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

Sadegh Eshaghnia James J. Heckman Goya Razavi

November 19, 2021

Abstract

Education in Denmark is available for free to all. Yet, despite equalized school expenditures, we observe substantial differences in school quality across areas due to sorting of individuals across neighborhoods. This paper aims to evaluate the willingness to pay for school quality and neighborhood socio-demographics and proposes a new methodology to do so. We draw on contiguous clusters, comprising of 250 households (nested within 1000-households clusters), which we show are relatively homogeneous neighborhoods. Using within-clusters variation allows us reduce the underlying issue of sorting that often plagues estimates of the willingness to pay in the literature. We find a willingness to pay of 3% for houses associated with a school whose average grade is one standard deviation above the mean - which is robust to various specifications. Using a similar strategy, we estimate the valuation of neighborhood attributes. Finally, using our rich longitudinal data, we reveal that better schools can have a significant impact on later life outcomes, increasing income at age 30 by around 5-10%.

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

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

1 Introduction

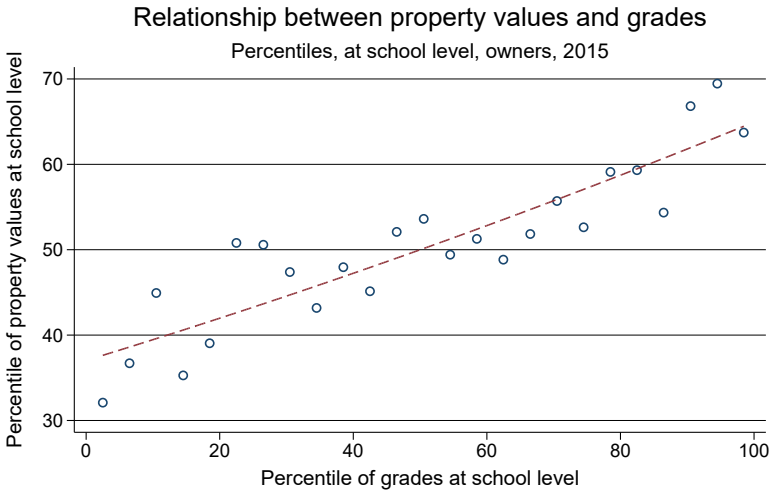
Estimation of household preferences for neighborhood attributes is a challenge due to endogeneity issues that arise from household sorting across neighborhoods. Still, this undertaking has great significance in better understanding the drivers behind the allocation of individuals to neighborhoods and the resulting later-life consequences of such allocations.

This paper aims to provide a novel empirical strategy to estimate 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 to reduce inequalities and equalize opportunities, inter alia, by providing an 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 (Gensowski *et al.* (2020); Heckman & Landersø, 2021). Residence-based assignments can further affect residential segregation and social mobility - issues that have been addressed in the theoretical literature (see for instance Durlauf, 1996; Benabou, 1993; Bénabou, 1996). By estimating the WTP for school quality, this paper provides a potentially important mechanism behind the presence of notable inequities in Denmark (Landersø & Heckman, 2017), despite the strong egalitarian welfare system.

The challenge in estimating the WTP for schools is exposed in Tiebout (1956) seminal work. Households sort based on their heterogeneous preferences for a vector of neighborhood characteristics. This leads to correlate neighborhood and individual characteristics and as a result neighborhood attributes themselves. Any unobserved neighborhood attribute, if not properly accounted for, would thus result in biased estimation of individuals' valuation of neighborhood amenities.

Figure 1 shows the positive relationship between percentiles of school quality (as measured by average school grades) on the percentiles of property values of the biological parents in 2015, averaged at the school level. This figure reveals the possibility that wealthier parents' indeed sort into neighborhoods where school quality is higher. Of course, sorting

into neighborhood would correlate school quality and other neighborhood amenities. Our goal here is to disentangle the role played by school quality on house prices, above and beyond other neighborhood attributes.



Note: The figure shows a binned scatterplot, with a quadratic 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 grades of students attending a given school in 2015. The sample contains property values for parents who have a child attending 9th grade in 2015.

