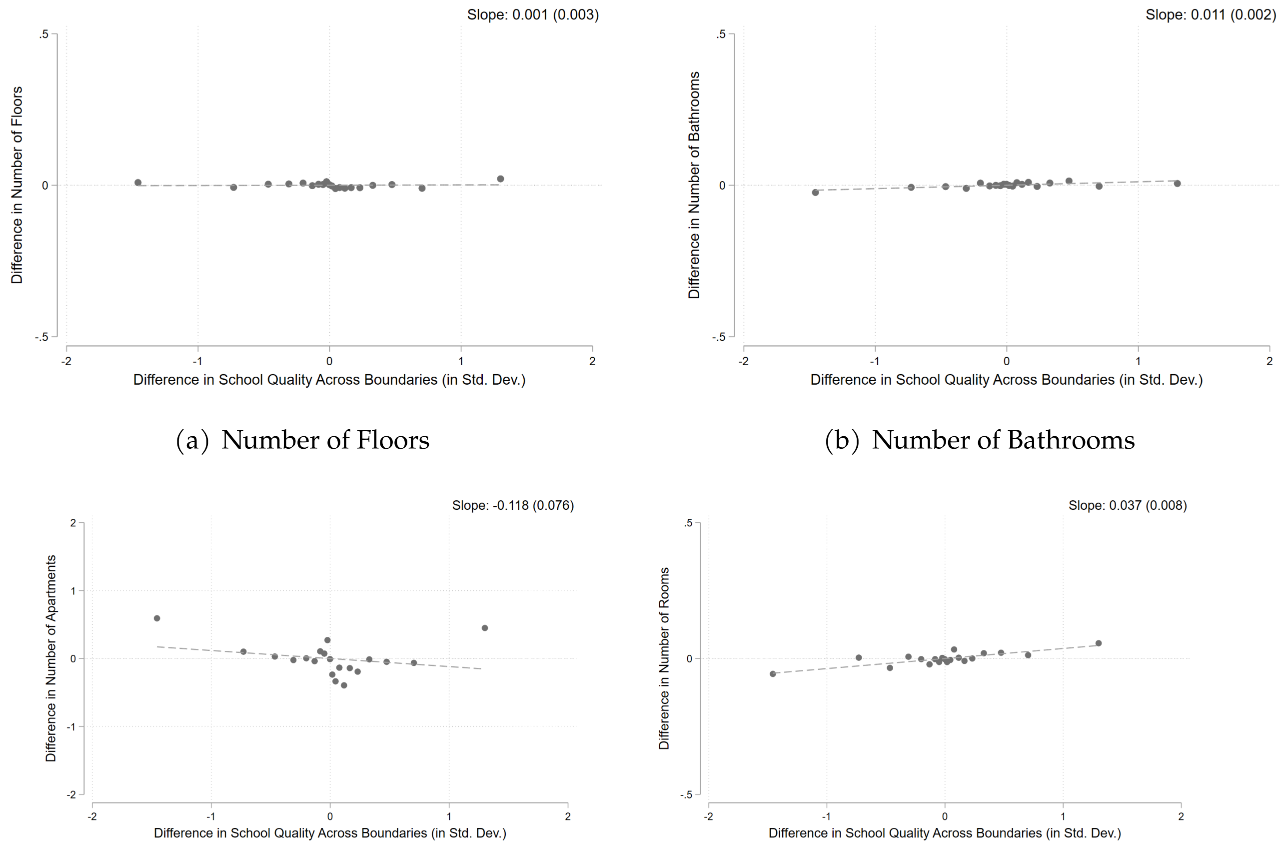
The previous section provides evidence on the homogeneity of our clusters with respect to housing structure, relative to broader neighborhood concepts, such as municipalities. In this section, we take a more direct approach and present evidence explaining the variation that still remains. In line with our identifying assumptions, we show that housing characteristics are rather homogeneous within large clusters. This should not come as a surprise, given that the algorithm generating these clusters aimed to minimize the within variance in housing types (as presented in Section 4.3.1). Still, we show below the strong sorting of individuals based on their own characteristics as well as characteristics of the neighborhood (in line with our analysis in Appendix A.6, which shows that although clusters are more homogeneous, there is a still a significant within cluster variation). Thus, the remaining heterogeneity may be driven by sorting across school boundaries, which occurs even within the large clusters we consider as fixed effects. This emphasises the need to control for these hyper-local neighborhood characteristics in hedonic price regressions.

To assess the level of sorting within large clusters in Denmark, we plot the relationship between school quality and different attributes, both at individual and household level after controlling for neighborhood-by-cohort fixed effects.

This analysis provides a test of our identifying assumption—unobserved neighborhood attributes should not vary within large clusters. First, we show in Figure A.8 that housing prices respond to better relative school quality. In Figures 3 and A.9, we then show that conditional on small neighborhood characteristics, and neighborhood-by-cohort fixed effects, housing characteristics are largely uncorrelated with differences in school quality across schools. This provides evidence that the neighborhoods we consider as fixed effects are rather homogeneous, at least with regards to the make up of their housing typology.

Still, household evidently sort across neighborhoods in Denmark, based on their heterogeneous preferences for a vector of neighborhood attributes. This is evidenced by the relationship between individual characteristics, as well as hyper-level characteristics with school

Figure 3: Sorting Within Large Clusters - Housing Characteristics



(a) Number of Floors

(b) Number of Bathrooms

(c) Number of Apartments

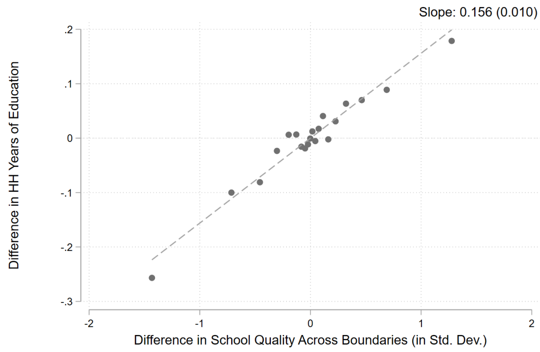
(d) Number of Rooms

Note: Relationship between the difference in housing characteristics and the difference in school quality within large clusters. Each panel is constructed by regressing various housing characteristics on school quality, controlling for neighborhood-by-cohort fixed effects and hyper-local neighborhood attributes—average income, years of education, fraction married, non-westerners, foreigners, non-intact households, private schools and average neighborhood school quality. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

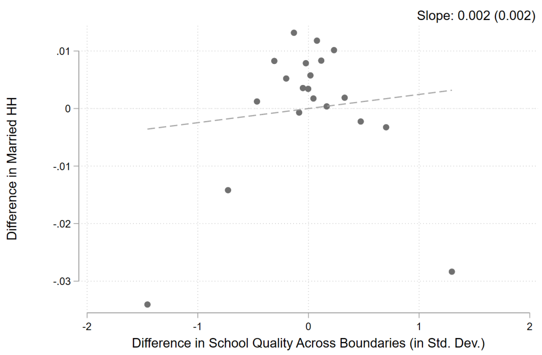
quality, within large clusters. On average, households on the high test score side have more gross income, education and stable family structures, as seen in Figures 4 and A.10. We also find that the high side neighborhood is on average richer, in terms of income, more educated, has less criminality, with more stable family structure and a smaller fraction of western or non-western foreigners, as depicted in Figures A.11 and A.12. Overall, this evidence showcases the importance of controlling for unobserved neighborhood characteristics through our cluster fixed effects strategy as well as hyper-local neighborhood attributes.

Figure 4: Sorting Within Large Clusters - Individual Characteristics

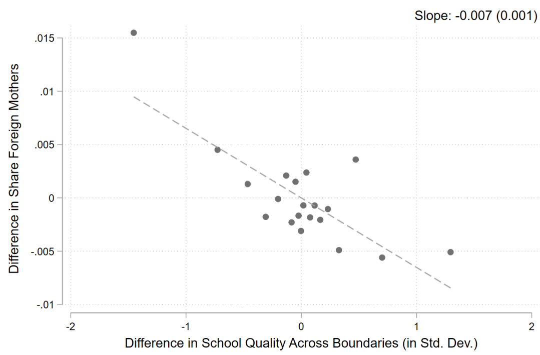
(a) Gross Income Excluding Transfers



(b) Years of Education



(c) Married Household



(d) Foreign Mother

Note: Relationship between the difference in individual characteristics and the difference in school quality within large clusters. Each panel is constructed by regressing various individual characteristics on school quality, controlling for neighborhood-by-cohort, hyper-local neighborhood as well as housing attributes—the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

5.3 Fixed Effects and Unobserved Preferences for Local Amenities

Section 5.1 provided evidence suggesting that the sorting pattern is stronger at smaller geographic units. Section 5.2 documented that families sort into schools given their local neighborhood units, while their access to other amenities does not vary within the small geographic units by design. Our hedonic approach exploits a fixed effect model to account for unobserved preferences for neighborhood amenities and public goods. In this section, we show that the unexplained variation in house price decreases with the granularity of the geographic units. This gives credibility to our estimation strategy for identifying the WTP parameter, which relies on the Proximity Theorem discussed in Section 3.

To do so, we analyze the adjusted R-squared of a set of regressions of house prices on school quality, with fixed effects at different neighborhood-by-cohort levels. To define the neighborhood unit, we use five alternatives, namely regions, municipality, parish, large cluster, and small cluster (by diminishing order of size).

Table 4: R-squared for a Set of Neighborhood Fixed Effect Models

	Region FE	Munic. FE	Parish FE	Large Cl. FE	Small Cl. FE
	(1)	(2)	(3)	(4)	(5)
Adj. R^2 (full sample)	.18	.26	.29	.30	.34
Adj. R^2 (Cph. Area)	.19	.27	.35	.46	.56
# FEs (full sample)	25	1,341	7,699	10,158	32,538

Note: Column (1) presents the adjusted R-squared of the regression of house prices on school quality using a region-by-cohort fixed effect model. Column (2) presents the adjusted R-squared of the regression of house prices on school quality using a municipality-by-cohort fixed effect model. Column (3) reports the adjusted R-squared of the regression of house prices on school quality using a parish-by-cohort fixed effect model. Column (4) shows the adjusted R-squared of the regression of house prices on school quality using a large cluster-by-cohort fixed effect model. Column (5) presents the adjusted R-squared of the regression of house prices on school quality using a small cluster-by-cohort fixed effect model. We do not add any controls. For each of these specifications we provide the corresponding number of area-by-cohort fixed effects for the full sample. We also provide a breakdown of the adjusted R-squared for the full sample as well as for a subset of our data focusing on the Copenhagen Metropolitan Area.

Table 4 shows that, consistent with our argument based on the Proximity Theorem, the unexplained variation in house prices decreases when we move towards more narrowly-defined geographic units. For example, the adjusted R-squared increases by about 67% when we shift from the region fixed effects to large cluster fixed effects model. As Table 4 shows, the adjusted R-squared of the regression of house prices on school quality using the micro-level data increases from 0.18 for regions fixed-effect to .26 for municipalities fixed effects, to 0.29 for parishes, to 0.30 for large clusters, and to 0.34 for small clusters fixed-effects.

We further show that our strategy would work best in more densely populated areas where the implied sizes of the cluster would be smaller and where adding cluster-level fixed effects would capture more of the unobserved heterogeneity. In particular, we see that including large cluster fixed effects captures 46% of the variation in house prices, while going up to 56% for the small cluster fixed effects specification. Thus, although our assumption is likely to hold better in urban areas, such as the Copenhagen Metropolitan Area, we show in Section 7.2 that our coefficients remain largely the same.

6 Hedonic Price Regressions

6.1 Baseline and FE Estimates of School Quality Valuation

We present a set of baseline estimates in the first two columns of Table 5. First, we run a simple OLS regression of house prices on school quality, without controlling for observed nor unobserved neighborhood characteristics. These estimates would imply that a one standard deviation increase in school test scores would increase house prices by 14.4%. In our second OLS specification, we also add a vector X of hyper-local neighborhood characteristics, such as average income and education (presented in more details in Section 6.2), as well as housing characteristics. The set of housing characteristics includes the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. The coefficient on test scores now decreases notably, from .14 to .02, emphasizing the role of of these neighborhoods and housing characteristics in the sorting process.

We can now turn to our cluster fixed effects strategy, to assess the role of unobservables, in the third and fourth columns of Table 5. Compared to the OLS regression with no covariates, we see that the cluster fixed effects specification significantly reduces the estimated capitalization of school quality in house prices, showcasing the importance of controlling for unobserved neighborhood characteristics.

The final column adds both the neighborhood fixed effects, ρ, and the hyper-local and housing characteristics, X, to recover the marginal WTP for school quality. For the average house, our estimate of 2.6% implies that a one standard deviation increase in average test score increases house prices by about $\$6,500$, holding housing and neighborhood characteristics constant. In percentage terms, this is a very similar estimate to those found in other countries, such as the US, UK or France (see Black & Machin, 2011).

The fact that this estimate is lower than when controlling only for neighborhood fixed effects suggests that households do not only care about their neighborhood at large[20], but also about much more local neighborhood attributes and sort on that basis. This finding reflects that of Bayer *et al.* (2007) in the US.

In contrast with the previous literature (e.g., Bayer *et al.* (2007) and Black (1999)), we find that the OLS specification with controls is downward biased compared to the fixed effects specification. This is likely to be driven by the fact that we study the WTP for the whole country, whereas previous studies have focused on narrower housing markets. In Section 7.2, we study how distinguishing between rural and urban areas changes the direction of the bias, without much effect on our estimate of interest in the fixed effects regression.

6.2 Valuation of Neighborhood Characteristics

We next turn to the valuation of neighborhood characteristics[21]. We contrast two specifications, a standard hedonic price regression and a fixed effect model, to emphasize the role of unobserved heterogeneity at the neighborhood level, which drive the household sorting process. Both models represented in columns (1) and (2) of Table 6 include school quality (average school test scores) as well as the usual set of housing characteristics as regressors.

Estimates from these two models diverge in important ways. A first set of coefficients, including neighborhood average school quality (measured by test scores) as well as average income and education are biased upwards. In particular, the coefficient on education and on neighborhood income are divided by 3.4 and 6.5 respectively. Thus, controlling for unobserved heterogeneity, a one year increase in average parental education at the neighborhood level, increases house prices by 7.8%—significantly less than the 26.6% estimate from the OLS specification. This reflects that such attributes are positively correlated with unob-

[20]Recall that the median size of these neighborhoods are around 0.3 square miles.
[21]These are constructed as reported in Section 4.4

Table 5: Regression Results: Contrast Between OLS and FE Estimates

	OLS	Nbhd Controls	Large Cl. FE	Controls and FE
	(1)	(2)	(3)	(4)
School Quality	0.144***	0.019***	0.042***	0.026***
	(0.005)	(0.003)	(0.003)	(0.002)
Nbhd characteristics	No	Yes	No	Yes
Housing characteristics	No	Yes	No	Yes
Large Cluster-Cohort FE	No	No	Yes	Yes
Cohort FE	Yes	Yes	No	No
Observations	131,951	130,220	131,951	130,220
Adjusted R^2	0.075	0.405	0.299	0.498

Note: Columns (1) and (2) show an OLS specification as a benchmark, while columns (3) and (4) show two different specifications with cluster-by-cohort as fixed effects. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2006 and own a property. Property values are logged and school quality is standardized such that the coefficients can be interpreted as the WTP, in percentage terms, for a one standard deviation increase in school quality. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. Neighborhood characteristics include household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Singleton groups were kept, but results are robust to dropping them, as their number is small. In model (4), 50% of the explained variation is due to the fixed effects, while the remaining is due to the controls. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

served neighborhood quality.

Within our framework, we are also able to precisely estimate household valuation of neighborhood average school quality, separately from school quality itself. This is the case, because some households who live in the same hyper-local neighborhood may send their children to different schools, as they may still be assigned to different school catchment area. On the other hand, households may live in different hyper-local neighborhoods and still send their children to the same school (again because of the shape of the school catchment area). All in all, we see that households value neighborhood school quality above and beyond the school quality of their catchment area. The estimate, standing at 1.7% can be interpreted as reflecting households' valuation of neighborhood peers, beyond that of school peers.

A further set of estimates experience changes in signs through the introduction of unobserved heterogeneity. First, the coefficient on fraction of individuals who have committed

Table 6: Regression Results: Valuation of Neighborhood Characteristics

	OLS	Large Cl. FE
	(1)	(2)
Log HH Gross Income (at Nbhd Level)	0.356***	0.055***
	(0.020)	(0.016)
HH Years of Educ. (at Nbhd Level)	0.266***	0.078***
	(0.006)	(0.007)
School Quality (at Nbhd Level)	0.014***	0.017***
	(0.002)	(0.002)
Share Criminals (at Nbhd Level)	1.384***	−0.247
	(0.231)	(0.231)
Share non-Westerners (at Nbhd Level)	−0.994***	−0.105
	(0.261)	(0.190)
Share Foreigners (at Nbhd Level)	2.761***	0.373**
	(0.236)	(0.170)
Share Private Schools (at Nbhd Level)	0.080***	0.012
	(0.013)	(0.013)
Housing characteristics	Yes	Yes
Large Cluster-Cohort FE	No	Yes
Cohort FE	Yes	No
Observations	130,220	130,220
Adjusted R^2	0.405	0.498

Note: Column (1) shows an OLS specification while column (2) adds cluster-by-cohort fixed effects. Sample includes all parents in Denmark whose children attend 9th grade in public schools between 2002 and 2006 and own a property. Property values are logged and school quality is standardized such that the coefficients can be interpreted as the WTP, in percentage terms, for a one standard deviation increase in school quality. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. Neighborhood characteristics include household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Singleton groups were kept, but results are robust to dropping them, as their number is small. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

a crime adjusts from 1.384 to -.247 and loses significance. This change in sign may reflect the positive correlation between crime and economic activity, which we are able to capture through our neighborhood fixed effects strategy. Moreover, the coefficients on fraction of non-western foreigners change from being, respectively positive and negative, to being insignificant. This again showcases the importance to control for unobserved neighborhood attributes through the fixed effects. However, findings from Bayer *et al.* (2007) may warrant us to be cautious in interpreting this as a mean preference parameter. In fact, given that individuals may be self-segregating, the share of non-westerners would not capitalize directly into house prices as it is not required to clear the market. The amount of capitalization into house prices would then not necessarily reflect mean preferences.

6.3 Valuation of School Characteristics

Parents in Denmark have access to a wide array of information on school characteristics and notably school test scores. In light of the evidence on teacher sorting across neighborhood in Denmark, it is also clear that parents may also place value on teacher quality, which may not be well captured by our measure of test score. Furthermore, parents may also care about the quality of peers themselves at the school level, beyond those that are in their immediate neighborhood.

Our data allow us to further decompose parental valuation of a vector of school characteristics and contrast it with the willingness to pay for neighborhood characteristics. We show in this section that parents not only care about school quality as measured by grades, but also about teacher quality, although to a lesser extent. This could be explained by the more challenging task to observe such characteristics. Furthermore, parents do also place important value on the quality of immediate neighbors, beyond that of their children's peers at school.

We first run a specification with only teacher quality instead of test scores as our measure of school quality. This is to assess whether in contrast with test score as a sole regressor, the capitalization of teacher quality into house prices would be of the same magnitude. Our results presented in Table 7 reveal that this is not the case. The coefficient of 1.6% is around half of that of the coefficient on average test scores presented in Table 5. There are two potential reasons for such result. First, test scores are much more widely publicized and available to

Table 7: Regression Results: Decomposing School Characteristics

	Teacher Quality	Peer Quality
	(1)	(2)
Teacher Quality	0.016***	0.008***
	(0.003)	(0.003)
School Quality		0.010***
		(0.003)
HH Years Schooling (at School Level)		0.022***
		(0.005)
Share Foreigners (at School Level)		−0.078
		(0.049)
Share non-Westerners (at School Level)		0.011
		(0.065)
Log HH Gross Income (at School Level)		0.116***
		(0.024)
Share Married HH (at School Level)		0.085***
		(0.025)
Share Non-Intact HH (at School Level)		−0.041**
		(0.020)
Nbhd characteristics	Yes	Yes
Housing characteristics	Yes	Yes
Large Cluster-Cohort FE	Yes	Yes
Observations	114,735	114,715
Adjusted R^2	0.503	0.504

Note: Column (1) shows a cluster-by-cohort fixed effects specification with teacher quality as only regressor at the school-level. Column (2) further adds a whole vector of school-level covariates. Sample includes all parents in Denmark whose children attend 9th grade in public schools between 2002 and 2006 and own a property. Property values are logged and school quality is standardized such that the coefficients can be interpreted as the WTP, in percentage terms, for a one standard deviation increase in school quality. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. Neighborhood characteristics include household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Singleton groups were kept, but results are robust to dropping them, as their number is small. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

parents than are teacher quality. Second, parents may actually care more about test scores, since it also reflects quality of peers at the school.

In the second column of Table 7 we present estimates of the whole vector of school characteristics. This further comforts the two hypothesis made above. Parents tend to care more about peer quality beyond that of teachers. This is reflected in the positive and significant coefficients on school average income, education and fraction married. In particular, a one percent increase in peer's household income increases house prices by .116%. On the other hand, the coefficient on teacher quality is cut in half, while that on average school test scores is divided by more than two compared to the specification in column (3) of Table 5.

We note that when adding this vector of school-level characteristics to our previous model with neighborhood-level characteristics, the coefficient on the latter regressors stay largely the same.

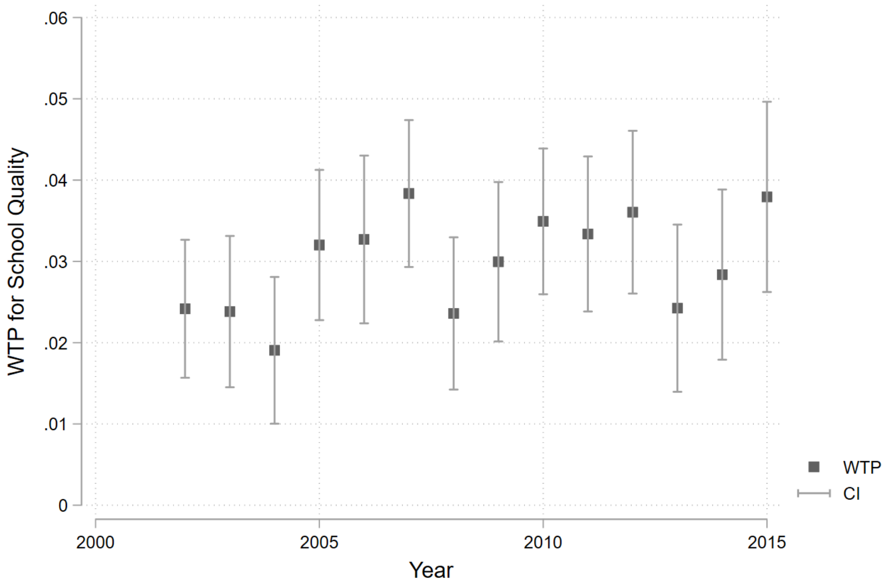
7 Extensions and Sensitivity Analyses

7.1 WTP Over Time

An important benefit from our strategy is that it does not rely on changes over time in school quality. In fact, Kuminoff & Pope (2014) present a theoretical framework demonstrating the issues that arise from using temporal variation in attributes when estimating the WTP. In particular, issues arise because the hedonic price function may not be invariant over time. In this section, we provide evidence that we only get modest fluctuation over time in the marginal WTP for school quality.[22] Figure 5 presents the WTP estimates over the years 2002–2015. To go beyond our estimation sample used in previous sections (2002 to 2006), we expand the mapping of housing units to clusters as explained in Appendix A.5 through probabilistic matching. We fail to reject the null hypothesis that the WTP over the studied years are not statistically different from each other using an F-test with a p-value of 0.56.

[22] Part of the variation in the WTP over time may be due to changes in the boundaries of the school catchment areas.

Figure 5: WTP for School Quality by Cohorts



Notes: This figure shows the WTP for school quality for different cohorts between 2002 to 2015. Sample includes all parents in Denmark whose children attend 9th grade in public schools between 2002 and 2015 and own a property. Property values are logged and school quality is standardized such that the coefficients can be interpreted as the WTP, in percentage terms, for a one standard deviation increase in school quality. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. Neighborhood characteristics include household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

7.2 Further Robustness Checks

In this section, we report several further robustness checks. We show that our main estimates remain robust to a number of sensitivity checks. Results are presented in Table 8 and Table Table 9.

School District Boundaries The first set of sensitivity analyses we present address the potential concern that school boundaries do not actually cross within our large clusters, given the small geographical area they span. Having data on the assigned school district for about 75% of houses in our sample from 2005 to 2015 allows us to verify that 80% of large clusters are composed of at least two distinct school districts. We report in column (1) of Table 8 estimates based on this sample and using the same specification as in (1).

Similarly, about 40% of small clusters do have at least a school boundary crossing. Using these small clusters, we can further address the concern that our estimates are potentially biased due to individuals sorting based on characteristics that are even more local than our previously included hyper-local neighborhood (spanning on average 0.1 square mile) covariates. To do so, we can replace the large fixed effects with small fixed effects in Equation (1). Moreover, we replace the small cluster-level characteristics, with attributes computed at the small-cluster-by-school-district level. Estimate of the WTP for school quality for such specification is reported in column (2) of Table 8. These include a set of controls for housing attributes, as in our main specification.

These two latter specifications closely approximate the idea of BDD initiated by Black (1999). This is the case, since we use only variation in very close proximity to school boundaries. In fact, for the latter specification, which uses small cluster fixed effects, we capitalize on variation that is particularly close to the boundary – retaining variation in school quality and house prices that are no further apart than within a 0.1 square mile cluster. In particular, this ensures that houses are not only close to the boundary, but also that they are in close proximity with each other. This is an important benefit of this methodology.

Treating Defiers. A different set of results we report aims to address concerns regarding the existence of school district defiers in Denmark. In this context, defiers are households which live in a specific school district, but send their children to a school in a different school district. Given our data on school districts for about 75% of the sample, we are able to get a better grasp on the importance of defiers in Denmark, as well as its potential impact on our estimate.

We treat for defiers in two distinct ways.[23] First, we drop all individuals who do not attend the most attended school in a given large-cluster-by-school-district (reported in column (3) of Table 8) or in a given small-cluster-by-school-district (reported in column (4) of Table 8).[24] Second, we drop any cluster-by-school-district which have *any* defiers.This leads to

[23]Our data does allow us to capture schools attended by students. However, we cannot tell whether this school is the one assigned. This arises because we do not have a mapping between school districts and schools. We therefore devise methods based on most attended schools in narrowly defined geographical areas.

[24]These specifications drop approximately 15% of individuals.

drop about 50% of the sample, if based on dropping large clusters with any defiers, while dropping 25% of the sample when dropping small clusters with any defiers. Results are reported in column (5) and (6). In all cases, we see that our estimate is extremely robust to these sensitivity checks.

Table 8: Robustness Checks: School Districts, Small Cluster Fixed Effects and Treating Defiers

	(1)	(2)	(3)	(4)	(5)	(6)
School Quality	.029***	.025***	.022***	.024***	.028***	.026***
adj. R^2	.53	.54	.54	.53	.54	.50
N	85,415	39,799	35,120	18,040	72,019	21,394

Note: Table shows estimates from various robustness checks. The sample includes parents in Denmark whose children attend 9th grade in public schools between 2002 and 2015 and own a property. Column (1) shows results from our main specification, using only clusters where school boundaries are crossing. Column (2) presents results of a specification using small cluster fixed effects. Controls include small-cluster-by-school-district attributes measuring average household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Column (3)-(6) shows our estimates from conducting a subsample analysis aimed at removing defiers, as explained in the text. Property values are logged and school quality is standardized such that the coefficients can be interpreted as the WTP, in percentage terms, for a one standard deviation increase in school quality. Neighborhood and housing characteristics are as above. Large cluster-by-cohort fixed effects are included. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Transaction Data. Next, we look at whether using house prices directly from sales data has an impact on the estimated coefficient. In fact, thus far we have used governmental valuations of housing prices. We show that our estimates are robust to this different measure of our dependent variable (Column (1) of Table 9). We note that since only a fraction of houses are sold on the market every year, the number of observations drops for this analysis.

Lags of Quality. Third, we look at the impact of using past values of our school quality measures. We show in column (2) of Table 9 that our results are rather robust to using lags of school quality.[25]

Copenhagen Metropolitan Area. Finally, we look at the heterogeneity of our results based on different geographic areas. Column (3) of Table 9 shows estimates for the Copen-

[25] Gibbons & Machin (2003) outline the potential endogeneity of school quality when measured by indicators of student performance. A potential test to this is to use lags of school quality, although we note that under serial correlation it would be a weak test.

hagen Metropolitan Area. The coefficient drops slightly to 2.3%, while the adjusted R-squared increases to .59. This showcases the possibility that our fixed effects strategy works better in more urban and denser areas, as we are able to control for more of the unobserved heterogeneity.

Urban and Rural Areas. Turning to columns (4) and (5), again we see that our estimates are very robust to focusing only on urban or rural areas.[26] Interestingly, a specification as in column (3) but without cluster fixed effects would lead to an estimate of the WTP for school quality of 5.1%, whereas in columns (4) and (5), the OLS estimate is downward biased compared to a fixed effects model. This provides evidence of the differential nature of unobserved attributes and their effect on prices and school quality, across places, in Denmark.

Table 9: Robustness Checks: Sales Data, School Quality Lags and Distinct Housing Markets

	(1) Sales	(2) Lags	(3) Cph. Met. Area	(4) Urban	(5) Rural
School Quality	.035***	.027***	.023***	.028***	.027***
adj. R^2	.37	.50	.59	.48	.40
N	17,441	79,087	14,916	95,404	19,900

Note: Table shows estimates from various robustness checks. The sample includes parents in Denmark whose children attend 9th grade in public schools between 2002 and 2015 and own a property. Column (1) shows our estimate from using data on property transactions. Column (2) shows results from replacing our variable for school quality by its second lag (using average test score as measure of school quality).Columns (3), (4) and (5) look at housing markets, respectively focusing on the Copenhagen Metropolitan Area, urban and rural areas. Property values are logged and school quality is standardized such that the coefficients can be interpreted as the WTP, in percentage terms, for a one standard deviation increase in school quality. Neighborhood and housing characteristics are as above. Large cluster-by-cohort fixed effects are included. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

[26] The definition of the United Nations is here used, where urban (as opposed to rural) denotes a built-up area with at least 200 inhabitants, where the distance between the buildings is not more than 200 metres, unless interrupted by public facilities, such as parks.

8 Controlling for Selection into Neighborhoods

In this section, we provide a test for whether our estimates are biased due to households self-selecting into neighborhoods, on the basis of taste dispersion. To this end, Dahl (2002) proposes a methodology to approximately control for selection in a setting with polychotomous choice. It approximates the Heckman (1979) correction model properly extendable in the multidimensional choice context by constructing an analog of the inverse Mill's ratio. This analog turns out to be a polynomial of choice probabilities, under an index sufficiency assumption, as described further in Appendix A.9. Heckman & Vytlacil (2007) show how the selection-corrected estimands under Dahl's (2002) framework can be interpreted as a local average treatment effect (LATE).

We implement this methodology in the context of the WTP for school quality by controlling for selection into neighborhoods in Denmark. Consider an individual i, who makes a choice of neighborhood j amongst M different neighborhood, i.e., the 0.3 square miles neighborhoods we considered as fixed effects in the previous hedonic specifications. Assume i chooses $j = 1$, then in this case we observe individual i's property value only for neighborhood $j = 1$. The hedonic price regression written for individual i, is given by:

$$y_{i1} = \alpha_1 + x_i'\delta_1 + s_i\beta_1 + u_{i1}$$

The choice of a given neighborhood amongst its choice set will be based on an individual's utility denoted by $V_{ij} = z_i'\gamma_j + \eta_{ij}$. Choices of neighborhoods are driven by household-level characteristics z_i, such as income, education, marital status and distance to work, as well as a set of dummies capturing whether each child lives in the same parish as their grandparents. We use the later set of dummies as an exclusion restriction to identify the model.

Dahl (2002) shows that consistent estimation of the WTP can be based on the following model:

$$y_1 = \alpha_1 + x_i'\delta_1 + s_i\beta_1 + \mu(P_{i,i\in S}) + w_{i1}$$

where y_{i1} is log house prices in chosen cluster 1, α_1 is a cluster level specific constant, x_i is a vector of neighborhood characteristics (including housing characteristics), s_i measures the school quality and w_{i1} is an error term. P_k is the probability that any neighborhood k is preferred:

$$P_k = \frac{\exp\left(z'\gamma_k\right)}{\sum_j \exp\left(z'\gamma_j\right)}$$

Based on this model, we present estimates of the WTP for school quality controlling for neighborhood selection, in the next section.