Figure 1: Relationship between percentiles of property values and school quality at school level, 2015 (for owners)

To deal with the sorting of households across neighborhoods, our strategy capitalizes on variation in house prices and school quality within very small neighborhoods (median size of 0.3 square mile) arising from discontinuous school catchment area within these clusters. Given the small geographic area that are spanned by these clusters, we are able to hold constant unobserved neighborhood amenities. In fact, within such small neighborhoods, access to public goods or other amenities would not vary, allowing us to identify the effect of changes in school quality on house prices.

To motivate our general approach, we begin with a descriptive analysis of the variation within and across different neighborhood levels in Denmark. We present evidence that fo-

cusing on smaller geographic units allows us to get more homogeneous clusters of individuals. We then show, how adding these geographic fixed effects decreases considerably 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 clearly illustrated by the fact that adding these fixed effects flips the sign of some coefficients from our standard OLS specification.

Of course, households would still sort on either side of the boundaries to get access to their preferred school, as documented 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 miles neighborhoods we consider, for instance, the fraction of migrants in the surrounding blocks.[1] Thanks to our rich data, we are able to control for hyper local neighborhood sociodemographics, which not only allows us to recover household valuation of school quality, but also of those other hyper local neighborhood attributes which we add in our model. 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.

We make use of 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 of housing units, we are able to observe a number of attributes including living area, number of bedrooms, building type and age, as well as number of floors. This further 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 this administrative data, building on students grades and teacher employment and academic records, allowing us to tease out the different features of school

[1] Put differently, the presence of sorting across school catchment areas (within these small clusters) would not necessarily bias our estimate of the mean marginal WTP, since the neighborhood quality would not vary within such small areas (except for school peers, which would be encapsulated in our estimate of school quality).

quality contributing to its capitalization into house prices.

Our strategy needs to satisfy two key identifying assumptions. First, unobserved housing and neighborhood characteristics do not vary systematically within the geographic clusters we use as fixed effects. Second, the measures of immediate neighborhood income, education and origins included in the regression control fully capture everything that is relevant about the attributes of one's neighbors.

Our main results reveals that households are willing to pay around 3% of house prices for a one standard deviation increase in school quality, measured by average school grades. This is in line with estimates found in other countries with greater income and wealth inequalities than Denmark, including Australia, France, the UK and the US (see Black & Machin, 2011 for a review). We also find that households positively value neighborhoods attributes such as being in a neighborhood with a larger fraction of more educated individuals and with higher average income.

Results on the willingness to pay for school quality are robust to various specifications. In particular we incorporate our model within the framework of a polychotomous neighborhood choice model. Estimates from this model tend to show that our results are not significantly biased by households self-selecting into neighborhoods.

Building on this framework, we devise an empirical model of the impact of school quality on later life outcomes. We show that students' attending better schools do indeed have a 10% higher probability of graduating college, conditional on neighborhood quality and parental characteristics. 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 outcome.

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, even in a country that has been touted for its egalitarian policies.

In fact, the combination of sorting together with the residence-based assignment rules of households to schools has a significant impact on later life outcomes, increasing chances of graduating college by around 10% and income by between 5% to 10%, for a one standard deviation increase in school quality attended.