8.1 Estimation Results

In our setting, individuals may in practice choose to live in one neighborhood amongst many hundreds of other neighborhoods. To make the above methodology tractable, we reduce the choice set of individuals in the following sense, based on the assumption that individuals choose where to locate based on average education level. We create 50 quantiles of neighborhood quality based on education levels of parents, and let individuals choose any neighborhood within that quantile of quality in which they currently live.[27]

Then, for each neighborhood in Denmark, we compute the WTP in two ways. First, we compute corrected coefficients using the above methodology. Second, we run a set of neighborhood-level hedonic regressions controlling for hyper-local neighborhood and housing characteristics.

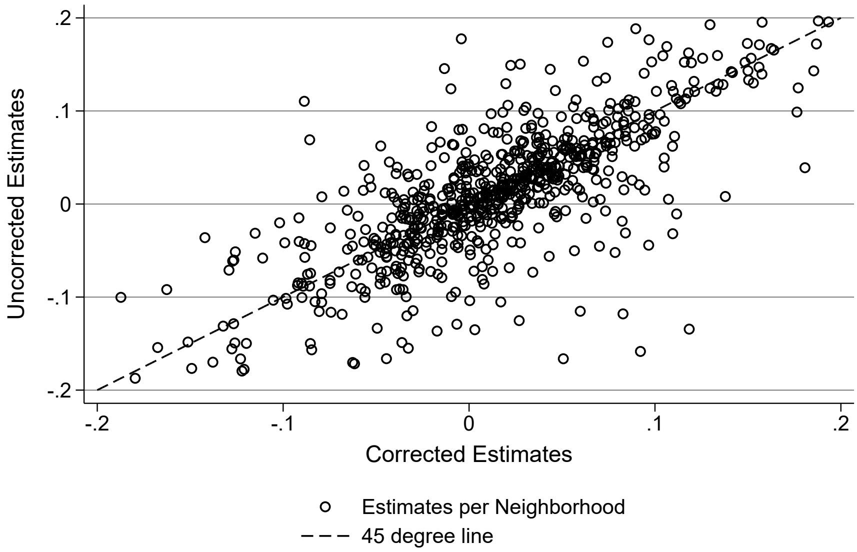
Figure 6, presents a scatter plot of these estimates for 779 neighborhoods. We see that both corrected and uncorrected estimates are highly correlated.

To better assess how these two sets of estimates differ, Figure 7 presents the densities of both the corrected and uncorrected estimates. We see that the density of the corrected co-

[27]Our results are robust to specifying the choice set differently, i.e., by letting individuals choose to live in any neighborhood within the municipality in which they live.

efficients is only very slightly shifted to the left of the density of uncorrected coefficients, providing evidence that our estimates are only slightly impacted by selection into neighborhoods. We fail to reject the null hypothesis that the two distributions are not statistically different from each other using a Kolmogorov–Smirnov test with a p-value of 0.117.

Figure 6: Scatter Plot of Corrected vs. Uncorrected Estimates

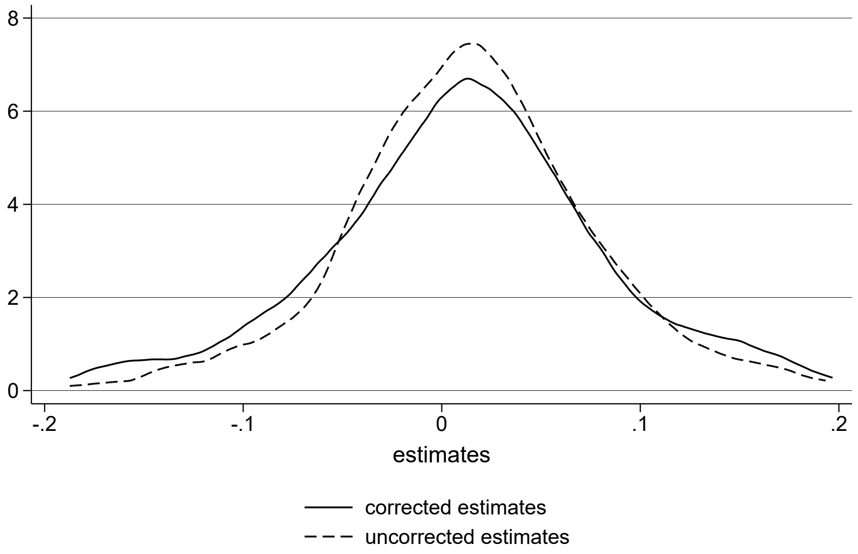


Notes: Scatter plot of corrected on uncorrected estimates of the WTP for school quality. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2015 and own a property. For the uncorrected estimates, we run the same specification as in the second column of Table 5, i.e., including household and neighborhood characteristics (but without large cluster fixed effects), for each large cluster. For the corrected estimates, we control for neighborhood choice, where the selection equation includes individual level controls—income, education, origin of the parents, criminal record, and dummies for whether or not the grandparents live in the same parish. Neighborhood characteristics include household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. The R-squared of a regression of uncorrected estimates on corrected estimates stands at .54. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

9 School Quality & Later Life Outcomes

Our analyses thus far presented evidence that despite an egalitarian redistribution of resources across schools, school quality differences remain (Gensowski *et al.*, 2021) and are capitalized into house prices. Do such parental preferences reflect later life consequences

Figure 7: Distribution of neighborhood-level corrected vs. uncorrected estimates.



Notes: Distribution of neighborhood-level corrected vs. uncorrected estimates of the WTP for school quality. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2015 and own a property. For the uncorrected estimates, we run the same specification as in the second column of Table 5, i.e., including household and neighborhood characteristics (but without large cluster fixed effects), for each large cluster. For the corrected estimates, we control for neighborhood choice, where the selection equation includes individual level controls—income, education, origin of the parents, criminal record, and dummies for whether or not the grandparents live in the same parish. Neighborhood characteristics include household gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools. Housing characteristics include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

anticipated by parents? This is the question we aim to address in this section. We do so by assessing the relationship between school quality and a number of economic, educational and social outcomes, including earnings, college completion, criminal record and teenage pregnancy. Such analysis sheds light on household's valuation for better later life outcomes of their offspring.

To do so, we build on rich longitudinal administrative data from Denmark, which allows us to observe outcomes up to 18 years after completing 9th grade (age 16, on average), for our earliest cohort. Data availability on school quality[28] commences in 2002. We restrict our

[28]Throughout this section, we use a measure of school quality that is slightly different from the one used in our previous analyses of the WTP. Although it is still based on test scores at the school level (based on exams taken in 9th grade), we average over all scores but now exclude one's own test score. This reduces the possibility that our results capture the student's own performance in school.

data to cohorts who complete 9th grade between 2002 and 2006, for the later life earnings analysis. This is to avoid capturing individual outcomes too early in their life cycles.

9.1 Conceptual Framework

Estimating the effect of school quality on later life outcomes requires us to account for several important endogeneity issues, as discussed in the literature (see for instance Dearden *et al.*, 2002b; Heckman *et al.*, 1996; and for a review Meghir & Rivkin, 2011).

First, family location and schooling decisions are not random and are part of a process of life cycle utility optimization. As we have seen above, we have evidence that our strategy allows us to capture relevant characteristics of the sorting process across neighborhoods. We thus build on this strategy to control for observed and unobserved neighborhood-level characteristics that may be correlated with school quality and could otherwise bias the estimates. This would be particularly problematic if parents with stronger preferences to invest in their offspring human capital choose better neighborhoods. Should these neighborhood characteristics relevant to the development of the child's human capital be omitted from the model, estimates of the effect of school quality on later life outcomes would be upward biased. The opposite bias could, however ensue, if parents who locate in better neighborhoods substitute away and invest less at home (as has been recently documented in Denmark by Gensowski *et al.* (2021).

This latter point leads to a second evident endogeneity issue, that of parental investment being related to choice of schools. In fact, parents who locate in catchment areas of better school may also invest more in their child, a relationship that would lead to an upward bias, in absence of the possibility to control for parental investment. In line with this story, our previous analysis on the WTP clearly showcases that there is sorting across school boundaries. To deal with this issue, we first control for a wide range of parental characteristics, such as income, education, family structure and crime. Despite our rich data and ability to control for proxies of parental investment, these may not be enough to capture all differences across parents. To do away with any remaining unobserved differences in parental investment across families, we build on a family fixed effect model of the education production function. More specifically, we look at within variation in school quality amongst family with multiple siblings. A large majority of these siblings attend the same school and neighborhood, but experience different school quality due to differences across cohorts related to differences in peers and teacher quality. We show how similar the results are between the family and neighborhood fixed effects models.

Finally, differential school finances and laws could differentially affect the allocation of resources to students or schools further potentially biasing estimates of the impact of school quality on later life outcomes. We have here provided evidence that school finances are very homogeneous across Danish neighborhoods and schools. Moreover, our focus on within neighborhood variation allows us to fully address this issue.

In Appendix A.10, we formalize our dynamic model of the education production function (drawing on Todd & Wolpin (2003)), which account for the different mechanisms explored above. We show how our reduced-form estimates map to structural parameters and clarify the necessary identifying assumption.

9.2 Data

We turn to describe the different outcomes we analyze.

Economic Outcomes We look at a number of economic outcomes, considering both intensive and extensive margin responses. Specifically, in our main specification we examine total gross income excluding transfers. We contrast this measures with annual disposable income, which is income after taxes, interests, and rental value of owner-occupied housing and stress the impact of redistribution in Denmark.[29] Furthermore, we look

[29] The disposable income is computed as follows: The following items are added: total salary income, remuneration, social security contributions, net profits from self employment, public transfers (social assistance, unemployment benefits, labor market leave, sick leave assistance, labor market activation, child benefits, education grants, housing support, early retirement pension, disability pension, and retirement pension), private pensions paid, interest income and realized gains on securities, and residual income including child support.

at wage growth of individuals, between 2015 and 2019, to explore the income path of individuals.

On the extensive margin, we look at whether individuals who attend better schools are more likely to work. This is defined based on whether individuals earn more than $35,000 a year.[30] Finally, we also explore homeownership as an outcome. An individual is defined as homeowner based on reporting a housing value greater than $20,000.[31]

With regards to data availability, the latest year for data on economic outcomes is in 2019. Since taking a single year would provide a noisy measure of earnings, we average over incomes between 2018 and 2019. This means that for the later cohorts, earnings are measured between ages 28 and 29, while they are measured between ages 32 and 33 for the earliest cohorts. Regarding employment and homeownership, these outcomes are measured between ages 29 and 33, depending on the cohort.

College Completion In Denmark, individuals tend to complete college until late in their 20s and beyond. Therefore, by focusing on cohorts in 9th grade between 2002 and 2006, we are able, for the latest cohort to observe college completion by age 34, given data availability up to 2020.

Criminal Record Our measure of criminal record is dichotomous, considering whether an individual has ever committed any type of crime. We conduct a similar sample restriction when looking at criminal record, by focusing on earlier cohorts who complete 9th grade between 2002 and 2006. Individuals are observed up to ages 30, since data availability on crimes is capped at 2016.

Age at First Birth Another aspect of this analysis pertains to the effect of school quality on

The following items are subtracted: interest expenses, taxes, labor market contributions and special pension, maintenance (contributions) paid to a former spouse as well as to children under age 18. Finally, the estimated rental value of own home is added.

[30]The law does not mandate a minimum wage, as it is negotiated through bargaining between unions and employers association. We set this amount based on the lowest wage agreed upon in an industry of 120 DKK per hour and a workweek of 37.5 hours.

[31]If house is co-owned by partner, this implies a property value of $40,000.

the timing of childbearing. We look at two measures, namely teenage birth and age at first birth. With regards to teenage birth we focus on the whole sample, while for age at first birth we consider the same restriction as above, which is to focus on cohorts between 2002 and 2006. Teenage birth is defined here as having a child by age 19. For this outcome we look at cohorts of children who graduate 9th grade between 2002 and 2015.

9.3 Research Designs

9.3.1 Selection on Observables

Our first approach to estimate our model of the education production function set out in Appendix A.10, builds on the results set out in the previous sections of this paper. In particular, to control for family sorting across neighborhoods, we use the same cluster fixed effects approach together with hyper-local neighborhood-level controls capturing average income, years of education, criminality, share of intact, married and foreign families as well as share of private schools. We also note that given the equalization of expenditures across schools in Denmark, we do not need to add any school-level measures of differential school spending.

We estimate the following model of the education production function:

$$y_{imjkt} = \alpha + \beta S_{imjkt} + \gamma F_{imjkt} + \rho_{kt} + \varepsilon_{imjkt}, \tag{3}$$

where y_{imjkt} denotes our outcome variable (e.g., college completion, earnings, criminal record or teenage pregnancy) of individual i who attends school m in small cluster j within large cluster k in time period t. S_{imjkt} denotes our measure of school quality. F_{imjkt} denotes our vector of housing, neighborhood and individual-level controls. Finally, ρ_{kt} denote neighborhood-by-cohort fixed effects.

Unbiased estimation of (3) relies on the following assumption:

Assumption 2 (Selection on Observables)**.** *School quality is orthogonal to unobserved determi-*

nants of outcomes conditional on our vector of covariates:

$$\mathrm{cov}(S_{imjkt}, \varepsilon_{imjkt} \mid F_{imjkt}, \rho_{kt}) = 0 \tag{4}$$

While this assumption remains strong, our rich data allows us to control for a number of key inputs in the production function, including individual, family and neighborhood-level inputs. Moreover, the results from this specification are in line with those from our within-family specification, which we present next, which lends credence to this assumption.

9.3.2 Within Family Variation

Our second approach, augments equation (3) by including family fixed effects.[32] In the context of the education production function, including family fixed effects are crucial to control for family-specific endowments and investments. We consider the following specification:

$$y_{imjkt} = \alpha + \beta S_{imjkt} + \gamma F_{imjkt} + \nu_t + \mu_l^f + \varepsilon_{imjkt}, \tag{5}$$

where we add family fixed effects μ_l^f for each family l and replace the neighborhood-by-cohort fixed effects with cohort fixed effects ν_t.

Unbiased and consistent estimation of the effect of school quality now relies on the following assumption:

Assumption 3 (Within Family Variation)**.** *School quality is orthogonal to unobserved determinants of outcomes conditional on our vector of covariates, within families:*

$$\mathrm{cov}(S_{imjkt}, \varepsilon_{imjkt} \mid F_{imjkt}, \rho_{kt}, \mu_l^f) = 0 \tag{6}$$

Our key identifying assumption in the family fixed effect model can be restated as follows. Within families, siblings who are the same in terms of observable but who attended schools

[32] We consider here biological parents to define families.

with varying levels of quality, do not differ on average, in terms of unobservables (for instance endowed mental capacity). There are two behavioral restrictions that arise from this modelling. First, parents do not change their input decisions based on innovations on sibling outcomes. Second, input choices may depend on family-specific endowment but not on child-specific endowment. These assumptions may be justified insofar as parents may have limited information on the exact level of the child-specific endowment as well as on the sibling innovations.

It is important to further clarify the source of exogenous variation we use here to identify the effect of school quality. Most siblings in Denmark attend the same school. Still, different cohorts within schools may be exposed to varying quality, due to differing teachers and peers. We note, however, that our approach does not allow us to distinguish whether the quality of the school is being driven by the change in the teacher quality and/or peers.

9.3.3 Within School Variation

To alleviate concerns on the behavioral restrictions implied by Assumption 3, we further propose a third specification, building on within schools and across cohort variation:

$$y_{imjkt} = \alpha + \beta S_{imjkt} + \gamma F_{imjkt} + \eta_m + \nu_t + \rho_k + \varepsilon_{imjkt}, \tag{7}$$

where η_m, ν_t and ρk respectively denote school, cohort and small cluster fixed effects.

Unbiased and consistent estimation of the effect of school quality now relies on the following assumption:

Assumption 4 (Within School Variation)**.** *Changes in school quality, occurring within schools across cohorts, are orthogonal to unobserved determinants of outcomes conditional on our vector of covariates:*

$$\operatorname{cov}(S_{imjkt}, \varepsilon_{imjkt} \mid F_{imjkt}, \eta_m, \nu_t, \rho_k) = 0 \tag{8}$$

Under this specification, variation in school quality comes from changes in teachers and peers in adjacent cohorts within the same school. Three implications are to be noted.

First, Assumption 4 implies that families experiencing idiosyncratic shocks to their children's 9th grade school quality, do not endogenously sort to a different neighborhood to attend a different school. This assumption is plausible given the cost for families to move to a new neighborhood to attend a different school. Moreover, our data on geographical locations of families over time allows us to verify that only very few do move in Denmark when their children reach 9th grade at school.

Second, underlying changes in students' quality may invalidate Assumption 4. To that end, we build on the rich Danish administrative data to control for a host of family specific covariates. We show that our estimates are insensitive to adding these large set of family and children specific characteristics, which include income, education, origins and marital status of parents, order of birth, age of mother at birth, household size and gender of the child.

Finally, school quality changes should be uncorrelated with changes to the neighborhood, which may itself contribute to students' longer term outcomes. To alleviate such concern, we concurrently add small cluster fixed effects to control for time invariant neighborhood characteristics, as well as a vector of small cluster-level covariates, measuring neighborhood-level income, years of education, criminality, share of intact, married and foreign families as well as share of private schools.

9.4 Results

We present our results on college completion in the first row of Table 10. As discussed above, we find that the naive OLS estimates are significantly upward biased. These would imply that a one standard deviation increase in school quality is associated with a 7.7 percentage point increase in college completion. In contrast, the neighborhood fixed effects regression which also includes hyper-local neighborhood controls provide an estimate of 2.9 percentage point, which relative to the sample mean, represents a 8.6% increase. This difference is driven by parents who willingly sort across neighborhoods and thus leads to correlate neighborhood attributes with school quality. When controlling for family fixed effects, in column (3), we find that the coefficient relatively close, at 1.7 percentage point. After adding controls, the coefficient reduces further to 0.71, which at the mean, represents an increase in college completion of 2.1%. The estimate from our fixed effects specification with controls is slightly more muted and stands at 0.29 percentage points.

This evidence shows that parents are willing to pay a significant amount for better schools, and they also recoup this investment through the improved educational outcomes of their offspring.

We now turn to presenting results on the effect of school quality on income in the second row of Table 10. Showing the same set of specification across columns, the effect on later life income ranges from 7.7% to 1.3%, depending on the model.[33] Since Danes tend to graduate from college in their late 20s, it is likely that this effect is a lower bound. Our previous estimates from the WTP for school quality, imply that households are willing to pay $6,500 for a one standard deviation increase in school quality. Results from this section relate this quality increase to approximately $340 to $1020 higher income (at the mean of the later life earnings of our sample) at ages 28–33 for the various models we used. Assuming this same differential in income recurring every year until retire until age 65) for two siblings of the median family yields internal rates of return ranging from 3.7% to 8.0%. To put these numbers in perspective, they are in the range of the estimates of internal rates of return in García *et al.* (2021) for the Perry Preschool Project (PPP)—a pioneering high-quality early childhood education program implemented in the US before Head Start that targeted disadvantaged African-Americans and was evaluated by a randomized trial.

Educational attainment is endogenous to the school quality and may be a mediator for the impact of school quality on income in adulthood. Hence, it is interesting to further examine the impact of school quality on adulthood income, through interactions with educational attainment, in equations (3) and (5). Row (3) and (4) of Table 10 presents results on the effect of school quality on income in adulthood when we interact income with college at-

[33] Using gross including transfers or wage income provides similar results.

Table 10: Regression Results of Models 3, 5 and 7: Impact of School Quality on Various Later Life Outcomes

	Baseline	Controls	Nbhd. FE	HH FE	HH & Cont.	School FE	Sch. & Cont.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) College Completion	0.0768***	0.0241***	0.0293***	0.0165***	0.0071*	0.0094***	0.0029***
	(0.0014)	(0.0014)	(0.0016)	(0.0039)	(0.0040)	(0.0016)	(0.0011)
(2) Income	0.0995***	0.0574***	0.0768***	0.0543***	0.0402*	0.0244***	0.0133**
	(0.0063)	(0.0069)	(0.0081)	(0.0203)	(0.0208)	(0.0065)	(0.0057)
(3) Income (| COL)	0.0149***	0.0004	0.0002	0.0425*	0.0424*	−0.0010	−0.0033
	(0.0055)	(0.0060)	(0.0075)	(0.0239)	(0.0246)	(0.0050)	(0.0050)
(4) Income (| HS)	0.0460***	0.0625***	0.0812***	0.0744*	0.0780**	0.0251***	0.0183**
	(0.0103)	(0.0107)	(0.0125)	(0.0382)	(0.0391)	(0.0092)	(0.0086)
(5) Disposable Income	0.0153***	0.0104***	0.0127***	0.0048	−0.0030	0.0049***	0.0026*
	(0.0020)	(0.0019)	(0.0022)	(0.0070)	(0.0066)	(0.0016)	(0.0014)
(6) Income Growth	1789***	536***	622***	567***	406**	191***	95.2**
	(60)	(59)	(67)	(191)	(195)	(50)	(43)
(7) Employment	0.0111***	0.0066***	0.0100***	0.0052	0.0044	0.0042***	0.0024**
	(0.0012)	(0.0013)	(0.0015)	(0.0039)	(0.0041)	(0.0011)	(0.0010)
(8) Homeownership	0.0019	0.0096***	0.0096***	0.0031	0.0006	0.0039***	0.0023**
	(0.0015)	(0.0014)	(0.0016)	(0.0042)	(0.0044)	(0.0014)	(0.0011)
(9) Crime	−0.0395***	−0.0126***	−0.0178***	−0.0158***	−0.0069*	−0.0061***	−0.0025***
	(0.0013)	(0.0013)	(0.0015)	(0.0040)	(0.0039)	(0.0012)	(0.0010)
(10) Teenage Birth	−0.0021***	−0.0013***	−0.0016***	−0.0006**	−0.0008*	−0.0006***	−0.0004*
	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0005)	(0.0002)	(0.0002)
(11) Age at First Birth	0.2346***	0.1019***	0.1016***	0.1580**	0.1307*	0.0293***	0.0182*
	(0.0127)	(0.0133)	(0.0172)	(0.0803)	(0.0787)	(0.0110)	(0.0103)
Individual Char.	No	Yes	Yes	No	Yes	No	Yes
Nbhd Char.	No	Yes	Yes	No	Yes	No	Yes
Housing Char.	No	Yes	Yes	No	Yes	No	Yes
Cohort FE	Yes	Yes	No	Yes	Yes	Yes	Yes
Family FE	No	No	No	Yes	Yes	No	No
Large Cluster-Cohort FE	No	No	Yes	No	No	No	No
School FE	No	No	Yes	No	No	Yes	Yes
Small Cluster FE	No	No	Yes	No	No	No	Yes
Observations		Approx. 170,000		Approx. 40,000		Approx. 200,000	

Note: Column (1) shows a specification without controls as a benchmark, while column (2) shows a specifications with a host of individual, neighborhood and housing characteristics. Individual controls include income, education, origins and marital status of parents, order of birth, age of mother at birth, household size and gender of the child. Neighborhood and housing characteristics are as in the previous hedonic models. We do not include origins of the parents in the within family strategy. Column (3) shows a neighborhood FE fixed effects specification, with controls. Column (4) and (5) depict our family fixed effect models, respectively, without and with controls. Finally columns (6) and (7) respectively report the school fixed effects specification, without and with controls. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2006. For teenage birth, the sample includes cohort of children from 2002 to 2015. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

tainment. Our estimates provide evidence that a one standard deviation increase in school quality leads to increase later life income by 4.2% for those who complete college, whereas the effect is greater for those who only complete high school, at 7.8%, in the family fixed effects specification. These results are more muted in the within school specification, albeit showing similar heterogeneity.

To explore the role of redistribution in Denmark, we consider disposable income as a further outcome variable in column (6) of Table 12.[34] In contrast to our previous gross measure of income, we see that there is here a precisely estimated null effect of school quality on disposable income. This speaks to the previous literature which stresses the importance of the tax system in Denmark to reduce inequality (see for instance Landersø & Heckman, 2017). Thus, although gross income is increased by better schools, this does not translate into much higher disposable income during the early stages in the working life cycle. The within school specification provides a similar conclusion, where the estimate stands at a mere 0.026% and is significant at the 10% significance level.

In columns (7) and (8), we evaluate the role of school quality, on the extensive margin, looking at the impact on the decision to take up work and to become homeowner. Results suggest that better schools play a small positive role here, by increasing the probability of working by between 1.0 to .24 percentage points (albeit noisily estimated for the family fixed effects specification). Results for homeownership are very similar. The school fixed effect model points to a small effect of 0.23 percentage point increase. These small effects may be driven by the fact that we observe these outcomes early in the life cycle of individuals.

Turning to crime in row (9), we see again the naive OLS specifications are strongly biased. When adding neighborhood and housing controls as well as large cluster fixed effects, we see that the effect of school quality is more muted—a one standard deviation increase in school quality leads to a 1.8 percentage point decrease in criminality. In the family and school fixed effect model, the estimate respectively stand at -0.7 and -0.25 percentage point decrease – the latter represents close to a 1% drop in crime at the mean.

In row (10), we provide evidence regarding teenage birth. The effect from both the household and school fixed effects specification imply that attending schools with one standard de-

[34]Using net of tax income provides similar estimates.

viation higher average grades leads to decrease teenage birth by between 4-9% at the mean.

Finally, column (11) looks at whether higher school quality is associated with a later arrival of children. Both model estimates presented in columns (3) and (5) provide similar estimates, reflecting a 0.10–0.13 year increase in the age at first birth, while the school fixed effects specification reflects a less pronounced effect, with an estimate standing at 0.018.

9.5 Heterogeneity

We turn to assessing the extent to which our results are heterogeneous across family characteristics. To this end, we run the same two specifications, namely the within neighborhood and within family models and include all family, neighborhood and individual-level controls. Due to our limited sample sizes in the heterogeneity analysis, we are not able to get precisely estimated coefficients in the within family model. Therefore, we report below estimated coefficients from the within neighborhood model. Although the latter specification relies on a stronger assumption, the fact that estimates presented in the previous section provide estimates in the same order of magnitude, lends credence to its validity.

The heterogeneity analysis looks at how the effect of school quality varies along several dimensions presetend in Table 11.

Table 12 presents results from the heterogeneity analysis. Each coefficient is estimated from a separate regression of an outcome on school quality conditional on a large set of covariates, within neighborhoods.

Table 11: Individual and Household Characteristics Definition for Heterogeneity Analysis

Gender	Columns (1) and (2) of Table 12 explore heterogeneity by gender of the children. In our within family specification this implies that we necessarily focus on families with at least two daughters or two sons.
Household Income	Columns (3), (4) and (5) split our sample in three categories, parents whose gross household income excluding transfers is below the first quartile, those who earn between the first and third quartile, and those who earn above the third quartile.
Household Education	Columns (6) and (7) consider heterogeneity based on the maximum level of education of the parents being either high school or college.
Origins	Columns (8) and (9) assess the role of parental origins, splitting the sample based on whether any parent or grandparent is foreign, versus having both parents with Danish ancestry.
Family Structure	Finally, columns (10) and (11) look at the role of family structure, in particular whether being raised in a non-intact family has a differential impact. We define non-intact households as those whose parents have separated during the first 18 years of the child.

Table 12: REGRESSION RESULTS: HETEROGENEITY

	Gender		Income			Education		Origins		Fam. Struc.	
	(1) Female	(2) Male	(3) Low	(4) Med Inc	(5) High	(6) HS	(7) COL	(8) Dane	(9) Foreign	(10) Non-Intact	(11) Intact
College	0.0043***	0.0051***	0.0028	0.0033***	0.0069***	0.0044***	0.0027*	0.0043***	0.0047	0.0038*	0.0043***
	(0.0012)	(0.0012)	(0.0018)	(0.0012)	(0.0022)	(0.0010)	(0.0015)	(0.0008)	(0.0075)	(0.0020)	(0.0009)
Income	0.0557***	0.0347***	0.0802***	0.0557***	0.0087	0.0685***	0.0172	0.0541***	0.0182	0.0778***	0.0419***
	(0.0141)	(0.0128)	(0.0253)	(0.0135)	(0.0194)	(0.0117)	(0.0153)	(0.0091)	(0.0672)	(0.0244)	(0.0097)
Crime	−0.0097***	−0.0219***	−0.0155***	−0.0138***	−0.0162***	−0.0148***	−0.0163***	−0.0146***	−0.0251**	−0.0239***	−0.0113***
	(0.0021)	(0.0027)	(0.0040)	(0.0024)	(0.0038)	(0.0021)	(0.0027)	(0.0017)	(0.0108)	(0.0040)	(0.0018)
Teen Birth	−0.0022***	−0.0007**	−0.0026***	−0.0016***	−0.0000	−0.0018***	−0.0006*	−0.0015***	−0.0015	−0.0038***	−0.0006**
	(0.0005)	(0.0003)	(0.0010)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0017)	(0.0008)	(0.0003)

Note: This Table presents results from the heterogeneity analysis where we estimate equations (3) with different sets of controls and on different split samples. Column (1) and (2) contrasts the effect on Females versus Males. Column (3) and (4) look at heterogeneity based on household income. Columns (6) and (7) turn to parental education heterogeneity, while columns (8) and (9) look at the differential role of origins. Finally, columns (10) and (11) provide results contrasting non-intact and intact families. The estimated models include a host of individual, neighborhood and housing characteristics. Individual controls include income, education, origins and marital status of parents, order of birth, age of mother at birth, household size and gender of the child. Neighborhood and housing characteristics are as in the previous hedonic models. Standard errors corrected for clustering at the school-cohort level are reported in parentheses. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2006. For teenage birth, the sample includes cohorts of students from 2002 to 2015. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Several lessons emerge from this analysis. First, there is a positive parental income gradient in terms of the impact of school quality on college completion. In fact, the estimated effect of school quality on college completion for low income parents stands at 0.28 percentage point and is insignificant, while it rises to 0.69 for children of high income parents. Apart from this heterogeneity in the returns by parental income, it seems that the impact of school quality on college outcomes is rather homogeneous across the various characteristics we consider.

Interestingly, the parental income gradient is here opposite to that of college completion. We see that the effect of school quality on income goes from 8% for low income families to being insignificantly different from zero for children of wealthier households. In the same vein, offspring of parents with high school degrees experience greater benefits to school quality than those whose parents have college degrees. This is similarly the case when looking at children with parents of Danish origins compared to those whose parents come from abroad.

The impact of school quality on crime seems to be rather homogeneous across the categories we study. Still, we see that the effect is greater for male compared to female, as well as for children who live in non-intact household, compared to those who live in intact households.

Overall, children from low income households seem to benefit relatively more, with greater increase in later life incomes, greater reduction in criminal activity and teenage birth, compared to higher income families. Similarly, offspring from less educated households showcase greater relative benefits from attending better schools, particularly with respect to college attendance, greater later life income and reduction in the probability of teenage birth.

10 Conclusion

The Scandinavian welfare state is often touted as an exemplary system to reduce inequalities and equalize opportunities, inter alia, by providing an education system that is free for all. Yet, despite an equalized school expenditure and teacher salary distribution, there exist substantial differences between schools in terms of quality of teachers and the skill levels of the peer students. These differences are, in part, due to residence-based assignment of students to public schools along sorting of families and teachers across neighborhoods. More advantaged families sort into neighborhoods where school quality is higher. We provided evidence that access to better schools through residential choices is capitalized into house prices. Using rich longitudinal administrative data from Denmark, we develop a novel empirical strategy to estimate the marginal willingness to pay (WTP) for schools in Denmark, where public schools are free. Our main results reveal that households are willing to pay around 3% of house prices for a one standard deviation increase in school quality. Measured as percentage of income, our results indicate that the willingness to pay for school quality is higher at tails of the income distribution. Our result is robust to various specification and robustness checks. Specially, we provide evidence that our estimates are not biased due to households self-selecting into neighborhoods.

References

Altonji, Joseph G, & Dunn, Thomas A. 1996. Using siblings to estimate the effect of school quality on wages. *The review of economics and statistics*, 665–671.

Bajari, Patrick, & Benkard, C. Lanier. 2005. Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of political economy*, **113**(6), 1239–1276.

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

Bernal, Pedro, Mittag, Nikolas, & Qureshi, Javaeria A. 2016. Estimating effects of school quality using multiple proxies. *Labour economics*, **39**, 1–10.

Betts, Julian R. 1995. Does school quality matter? evidence from the national longitudinal survey of youth. *The review of economics and statistics*, 231–250.

Betts, Julian R. 1996. Is there a link between school inputs and earnings. *Fresh scrutiny of an old literature*, 141–191.

Black, Sandra E. 1999. Do better schools matter? Parental valuation of elementary education. *Quarterly journal of economics*, **114**(2), 577–599.

Black, Sandra E., & Machin, Stephen. 2011. *Housing Valuations of School Performance*. 1 edn. Vol. 3. Elsevier B.V.

Bogart, William T., & Cromwell, Brian A. 2000. How much is a neighborhood school worth? *Journal of urban economics*, **47**(2), 280–305.