2 Danish Context and Data

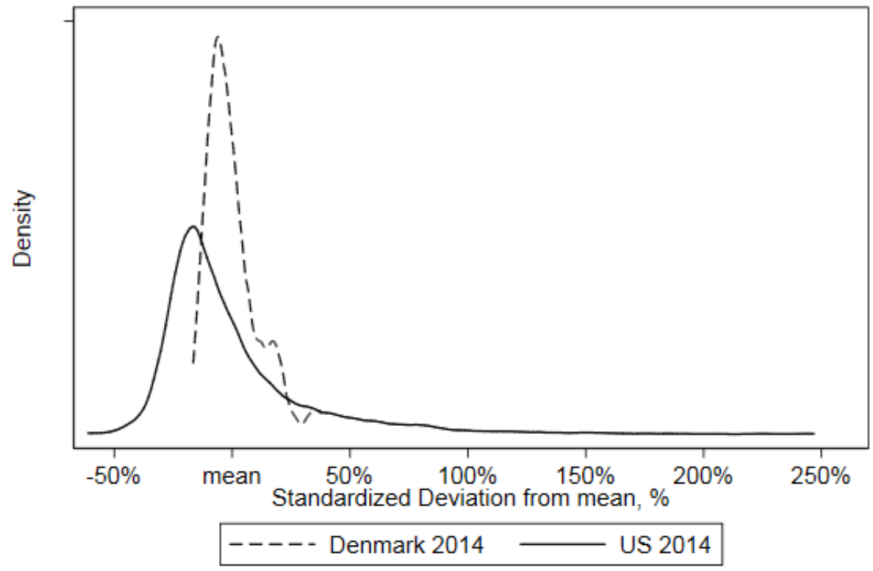
2.1 Public Schooling System in Denmark

The Danish schooling system is based upon the principle of schooling for all, which is provided at no charge in public schools between the ages of 6 to 16.

Danish municipalities' primary revenue in Denmark (about 70%) comes from local taxes, which vary only minimally across municipalities. Municipalities can decide on how to distribute their expenditures between task areas (such as public primary and lower secondary schools, employment agencies, elderly care, etc.), providing they meet the statutory requirements for each of them. Even though municipalities have some discretion in deciding the municipal tax rates, they need to take the financial policy of the central government into account. These are formalized through financial agreements and constrain the level of taxes set by municipalities. These agreements set out the level of aggregate expenditure, tax rates for local governments as well as the block grants from the central government. In particular, this constrains the amount of variation in per pupil expenditure through redistribution across municipalities, as can be seen in Figure 2. Another way to state this is that teaching resources are equalized and in particular teacher pay as well. In fact, the literature presents evidence pointing to this by showcasing the lack of relationship between observable characteristics of teachers and their pay (see Gensowski *et al.* (2020)).

Despite an equalized school expenditure and teacher salary distribution, teachers still do sort based on the potential quality of the students they may be teaching, as exemplified by Gensowski *et al.* (2020). 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.



Note: The figure shows average per pupil school expenditures in public schools in 2014 relative the to the country average. Source: Gensowski *et al.* (2020)

Figure 2: Comparing school expenditure between Denmark and US

This paper relies on administrative data from Statistics Denmark, which provides data covering the whole population. Below we describe in more details the different key variables used in the analysis that follows.

2.2 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.

Various studies in the hedonic analysis tradition have used so-called input-based measures of education quality, such as per-pupil spending. Hanushek (1986, 1997) found that school inputs have no apparent impact on student achievement and are therefore inappropriate as measures of school quality. This has led to the more prevalent use of output-based measures, such as standardized test scores.

Some authors, however, have expressed concerns about the potential endogeneity of school quality when it is measured by indicators of student performance. Gibbons & Machin (2003), for example, argue that better school performance in neighborhoods with high house prices may reflect that wealthy parents buy bigger houses with more amenities and therefore devote more resources to their children.

The research on education production functions also has made the case that value-added measures of achievement—often measured as the marginal improvement in a particular cohort's performance over a period of time—would be more appropriate as measures of quality in capitalization studies. However, constructing value-added measures requires tracking groups of students over time and implies more sophistication in the decision making process of potential buyers, as value-added measures are not commonly available to the public. Brasington (1999), Downes & Zabel (2002) and Brasington & Haurin (2006) found little support for using value-added school quality measures in the capitalization model; they argued that home buyers favor, in contrast, more traditional measures of school quality in their housing valuations.

We take two different approaches to measure school quality, which we describe in turn below.

First, we construct a measure that is based solely on students' grades. Specifically, for each school in any given year, we average over students' grades in exams that are taken in their last year of compulsory education (9th grade). These students take national exams in a wide range of subjects and complete them, for the majority, at age sixteen. 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.

Second, we measure school caliber through an index of teacher quality. This measure is construed as follows and is based on Gensowski et al. (2020). 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 prediction.

2.3 House prices

Our outcome variable is defined as the value of the property owned by the biological parents. It is measured at the start of the school year in which the final exam is passed. We take the natural logarithm of house prices in the empirical analysis. Moreover, we drop outliers in terms of property values, when below $1,000 and above $10 million (2010 USD).