Bourguignon, François, Fournier, Martin, & Gurgand, Marc. 2007. Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. *Journal of economic surveys*, **21**(1), 174–205.

Brasington, David, & Haurin, Donald. 2006. Educational outcomes and house values: A test of the value added approach*. *Journal of regional science*, **46**(2), 245–268.

Brasington, David M. 1999. Which Measures of School Quality Does the Housing Market Value? *Journal of real estate research*, **18**(3), 395–414.

Bénabou, Roland. 1993. Workings of a city: Location, education, and production. *The quarterly journal of economics*, **108**(3), 619–652.

Bénabou, Roland. 1996. Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance. *The american economic review*, **86**(3), 584–609.

Card, David, & Krueger, Alan B. 1992. Does school quality matter? returns to education and the characteristics of public schools in the united states. *Journal of political economy*, **100**(1), 1–40.

Chetty, Raj, Friedman, John N, Hilger, Nathaniel, Saez, Emmanuel, Schanzenbach, Diane Whitmore, & Yagan, Danny. 2011. How does your kindergarten classroom affect your earnings? evidence from project star. *The quarterly journal of economics*, **126**(4), 1593–1660.

Chetty, Raj, Friedman, John N, & Rockoff, Jonah E. 2014. Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American economic review*, **104**(9), 2633–79.

Clark, Damon. 2010. Selective schools and academic achievement. *The be journal of economic analysis & policy*, **10**(1).

Coleman, James S. 1968. Equality of educational opportunity. *Integrated education*, **6**(5), 19–28.

Cullen, Julie Berry, Jacob, Brian A, & Levitt, Steven D. 2005. The impact of school choice on student outcomes: an analysis of the chicago public schools. *Journal of public economics*, **89**(5-6), 729–760.

Dahl, Gordon B. 2002. Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica*, **70**(6), 2367–2420.

Damm, Anna Piil, & Schultz-Nielsen, Marie Louise. 2008. Danish neighbourhoods: Construction and relevance for measurement of residential segregation. *Nationaloekonomisk tidsskrift*, **146**(3), 241–262.

Dearden, Lorraine, Ferri, Javier, & Meghir, Costas. 2002a. The effect of school quality on educational attainment and wages. *Review of economics and statistics*, **84**(1), 1–20.

Dearden, Lorraine, Ferri, Javier, & Meghir, Costas. 2002b. The effect of school quality on educational attainment and wages. *The review of economics and statistics*, **84**(1), 1–20.

Deming, David J, Hastings, Justine S, Kane, Thomas J, & Staiger, Douglas O. 2014. School choice, school quality, and postsecondary attainment. *American economic review*, **104**(3), 991–1013.

Downes, Thomas, & Zabel, Jeffrey. 2002. The impact of school characteristics on house prices: Chicago 1987-1991. *Journal of urban economics*, **52**(1), 1–25.

Durlauf, Steven. 1996. A Theory of Persistent Income Inequality. *Journal of economic growth*, **1**(1), 75–93.

Ekeland, Ivar, Heckman, James J, & Nesheim, Lars. 2004. Identification and Estimation of Hedonic Models Source. *Journal of political economy*, **112**(February), S60–S109.

Epple, Dennis. 1987. Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. *Revealed preference approaches to environmental valuation volumes i and ii*, **95**(1), 31–52.

Eshaghnia, Sadegh, Heckman, James J., Landersø, Rasmus, & Qureshi, Rafeh. 2021. *The intergenerational transmission of lifetime wellbeing*.

Eshaghnia, S.M. Sadegh. 2020. *Re-visiting long-run neighborhood effects: causality or selection and sorting*. Unpublished manuscript, University of Chicago, Department of Economics.

Fisher, Franklin M. 1966. *The identification problem in econometrics*. New York: McGraw-Hill.

García, Jorge Luis, Bennhoff, Frederik H, Leaf, Duncan Ermini, & Heckman, James J. 2021 (July). *The dynastic benefits of early childhood education*. Working Paper 29004. National Bureau of Economic Research.

Gensowski, Miriam, Landersø, Rasmus, Bleses, Dorthe, Dale, Phillip, Højen, Anders, & Justice, Laura. 2021. *Public and parental investments, and children's skill formation*. Tech. rept. Rockwool Foundation Working Paper.

Gibbons, Stephen, Machin, Stephen, & Silva, Olmo. 2013. Valuing school quality using boundary discontinuities. *Journal of urban economics*, **75**(C), 15–28.

Gibbons, Steve, & Machin, Stephen. 2003. Valuing english primary schools. *Journal of urban economics*, **53**(2), 197–219.

Glewwe, Paul W, Hanushek, Eric A, Humpage, Sarah D, & Ravina, Renato. 2011. *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010*.

Hanushek, Eric A. 1986. The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of economic literature*, **24**(3), 1141–1177.

Hanushek, Eric A. 1997. Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis*, **19**(2), 141–164.

Hanushek, Eric A, Rivkin, Steven G, & Taylor, Lori L. 1996. *Aggregation and the estimated effects of school resources*.

Hastings, Justine S, & Weinstein, Jeffrey M. 2007. *No child left behind: Estimating the impact on choices and student outcomes*.

Heckman, James, Layne-Farrar, Anne, & Todd, Petra. 1996. Human capital pricing equations with an application to estimating the effect of schooling quality on earnings. *The review of economics and statistics*, **78**(4), 562–610.

Heckman, James J. 1979. Sample selection bias as a specification error. *Econometrica*, **47**(1), 153–162.

Heckman, James J, & Landersø, Rasmus. 2021 (March). *Lessons from denmark about inequality and social mobility*. Working Paper 28543. National Bureau of Economic Research.

Heckman, James J., & Vytlacil, Edward J. 2007. Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative economic estimators to evaluate social programs, and to forecast their effects in new environments. *Chap. 71, pages 4875–5143 of:* Heckman, James J., & Leamer, Edward E. (eds), *Handbook of econometrics*, vol. 6B. Amsterdam: Elsevier B. V.

Heckman, James J, Layne-Farrar, Anne, & Todd, Petra E. 1995. *Does measured school quality really matter? an examination of the earnings-quality relationship*.

Heckman, James J., Matzkin, Rosa, & Nesheim, Lars. 2010. Nonparametric Identification and Estimation of Nonadditive Hedonic Models. *Econometrica*, **78**(5), 1569–1591.

Heyns, Barbara. 1997. Does money matter? the effect of school resources on student achievement and adult success. *Contemporary sociology*, **26**(4), 488.

Jackson, C Kirabo. 2013. Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of economics and statistics*, **95**(4), 1096–1116.

James, David R, & Taeuber, Karl E. 1985. Measures of segregation. *Sociological methodology*, **15**, 1–32.

Johnson, George E, & Stafford, Frank P. 1973. Social returns to quantity and quality of schooling. *Journal of human resources*, 139–155.

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

Landersø, Rasmus, & Heckman, James J. 2017. The scandinavian fantasy: The sources of intergenerational mobility in denmark and the us. *The scandinavian journal of economics*, **119**(1), 178–230.

Lee, Barrett A, Reardon, Sean F, Firebaugh, Glenn, Farrell, Chad R, Matthews, Stephen A, & O'Sullivan, David. 2008. Beyond the census tract: Patterns and determinants of racial segregation at multiple geographic scales. *American sociological review*, **73**(5), 766–791.

Meghir, Costas, & Rivkin, Steven. 2011. Econometric methods for research in education. 1 edn., vol. 3. Elsevier.

OECD. 2016. Denmark. *OECD Publishing*, October.

Pielou, Evelyn C. 1977. *Mathematical ecology*. Wiley.

Reardon, Sean F, & Bischoff, Kendra. 2011. Income inequality and income segregation. *American journal of sociology*, **116**(4), 1092–1153.

Reardon, Sean F, Yun, John T, & Kurlaender, Michal. 2006. Implications of income-based school assignment policies for racial school segregation. *Educational evaluation and policy analysis*, **28**(1), 49–75.

Reardon, Sean F, Farrell, Chad R, Matthews, Stephen A, O'Sullivan, David, Bischoff, Kendra, & Firebaugh, Glenn. 2009. Race and space in the 1990s: Changes in the geographic scale of racial residential segregation, 1990–2000. *Social science research*, **38**(1), 55–70.

Ries, John, & Somerville, Tsur. 2010. School Quality and Residential Property Values: Evidence from Vancouver Rezoning. *The review of economics and statistics*, **92**(4), 928–944.

Rivkin, Steven G, Hanushek, Eric A, & Kain, John F. 2005. Teachers, schools, and academic achievement. *Econometrica*, **73**(2), 417–458.

Rockoff, Jonah E. 2004. The impact of individual teachers on student achievement: Evidence from panel data. *American economic review*, **94**(2), 247–252.

Rosen, Sherwin. 1974. Hedonic Prices and Implicit Markets : Product Differentiation in Pure Competition. *Journal of political economy*, **82**(1), 34–55.

Theil, Henri, & Finizza, Anthony J. 1992. A note on the measurement of racial integration of schools by means of informational concepts. *Pages 1033–1040 of: Henri theil's contributions to economics and econometrics*. Springer.

Theil, Henry. 1972. *Statistical decomposition analysis; with applications in the social and administrative sciences*. Tech. rept.

Tiebout, Charles M. 1956. A pure theory of local expenditures. *Journal of political economy*, **64**(5), 416–424.

Tinbergen, Jan. 1956. On the theory of income distribution. *Weltwirtschaftliches archiv*, **77**, 155–175.

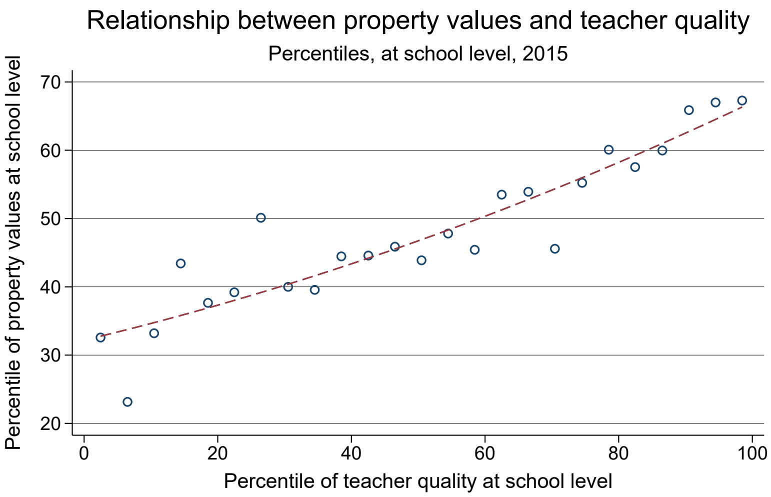
Todd, Petra E., & Wolpin, Kenneth I. 2003. On The Specification and Estimation of The Production Function for Cognitive Achievement. *Economic journal*, **113**(485), 3–33.

Zoloth, Barbara S. 1976. Alternative measures of school segregation. *Land economics*, **52**(3), 278–298.

A Appendix

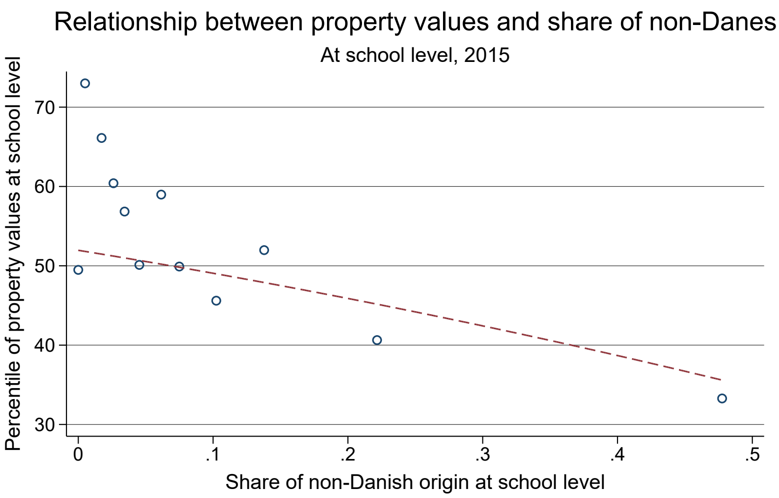
A.1 Sorting

Figure A.1: Relationship between property values and teacher quality



Notes: The figure shows a binned scatterplot with quadratic fit of property values percentiles on estimated teacher quality percentiles. Both measures are computed at school level. To proxy average teacher quality in a given school, we use a unique link between all school teachers in Denmark, the schools they work in, and the children that attend those schools. The multiple dimensions of teacher characteristics are condensed to an index ranging from 0 (the lowest quality teacher by observable characteristics) to 1 (the highest quality teacher), as explained in the main text.

Figure A.2: Origin of students and property values



Notes: The figure shows a binned scatterplot, with a quadratic fit, of property values percentiles on share of students with non-Danish origins at the school level in 2015.

A.2 Correlation Matrix of Neighborhood Attributes

Table A.1 reports the correlation between different neighborhood attributes.

Table A.1: Correlation Between Neighborhood Attributes

	Log HH Gross Income	HH Years of Educ.	School Quality	Share Criminals	Share Married HH	Share Foreigners	Share non-Westerners	Share Private Schools
Log HH Gross Income	1							
HH Years of Educ.	0.838***	1						
School Quality	0.431***	0.490***	1					
Share Criminals	-0.196***	-0.207***	-0.203***	1				
Share Married HH	0.662***	0.360***	0.197***	-0.275***	1			
Share Foreigners	-0.238***	-0.0401***	-0.175***	0.301***	-0.390***	1		
Share non-Westerners	-0.290***	-0.136***	-0.232***	0.304***	-0.337***	0.958***	1	
Share Private Schools	0.0589***	0.131***	0.102***	0.0322***	-0.0769***	0.0954***	0.0657***	1
Share Non-Intact HH	-0.316***	-0.217***	-0.194***	0.126***	-0.347***	0.0703***	0.0595***	0.00251

Note: This table reports the correlation between the different small cluster level neighborhood attributes we consider in our main strategy. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

A.3 Hedonic Theory

Consider the following hedonic regression model, omitting time variation, and after demeaning observable housing and neighborhood characteristics:

$$p_{imj} = \alpha + \beta S_{mj} + \varepsilon_{imj}, \tag{9}$$

where p_{imj} denotes log property values of individual i who attends school m in cluster j. S_{mj} denotes our measure of school quality for school m in cluster j. Finally ε_{imj} represents unobserved neighborhood and housing attributes that are assumed to be iid.

Assuming that houses within the same cluster share the same unobservable neighborhood and housing characteristics, we can estimate the mean marginal WTP for school as the difference in average house price across school boundaries, normalized by the difference in school quality over the two sides, within each cluster (provided that school boundaries cross within clusters).

Still, one may be wary that this estimate of the mean marginal WTP is biased due to sorting based on individual preferences. For instance, if individuals who value highly school quality live in areas with better schools, the marginal WTP may reflect the preferences of this sub-population. With heterogeneous tastes, the marginal WTP recovered may not align with the average marginal WTP.

One way to tackle this issue, is to essentially assume a certain level of homogeneity in preferences.

Let $U_j = U(\theta_j, X_j)$, where θ_j are individual preferences and X_j are neighborhood characteristics. The WTP is equal to the MRS between income paid (price) and the amenity (schooling).

Consider a consumer's utility and budget constraint:

$$U(\theta_j, X_j)$$
$$\sum_{i=1}^{I_j} P_{i,j} X_{i,j} = Y(i,j)$$

where $Y(i,j)$ represents income of i in j and one $X_{i,j}$ is residual consumption.

Maximizing the consumer's utility function given the budget constraint, we have:

$$\frac{\frac{\partial U}{\partial X_{i,j}}}{\frac{\partial U}{\partial P_{i,j}}} = \text{willingness to pay within } j$$

Consumption is as follows:

$$C_{i,j} = Y^*_{i,j} - \sum_{i=1}^{I_{j-1}} P_{i,j} X_{i,j}$$

Within a cluster j, we assume that characteristics X_j and preferences of consumers U_j are correlated. Moreover, within a cluster j, suppose the distribution of U_j collapses to point mass (common preferences $U(\theta_j, X') = U(\theta_j, X)$, for all X, for θ_j as a vector of preference parameters). Then, this is akin to a fixed effect where prices trace out an isoutility curve:

$$U(\theta_j, X') = U(\theta_j, X) \text{ all } X, X'.$$

Thus, in the special case where we assume homogeneity in preferences, the WTP recovered through a hedonic price regression reflects the willingness to pay of the population.

A.3.1 Preference Heterogeneity

Assuming homogeneity of preferences, even within clusters, may however be a too strong assumption to make. Bayer *et al.* (2007) provides a framework to theoretically underpin the relationship between hedonic price regression estimates under heterogeneous preferences.

In the context of heterogeneity, one can recover estimates of the WTP in a hedonic framework that are in line with mean preferences, when the attribute is supplied at different levels of quality in numerous different locations. Otherwise there may be significant divergence. For instance, the variation in house prices with a view of the Golden Gate bridge may not reveal mean preference for a good view, given that it is individuals at the margin, with strong preferences for a view, who drive up the price.

To clarify when the coefficients in the hedonic price regression likely provide an approximation to the mean marginal WTP (mean MWTP) of the population, we present the following heterogeneous model of residential sorting. Each household chooses its residence to maximize its indirect utility function V_h^i:

$$\max V_h^i = \alpha_X^i X_h - \alpha_p^i p_h - \alpha_d^i d_h^i + \theta_{bh} + \xi_h + \varepsilon_h^i$$

where X_h denotes observable characteristics of housing choice h, p_h is the price of housing choice h, d_h^i denote the distance from residence h to work, and finally θ_{bh} are a set of cluster fixed effects. We allow each household's valuation of choice characteristics to vary with individual's i own characteristics z^i:

$$\alpha_j^i = \alpha_{0j} + \sum_{k=1}^{K} \alpha_{kj} z_k^i$$

where $j \in \{X, Z, d, p\}$

$$\begin{aligned} V_h^i \quad = \quad & \underbrace{\alpha_{0X} X_h - \alpha_{0p} p_h + \theta_{bh} + \xi_h}_{\delta_\mathbf{h}} \\ & \underbrace{+ \left(\sum_{k=1}^{K} \alpha_{kX} z_k^i \right) X_h - \left(\sum_{k=1}^{K} \alpha_{kp} z_k^i \right) p_h - \left(\sum_{k=1}^{K} \alpha_{kd} z_k^i \right) d_h}_{\lambda_\mathbf{h}^\mathbf{i}} \\ & + \varepsilon_h^i \end{aligned}$$

Where δ_h : mean indirect utility provided by housing choice h

The first step of the estimation procedure recovers estimates of δ_h and parameters in λ

by finding the likelihood that maximizes the probability that the model correctly matches each household with its chosen housing choice. Assume ε_h^i is drawn from extreme value distribution, then:

$$P_h^i = \frac{\exp\left(\delta_h + \lambda_h^i\right)}{\sum_k \exp\left(\delta_k + \lambda_k^i\right)}$$

The log likelihood function is as follows:

$$\mathcal{L} = \sum_i \sum_h I_h^i \ln\left(P_h^i\right)$$

We can search over the parameters in λ and the vector of mean indirect utilities δ_h to maximize the log likelihood function.

The second stage decomposes δ into observable and unobservable components. Recall that:

$$\delta_h = \alpha_{0X} X_h - \alpha_{0p} p_h + \theta_{bh} + \xi_h$$

Then we have:

$$p_h + \frac{1}{\alpha_{0p}}\delta_h = \frac{\alpha_{0X}}{\alpha_{0p}} X_h + \frac{1}{\alpha_{0p}}\theta_{bh} + \frac{1}{\alpha_{0p}}\xi_h \tag{10}$$

Therefore, in presence of heterogeneous preferences, the mean indirect utility δ_h estimated in the first stage provides an adjustment to the hedonic price equation.

When households are homogenous, MWTP curve is horizontal line, then δ_h "correction" disappears. In fact:

$$\frac{\partial \mathcal{L}}{\partial \delta_h} = 1 - \sum_i \left(P_h^i\right) = 0 \Rightarrow \sum_i \left(P_h^i\right) = 1 \text{ for all } h$$

This implies that the ML estimates of δ_h must be identical (equal to constant K) for all houses. Then equation (2) simply becomes the hedonic regression (1) above.

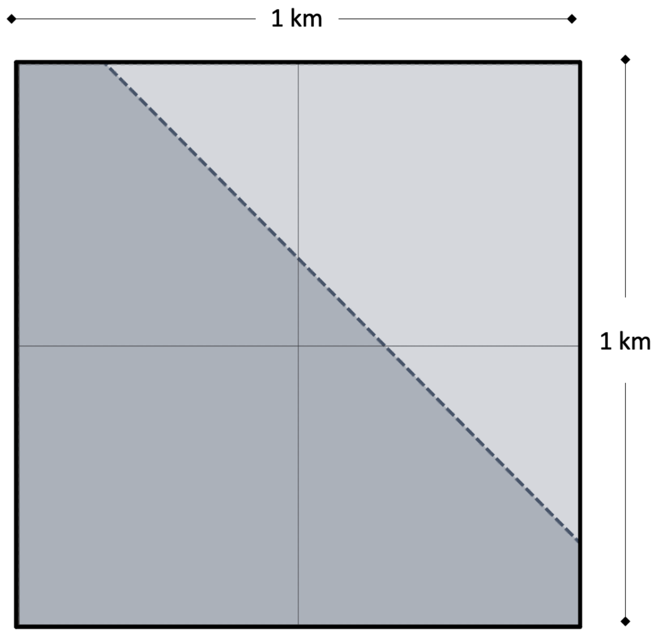
The model shows that a hedonic price regression could be rationalized under this heterogenous sorting model and provide an estimate of the mean MWTP, specifically "if there are roughly an equal number of students in each school, averaging the equilibrium price over all the houses in the sample corresponds roughly to the mean MWTP of all households. Consequently, for attributes that vary more continuously throughout the region, there is likely to be only a slight difference between the mean preferences estimated in the heterogeneous sorting model and the coefficients of the hedonic price regression." The first step of the estimation strategy, a multinomial logit model of housing choices, effectively provides an adjustment to the hedonic price regression by accounting for differences in valuation between the mean and marginal households—a necessary adjustment for attributes in limited supply in the context of heterogeneous preferences.

This conclusion is also in line with the results from our neighborhood choice model, which builds on Dahl (2002) to retrieve selection-corrected estimates of the WTP for school quality, which are close to the uncorrected estimates.

A.4 School Districts and Large Clusters

This appendix provides a visual depiction to fix ideas on the variation used to recover estimates of the WTP for school quality and neighborhood attributes. Our main specification uses variation in housing prices and school quality within large clusters. Variation in school quality arises from school districts' boundaries crossing within different large clusters. Figure A.3 depicts, schematically, a large cluster in which four small clusters are contained. The two shades of blue denote two different school districts which are bounded by the dashed line.

Figure A.3: School Districts and Large Clusters



Notes: This figure depicts one large cluster, with boundaries of the four small clusters represented with solid lines. The dashed line represents the boundary of two school districts, which are illustrated in two different shades of blue. The variation in school quality we utilize, arises from school district boundaries crossing within large clusters.

A.5 Expanding Cluster Assignment to Other Years

Since a given address always holds the same cluster ID, we were able to expand the cluster IDs for all individuals (including children), beyond 2004 for addresses that existed before 2004. We do so for 2005 and 2006, achieving a matching rate of 94.4% and 92.5% respectively. However, unique identifiers of addresses changed beyond 2006, making it more difficult to expand cluster IDs further. We exploit data on a number of housing attributes[35] to conduct probabilistic matching on housing units to recover a unique mapping between housing identifiers across 2006 and 2007. This mapping then allows us to expand our dataset of clusters further from 2007 to 2015, which we use in Sections 7.1 and 8.

Given the above challenges, our main estimating sample uses data from 2002 to 2006. In Section 7.1, we run our main model separately for each year between 2002 and 2015 and get similar results of the WTP for school quality. We therefore use this extended sample in Section 8 when controlling for potential bias arising from selection into neighborhoods.

The construction of these clusters ensures that their size remains small, such that we are able to control for unobserved neighborhood-level attributes. However, the requirement that each cluster comprises of a minimum number of individuals (150 for small and 600 for large clusters), implies that in more rural areas the size of these clusters can remain relatively large. This has two impacts on our empirical strategy. First, the presence of larger clusters reduces our ability to control for unobserved neighborhood level characteristics. Second, we are less likely to have different school district within each cluster. To avoid such issues, we compute the density (inhabitants per square kilometers) of parishes in which clusters lie as a proxy for clusters' size. We then remove from our sample all observations which lie within the bottom 25th percentile of parishes in terms of density.[36] This reduces our sample by less than 1%, while allowing us to focus on denser areas, where our cluster fixed effects strategy is likely to capture more of the unobserved attributes of the space common to all houses.

A.6 Spatial Decomposition of Inequality

Figure A.4 presents the Theil's T decomposition of building characteristics across different neighborhood units in Denmark, i.e., municipality, parish, large cluster, and small cluster levels.[37] Panel (a) focuses on the number of floors, Panel (b) shows the statistics for the

[35] These attributes include the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area, the age of the property as well as parish code.

[36] An alternative strategy would be to remove clusters based on their geographical size. However, as we do not observe the size of clusters in our data, we base our measure on the size of the parishes in which they lie.

[37] Consider the population of Danish households, $i = 1, ..., n$, with income y_i, and weight w_i (we assume $w_i = 1$ in the decomposition analysis in this paper). Let $f_i = w_i/N$, where $N = \sum w_i$. When the data are unweighted, $w_i = 1$ and $N = n$. Arithmetic mean income is m. Now, consider the exhaustive partition of the population into mutually-exclusive neighborhoods $k = 1, ..., K$.

The Generalized Entropy class of inequality indices is given by:

$$GE(a) = \frac{1}{a(a-1)}[\sum f_i(\frac{y_i}{m}^a) - 1], a \neq 0\&a \neq 1,$$

$$T_t = GE(1) = \sum f_i \frac{y_i}{m} \log \frac{y_i}{m},$$

number of apartments, Panel (c) uses the age of the building, and Panel (d) considers the living area.

$$GE(0) = \sum f_i \log \frac{y_i}{m},$$

$GE(a)$ index can be additively decomposed as

$$GE(a) = GE_W(a) + GE_B(a)$$

where $GE_W(a)$ is Within-group Inequality and $GE_B(a)$ is Between-Group Inequality and

$$GE_W(a) = \sum v_k^{1-a} s_k^a GE_k(a)$$

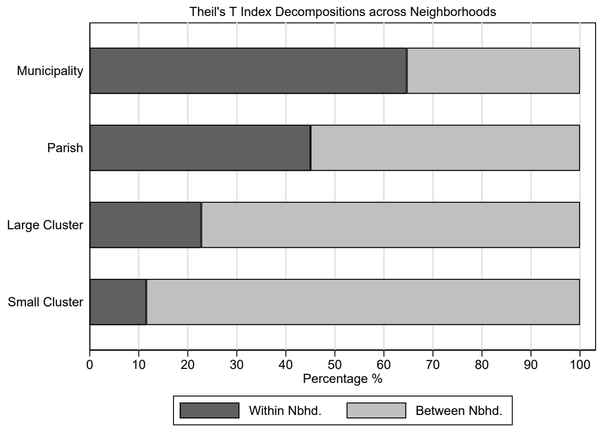
Which, for Theil's T index is as follows:

$$GE_W(1) = \sum s_k GE_k(1)$$

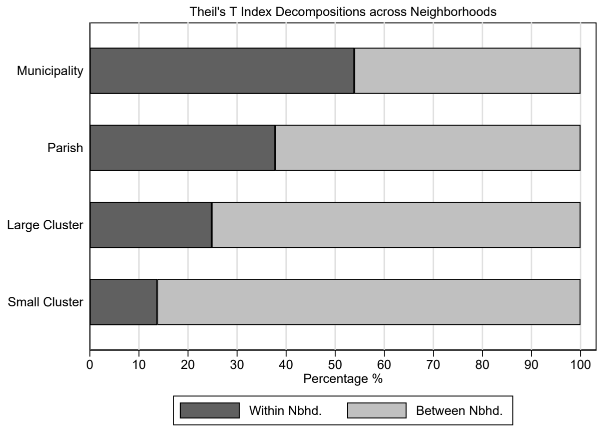
where $v_k = \frac{N_k}{N}$ is the number of persons in subgroup k divided by the total number of persons (subgroup population share), and s_k is the share of total income held by k's members (subgroup income share), i.e., v_k is the sum of the weights in subgroup k divided by the sum of the weights for the full estimation sample).

$GE_k(a)$, inequality for subgroup k, is calculated as if the subgroup were a separate population, and $GE_B(a)$ is derived assuming every person within a given subgroup k received k's mean income, m_k.

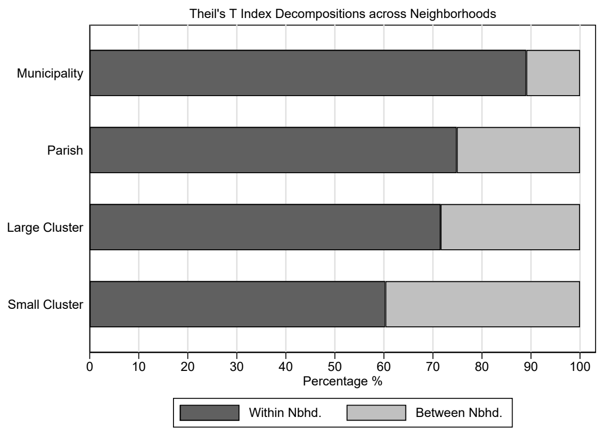
Figure A.4: Variance Decomposition of Housing Characteristics across Neighborhoods



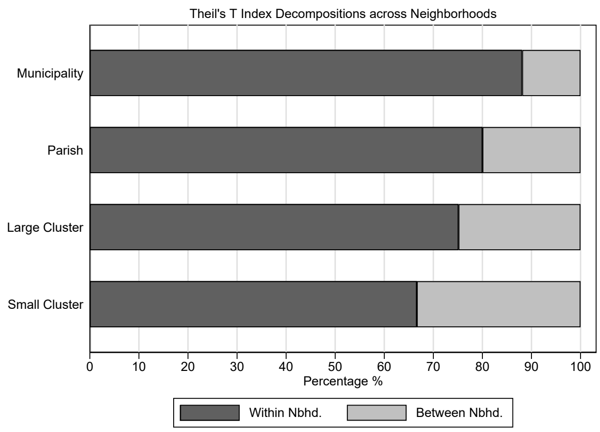
(a) Number of Apartments



(b) Number of Floors



(c) Age of the Building



(d) Living Area

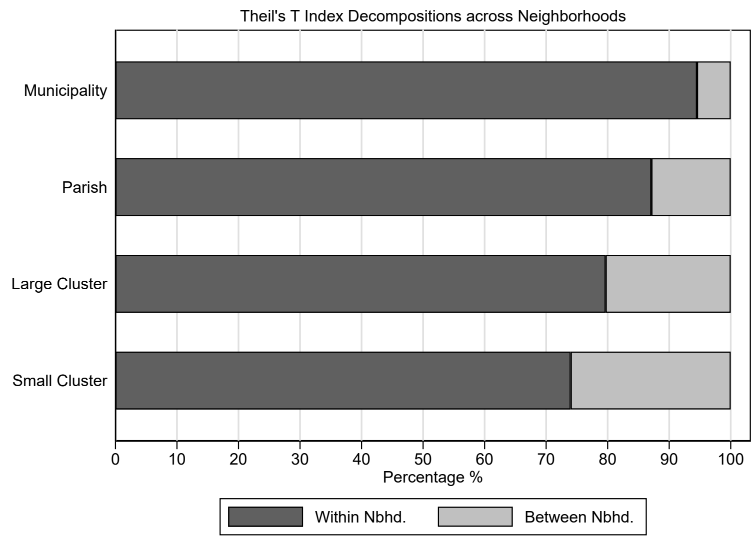
Notes: This figure presents the Theil's T decomposition of building characteristics across different neighborhood units in Denmark. Panel (a) focuses on the number of floors, Panel (b) shows the statistics for the number of apartments, Panel (c) uses the age of the building, and Panel (d) considers the living area. See Appendix A.6 for details.

A.6.1 Decomposition of Income Inequality

We also analyze the spatial decomposition of income inequality across neighborhoods in Denmark using the Theil's T Index. Figure A.4 shows the decomposition across neighborhoods by different units of neighborhood, i.e., municipality, parish, large cluster, and small cluster levels. Results suggest that while at municipality-level only about 5% of the income inequality can be contributed to between-neighborhood component, the share of between-neighborhood component increases by a factor of 5 (to more than 26%) when we analyze the income inequality across small clusters. We also use the Gini Index measure to analyze the the income inequality at different neighborhood units in Denmark. Figure A.6 plots the distribution of neighborhood Gini index for various neighborhood units, i.e., municipality, parish, large cluster, and small cluster. Figure A.6 shows that the income inequality is dramatically lower at cluster levels, suggesting more homogeneity among individuals living close to each other in a our neighborhood unit.

Figure A.5 presents the Theil's T decomposition of family income across different neighborhood units in Denmark, i.e., municipality, parish, large cluster, and small cluster levels.

Figure A.5: Variance Decomposition across Neighborhoods



Note: This figure presents the Theil's T decomposition of family income across different neighborhood units in Denmark. The disposable income is used as the measure of income. Family income is averaged over 2010-2015. See Appendix A.6 for details.

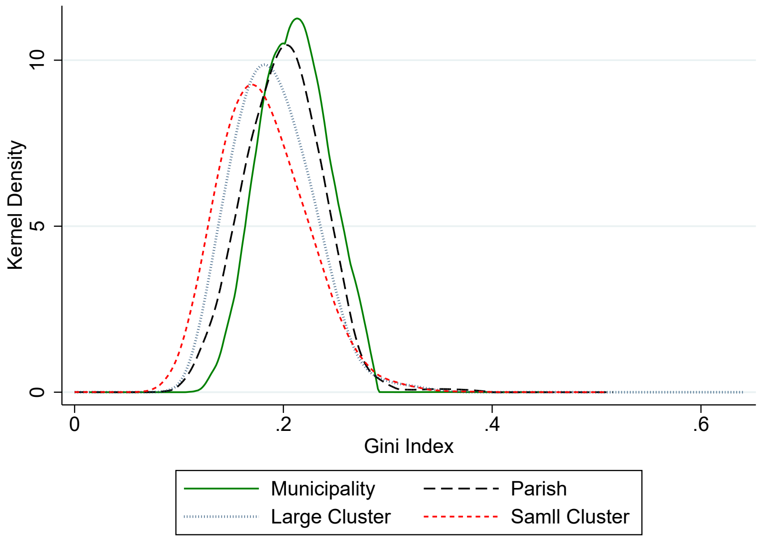
Overall, our results suggest that families who reside in the same cluster are much more alike in terms of their observable characteristics, compared to the pool of families who live in the same parish or municipality. The results presented here focuses on income measures, but we observe a similar pattern when we look at other characteristics such as education level.

Next, We also analyze the income inequality using the Gini coefficient and segregation intensity over the income distribution by neighborhood unit.

A.6.2 Gini Index

We use the Gini Index measure to analyze the the income inequality at different neighborhood units in Denmark. Figure A.6 plots the distribution of neighborhood Gini index for various neighborhood units, i.e., municipality, parish, large cluster, and small cluster, which shows that the income inequality is dramatically lower at cluster levels, suggesting more homogeneity among individuals living close to each other in a our neighborhood unit.

Figure A.6: Gini Index Kernel Density by Neighborhood Unit



Note: Kernel density of neighborhood Gini coefficient of household income inequality in Denmark, 2015. The disposable income is used as the measure of family income. Family income is averaged over 2010-2015. The right axis (dash-dot line) indicates the ratio of segregation using parishes to that using clusters.

A.6.3 Segregation Index across Neighborhoods

Measure of segregation in income in neighborhoods: Reardon & Bischoff (2011), can be used to form a scale from 0-1, where zero indicates no income segregation; i.e., all income percentiles equally represented in all neighborhood, and one suggests perfect segregation; i.e., each neighborhood consists of families from same part of income distribution.

Reardon *et al.* (2006) describe the rank-order information theory index in detail. First,let p denote income percentile ranks (scaled to range from zero to one) in a given income distribution (i.e., $p = F(Y)$) where Y measures income and F is the cumulative income density function). Now,for any given value of p, we can dichotomize the income distribution at p and compute the residential (pairwise) segregation between those with income ranks less than p and those with income ranks greater than or equal to p. Let $H(p)$ denote the value of the traditional information theory index (James & Taeuber, 1985; Theil & Finizza, 1992; Theil, 1972; Zoloth, 1976, see) of segregation computed between the two groups so defined. Likewise, let $E(p)$ denote the entropy of the population when divided into these two groups (Theil & Finizza (1992); Theil (1972); Pielou (1977)). That is,

$$E(p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$$

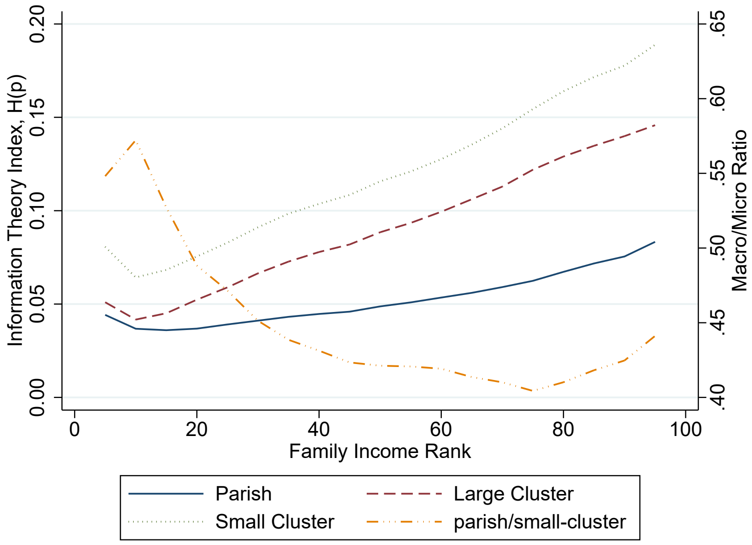
$$H(p) = 1 - \sum_j \frac{t_j E_j(p)}{TE(p)}$$

where T is the population of the metropolitan area and tj is the population of neighborhood j, and p is the income percentile ranks from 0-1, i.e., $F(Y)$, where Y is the total income from all resources.

Figure A.7 shows family income segregation for the municipality of Copenhagen in 2015. Results indicate that across the whole distribution of income, the segregation is more intensive at small neighborhood units (i.e., cluster level) compared to larger units (i.e., parish). In addition, Figure A.7 shows the macro/micro segregation ratio, which measures the proportion of micro-scale segregation (segregation among small cluster environments) that is due to macro-scale segregation patterns (segregation among parish environments). This ratio can be interpreted as a measure of the geographic scale of segregation, with larger values indicating that more of the measured segregation is due to the separation of groups over large distances (see Lee *et al.*, 2008; Reardon & Bischoff, 2011; Reardon *et al.*, 2009). Results suggest that around 50% (varying between $40\% - 60\%$) of the small cluster level segregation can be attributed to parish-level segregation patterns.

Figure A.7 shows family income segregation for the municipality of Copenhagen in 2015. The figure indicates estimated between-parish (cluster) segregation (as measured by the information theory index, H) between families with incomes above and at or below each percentile of the municipality-wide family income distribution. In addition, Figure A.7 shows the macro/micro segregation ratio (dash-dot line, with scale on the right-hand axis), which measures the proportion of micro-scale segregation (segregation among small cluster local environments) that is due to macro-scale segregation patterns (segregation among parish environments). This ratio can be interpreted as a measure of the geographic scale of segregation, with larger values indicating that more of the measured segregation is due to the sep-

Figure A.7: Segregation Index across Neighborhoods



Note: Family income segregation, municipality of Copenhagen, 2015, by income percentile and neighborhood units. The figure indicates estimated between-parish/cluster segregation (as measured by the information theory index, H) between families with incomes above and at or below each percentile of the municipality-wide family income distribution. The disposable income is used as the measure of family income. Family income is averaged over 2010-2015 The right axis (dash-dot line) indicates the ratio of segregation using parishes to that using clusters.

aration of groups over large distances (see Lee *et al.*, 2008; Reardon & Bischoff, 2011; Reardon *et al.*, 2009). Figure A.7 suggests that across the whole distribution of income, the segregation is more intensive at small neighborhood unite (i.e., cluster level).

A.7 Sorting Within Clusters

A.7.1 Property Price

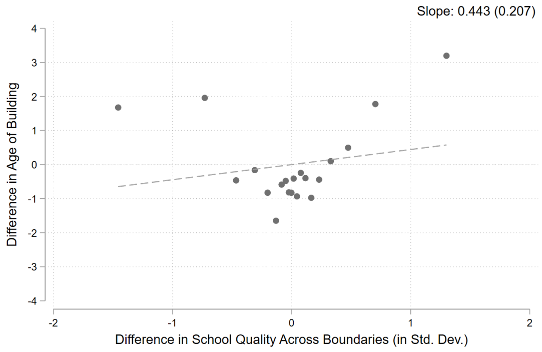
Figure A.8: Sorting Within Large Clusters - Property Price

(a) Property Price

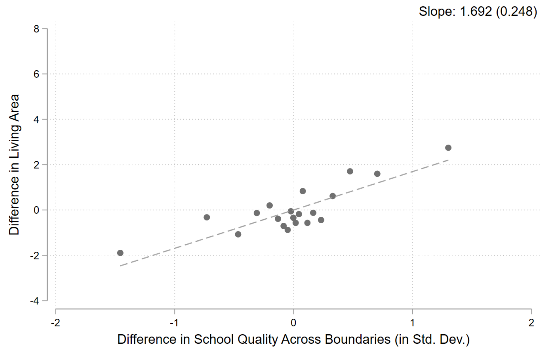
Note: Relationship between the difference in property prices and the difference in school quality within large clusters. This is constructed by regressing property prices on school quality, controlling for neighborhood-by-cohort fixed effects and hyper-local neighborhood attributes—average income, years of education, fraction married, non-westerners, foreigners, non-intact households, private schools and average neighborhood school quality. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

A.7.2 Housing Characteristics

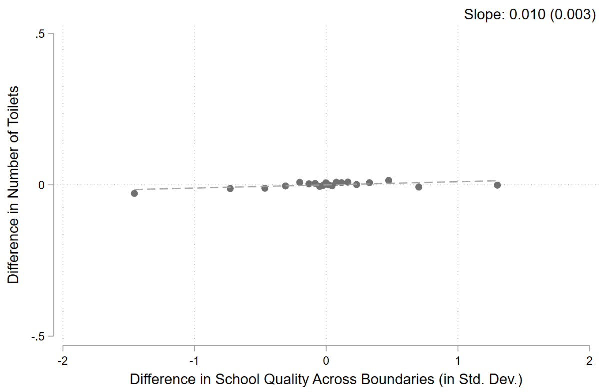
Figure A.9: Sorting Within Large Clusters - Housing Characteristics



(a) Building Age



(b) Living Area

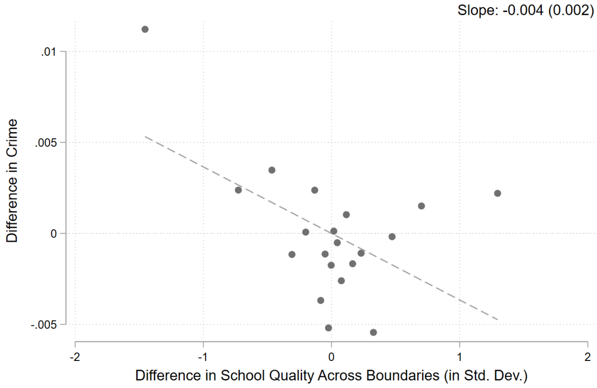


(c) Number of Toilets

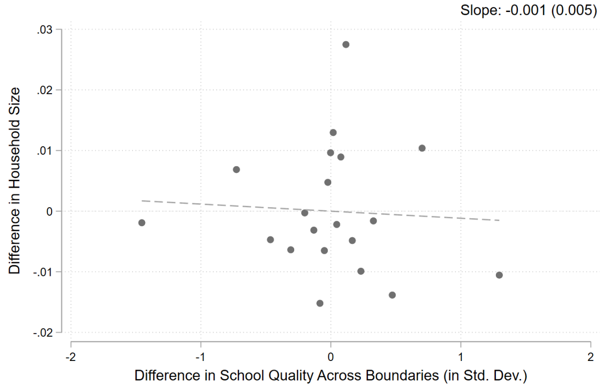
Note: Relationship between the difference in housing characteristics and the difference in school quality within large clusters. This panel is constructed by regressing various housing attributes on school quality, controlling for neighborhood-by-cohort fixed effects and hyper-local neighborhood attributes—average income, years of education, fraction married, non-westerners, foreigners, non-intact households, private schools and average neighborhood school quality. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

A.7.3 Individual Characteristics

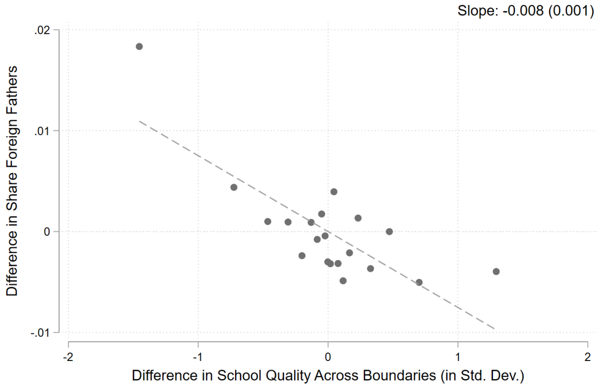
Figure A.10: SORTING WITHIN LARGE CLUSTERS - INDIVIDUAL CHARACTERISTICS



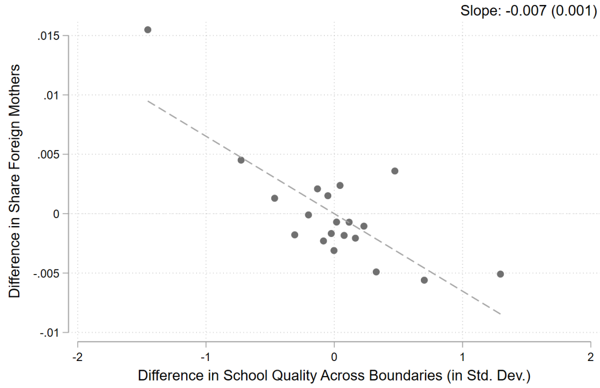
(a) Crime



(b) Household Size



(c) Distance to Work

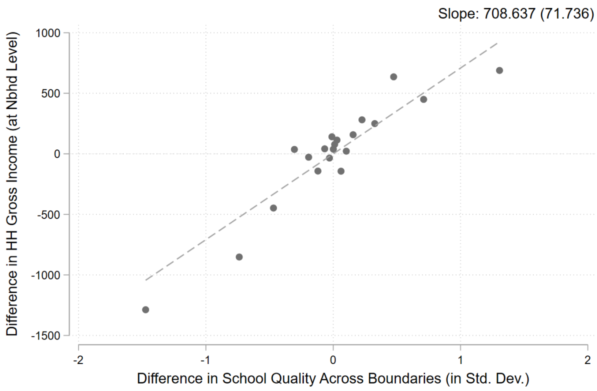


(d) Foreign Father

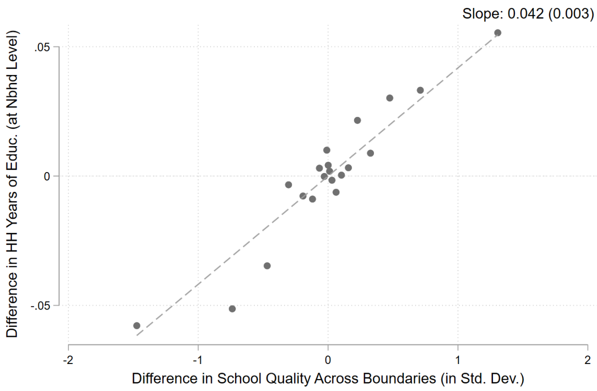
Note: Relationship between the difference in individual characteristics and the difference in school quality within large clusters. Each panel is constructed by regressing various individual characteristics on school quality, controlling for neighborhood-by-cohort fixed effects, hyper-local neighborhood as well as housing attributes—the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

A.7.4 Neighborhood Characteristics

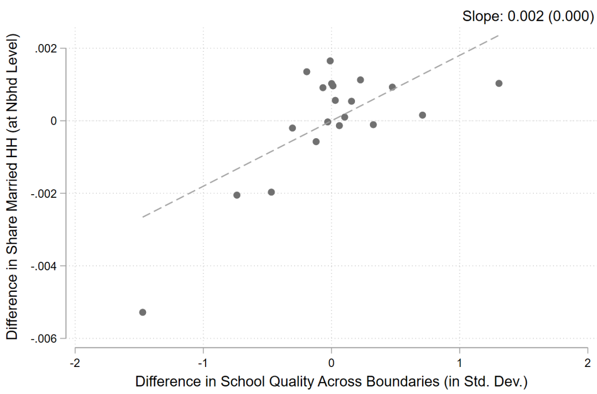
Figure A.11: Sorting Within Large Clusters - Neighborhood Characteristics



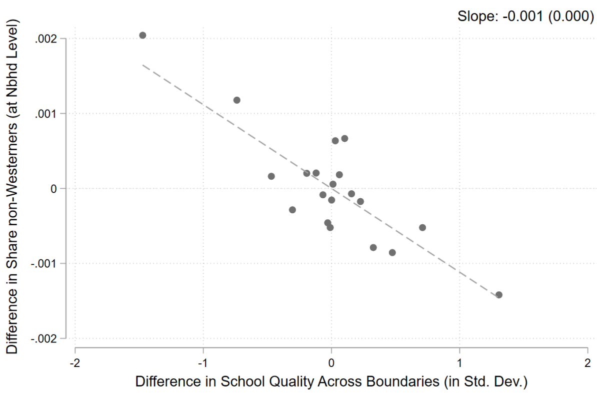
(a) Neighborhood Average Income



(b) Neighborhood Average Education



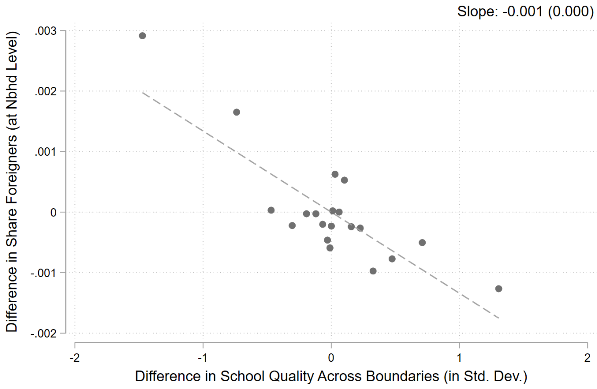
(c) Neighborhood Share of Married Households



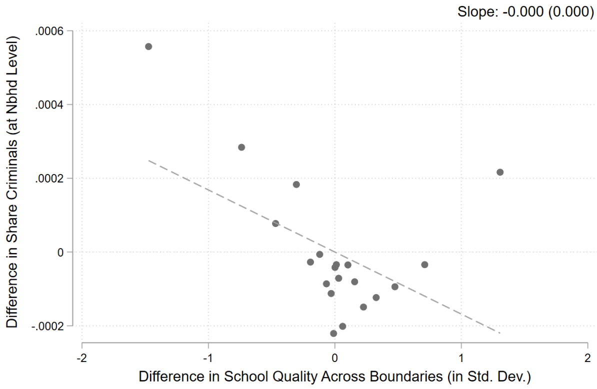
(d) Neighborhood Share of Non-western Foreigners

Note: Relationship between the difference in neighborhood characteristics and the difference in school quality within large clusters. Each panel is constructed by regressing various small cluster characteristics on school quality, controlling for neighborhood-by-cohort fixed effects and housing attributes—the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

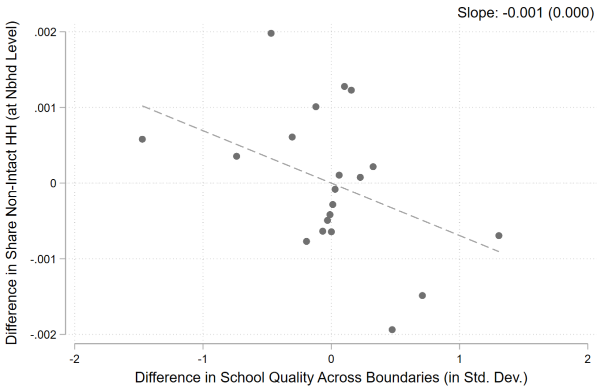
Figure A.12: Sorting Within Large Clusters - Neighborhood Characteristics



(a) Neighborhood Share of Foreigners



(b) Neighborhood Share of Criminals



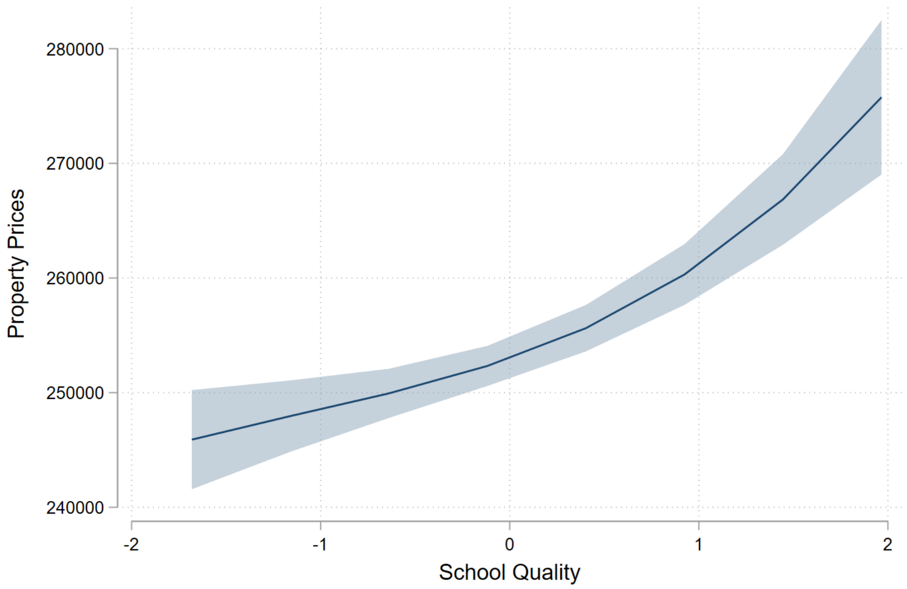
(c) Neighborhood Share of Non-intact Households

Note: Relationship between the difference in neighborhood characteristics and the difference in school quality within large clusters. Each panel is constructed by regressing various small cluster characteristics on school quality, controlling for neighborhood-by-cohort fixed effects and housing attributes—the type of building, the number of floors, the number of units per building, the number of bedrooms, toilets and bathrooms, the size of the living area and the age of the property. Standard errors corrected for clustering at the large cluster-cohort level are reported in the top right corner.

A.8 Local Linear Hedonic Price Function

To explore the heterogeneity of the WTP over the distribution of school quality, this Appendix provides local linear estimates of the WTP, controlling for the same family and housing characteristics as well as fixed effects at the large cluster-by-cohort level. Specifically, Figure A.13 depicts the nonlinear relationship between house prices and school quality—the hedonic price function. Its slope reflects the WTP for school quality.

Figure A.13: Local Linear Estimates of the WTP



Notes: The figure shows a local linear fit of the relationship between property prices and school quality, after controlling for neighborhood and housing characteristics, as well as large cluster-by-cohort fixed effects. School quality is measured by taking the average test scores of students attending a given school. The sample contains property values for parents who have a child attending the 9th school grade between 2002 and 2006.

A.9 Details on the Selection Model

We consider the following selection model.[38] An individual i makes a choice, j, amongst M different neighborhood alternatives, i.e., the large clusters we considered as fixed effects in the previous specifications. Assume i chooses $j = 1$, then in this case we observe individual i's property value only for neighborhood $j = 1$.

The hedonic price regression written for individual i, is given by

$$y_{i1} = \alpha_1 + x_i'\delta_1 + s_i\beta_1 + u_{i1}$$

where y_{i1} is log house prices in chosen cluster 1, α_1 is a cluster level specific constant, x_i is a vector of neighborhood characteristics (including housing characteristics), s_i measures the school quality and u_{i1} is an error term. Individuals' utilities, V_{ij}, are specified as follows:

$$V_{ij} = z_i'\gamma_j + \eta_{ij}, \qquad j = 1, \ldots, M$$

where the disturbance u_{ij} is not parametrically specified and verifies $E[u_{ij} \mid x_i, z_i] = 0$ and $V[u_{ij} \mid x_i, z_i] = \sigma^2$. Moreover, j is a categorical variable that describes the choice of an economic agent among M large cluster alternatives based on utilities V_{ij}.

We assume that the model is non-parametrically identified from exclusion of some of the variables in z from the variables in x; letting z_i be comprised of a set of dummies capturing whether each child lives in the same parish as their grandparents. The exclusion restriction for the identification purpose is that an individual's house price does not depend on whether their parents live in the same neighborhood. This specification allows individuals to have preferences for living close to their parents as well as the neighborhood characteristics, but we restrict mean house price to be a function only of neighborhood characteristics in which the house is located. We further add a host of household-level characteristics, such as income, education, marital status and distance to work.

Without loss of generality, the outcome variable y_{i1} is observed if and only if large cluster 1 is chosen, which happens when:

$$V_{i1} > \max_{j \neq 1}\{V_{ij}\}$$

Now, define:

$$\varepsilon_{i1} = \max_{j \neq 1}\{V_{ij} - V_{i1}\} \tag{11}$$
$$= \max_{j \neq 1}\{z_i'\gamma_j + \eta_{ij} - z_i'\gamma_1 - \eta_{i1}\} \tag{12}$$

which is equivalent to $\varepsilon_{i1} < 0$

Assume that the η_{ij}'s are independent and identically Gumbel distributed. As shown by McFadden (1973), this specification leads to the multinomial logit model with

$$P(\varepsilon_1 < 0 \mid z) = \frac{\exp(z'\gamma_1)}{\sum_j \exp(z'\gamma_j)}$$

[38] The technical exposition is based on Bourguignon *et al.* (2007).

Based on this expression, consistent maximum likelihood estimates of the γ_j's can be easily obtained.

The problem is to estimate the parameter vector β_1 while taking into account that the disturbance term u_{i1} may not be independent of all η_{ij}'s. This would introduce some correlation between the explanatory variables and the disturbance term in the hedonic price regression. Because of this, least squares estimates of β_1 would not be consistent.

Define Γ as follows:

$$\Gamma = \{z'\gamma_1, z'\gamma_2, \ldots, z'\gamma_M\}$$

Generalizing the model from Heckman (1979), bias correction can be based on the conditional mean of u_1 :

$$E\left(u_1 \mid \varepsilon_1 < 0, \Gamma\right) = \iint_{-\infty}^{0} \frac{u_1 f\left(u_1, \varepsilon_1 \mid \Gamma\right)}{P\left(\varepsilon_1 < 0 \mid \Gamma\right)} d\varepsilon_1 du_1 = \lambda(\Gamma)$$

where $f\left(u_1, \varepsilon_1 \mid \Gamma\right)$ is the conditional joint density of u_1 and ε_1. For notational simplicity, call P_k the probability that any neighborhood k is preferred:

$$P_k = \frac{\exp\left(z'\gamma_k\right)}{\sum_j \exp\left(z'\gamma_j\right)}$$

Given that the relation between the M components of Γ and the M corresponding probabilities is invertible, there is a unique function μ that can be substituted for λ such that:

$$E\left(u_1 \mid \varepsilon_1 < 0, \Gamma\right) = \mu\left(P_1, \ldots, P_M\right)$$

Therefore, consistent estimation of β_1 can be based on either regression:

$$\begin{aligned} y_{i1} &= \alpha_1 + x_i'\delta_1 + s_i\beta_1 + \mu\left(P_1, \ldots, P_M\right) + w_{i1} \\ &= \alpha_1 + x_i'\delta_1 + s_i\beta_1 + \lambda(\Gamma) + w_{i1} \end{aligned}$$

where w_1 is a residual that is mean-independent of the regressors.

As argued by Dahl (2002), semi-parametric estimation of this model would have to face the curse of dimensionality. Whenever the number of alternatives is large it implies the estimation of a large number of parameters, rapidly making it intractable for practical implementation. Thus, restrictions over $\mu(P_1, \ldots, P_M)$, or equivalently $\lambda(\Gamma)$, are required.

Dahl (2002) makes the following assumption:

A1 : Dahl's index sufficiency assumption

$$f\left(u_1, \varepsilon_1 \mid \Gamma\right) = f\left(u_1, \varepsilon_1 \mid P_{i,i=1\ldots M-1}\right) = f\left(u_1, \varepsilon_1 \mid P_{i,i\in S}\right), S \subset \{1 \ldots M-1\}$$

This means that consistent estimation of β_1 can be based on

$$y_1 = \alpha_1 + x_i'\delta_1 + s_i\beta_1 + \mu(P_{i,i\in S}) + w_{i1}$$

Based on this model, we present estimates of the WTP for school quality controlling for neighborhood selection, in Section 8.1.

A.10 Structural Interpretation of Reduced-form Estimates

In this Appendix, we formalize how the reduced-form estimates of the impact of school quality on later life outcomes should be interpreted through the lens of a dynamic model of the school-quality-achievement relationship. We draw extensively on Todd & Wolpin (2003) in the below description of the model.

Conceptual Framework Consider a three-period model, where $t = 0$ corresponds to the time interval prior to the age the child enters school, while $t = 1$ and $t = 2$, respectively denote the first and second year of school. Moreover, let A_1 denote the child's achievement level prior to entering the first year of school while F_0 represents family inputs into the education production during $t = 0$. Finally, let μ be a measure of the child's ability, determined at birth.

Achievement at the time of school entry depends only on family inputs and ability:

$$A_1 = g_0(F_0, \mu)$$

Moreover, family inputs in the preschool period are assumed to be determined by the family's permanent resources, W, and the child's endowment μ.

Achievement at the start of the second year of school depends on the entire history of family inputs (F_0 and F_1) and school inputs S_1 as well as endowments:

$$A_2 = g_1(S_1, F_1, F_0, \mu)$$

Along with the above technology for combining inputs to create outcomes, the level of input is determined by decision rules from the parents and the schools. Let S_1 denote the actual school-input level that is relevant to the child and $\bar{S}_1$ the average school-level investment. This distinction can arise due to parents having incomplete information, at the time of the neighborhood choice decision, about the actual level of school-input. The decision rule of the family and the school inputs associated with their neighborhood decisions are given by:

$$F_1 = \phi(A_1, W, \mu, S_1 - \bar{S}_1)$$

$$\bar{S}_1 = \theta(A_1, W, \mu)$$

We assume here that the school chooses input levels purposefully, accounting the child's achievement level and the endowment. School's input decision rule is $S_1 = \psi(A_1, \mu)$, where family resources do not play a role.

At the beginning of the second year, the family makes a new decision about neighborhood location, which depends on the child's achievement, resources of the family and child's ability at birth.

In this simple model, the total effect of an increase in the school input in the first year, S_1, on achievement in the second year A_2 is given by:

$$\frac{dA_2}{d(S_1 - \bar{S_1})} = \frac{dA_2}{dS_1}$$
$$= \frac{\partial g_1}{\partial S_1} + \frac{\partial g_1}{\partial F_1}\frac{\partial F_1}{\partial (S_1 - \bar{S_1})}$$

Thus, to disentangle the different channels, one requires knowledge of the family input decision rule. Without such knowledge, one can learn about the total effect of an exogenous change in school quality on achievement, without holding other inputs constants.

Statistical Model Building on the above model, we consider the below regression analog of the true technology:

$$\begin{aligned} Y_{il} \;\; = \;\; & F_{ila}\gamma_1 + F_{ila-1}\gamma_2 + \ldots + F_{il1}\gamma_a \\ & + S_{il}\beta \\ & + \mu^f_{il0} + \mu^c_{il0} + \varepsilon_{ila} \end{aligned}$$

where F_{ila} denotes the vector of parent-supplied inputs for individual i in household l at age a and S_{il} denotes 9th grade school-supplied input. Y_{il} denotes later life achievements. Further, let a child's endowed mental capacity be denoted as μ_{il0}. The superscript c and f, respectively correspond to a sibling-specific and family-specific endowment. Measurement error in test scores is denoted by ε_{ila}. For simplicity of notation, we omit the school and neighborhood subscripts present in the main text.

Differencing the above equation, we get:

$$Y_{il} - Y_{i'l} = [S_{il} - S_{i'l}]\,\beta + \{[\mu^c_{il0} - \mu^c_{i'l0}] + \varepsilon_{ila} - \varepsilon_{i'la}\}$$

Under the above specification with sibling fixed effects, where input effects are constant across ages, the parameter β is identified under the following assumptions:

1. Inputs are uncorrelated with sibling error terms (in other words, parents do not change their input decisions based on innovations on sibling outcomes) and prior own outcome;

2. Any omitted input, varying across sibling, is orthogonal to included inputs;

3. Input choices may depend on family-specific endowment but not child-specific endowment;

4. 9th grade school quality is a sufficient statistic for the vector of school-supplied inputs over age.

In further work, we aim to empirically test assumptions (3) and (4). Assumption (3) can be tested by adding some proxies of latent measures of ability, such as birth weight. Assumption (4) can be tested by adding a vector of school-supplied inputs over age.