2.4 Control variables

To complement the above 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 to reduce any potential bias arising from sorting. More specifically, we use variables pertaining to the household, including income, education level, crime as well as information on family structure. We aggregate these measures both at the large cluster level and at the school level. Moreover, we include a host of housing characteristics such as number of bedrooms, number of floors, living area, age of building and house type. This paper relies on administrative data available for the whole population of Denmark as provided by Statistics Denmark. We present summary statistics in Table 1

		Mean	St. Dev.
Neighborhood Level			
	HH Gross Income (Excl. Tr.)	59,469	25,892
	HH Avg. Years of Schooling	13.2	1.2
	Married Household (%)	.756	.232
	Not Intact Household (%)	.524	.252
	Foreigner (%)	.061	.058
	Non-Western Foreigner (%)	.037	.049
	Crime (%)	.023	.010
School Level			
	HH Gross Income (Excl. Tr.)	108,002	39,982
	HH Avg. Years of Schooling	13.6	1.1
	Married Household (%)	.753	.105
	Not Intact Household (%)	.524	.116
	Foreigner (%)	.089	.082
	Non-Western Foreigner (%)	.051	.068
Housing Attributes			
	Age Building	54.2	35.5
	Living Area	148.5	48.2
	Number of Floors	1.24	.79
	Number of Appartments	9.31	39.2

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 2002-2015. We focus on homeowners, since we can observe their housing values. Incomes are converted from DKK to 2010 USD.

Table 1: Summary statistics.

In our analysis, we focus solely on homeowners, given the construction of clusters mentioned in our introduction.

2.5 Clusters as Neighborhoods

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

1. Clusters correspond to geographical areas within which an individual has social contact;
2. They should be unaltered over a specified period of time;
3. Allow to be combined with administrative register information.

The last rule enforces cluster sizes 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.

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 and following the above mentioned rules.

In order to capture neighborhood in which individuals would experience most of their social interactions, the authors created the small (minimum 150-households) and large (minimum 600-households) clusters such as to be just above the confidentiality requirements of statistics Denmark. 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 that within cluster differences in house prices are not driven by these barriers.

The clusters are defined based on housing type and ownership information[2].

[2]Housing type in the register data is divided into four categories: farmhouse or detached house; townhouse

	Median	Mean	Std.Dev.	N
Small Neighborhoods				9086
Households 2004	245	272.7	115.0	
Persons 2004	526	592.2	285.5	
Size (in hectares)	22	47.5	64.46	
Large Neighborhoods				2295
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)

Table 2: Summary Statistics of Clusters

Since a given address always holds the same cluster ID, we were able to expand the cluster IDs beyond 2005 for the all addresses that existed before 2005. We followed two further rules to expand the cluster ID. First, if the date of a person's last recorded move was before 2005, then they would be matched to their cluster ID in 2005. Second, if a person did not move in the last two years and did not have a non-missing cluster ID in the last year, their current year cluster ID was replaced with it. As a robustness checks, we ran our estimates in 2002, 2003, and 2004, where original clusters are available, and our results did not differ qualitatively.

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.

3 Methodology

This paper develops a framework to recover household preferences for school quality. Our empirical strategy builds on Tinbergen (1956) and Rosen's subsequent seminal work (1974). Rosen's analysis rationalizes the interaction between consumers and suppliers in a competitive markets with differentiated goods. An equilibrium is reached where consumers bid functions and suppliers offer functions kiss each other along the hedonic price function $P(z)$ where $z = (z_1, z_2, \ldots, z_n)$ is a 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 lower housing prices, to attract potential buyers. In this framework, at each point on the hedonic price function, the marginal price of a housing characteristics is the individual consumer's marginal willingness to pay for that characteristic and will be equal to the individual supplier's marginal cost of producing 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). In the context of neighborhood amenities, a key issue that the literature has aimed to address is the endogeneity of school and neighborhood characteristics which arises from household sorting.

One strand of the literature builds on Black (1999) who developed the boundary discontinuity design (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.

Bayer *et al.* (2007) further build on this methodology by developing a heterogeneous model of residential sorting and clarify conditions under which coefficients in a hedonic price regression are likely to provide an approximation to the mean marginal WTP (mean MWTP) of the population. Specifically, when attributes - such as school quality - vary more or less continuously across geographic areas, mean preferences estimated in the heterogeneous sorting model and the coefficients of the hedonic price regression would differ only slightly.

While BDD has been popular, issues of sorting around the borders have also led to the exploitation of temporal variation in either school characteristics or assignment to identify the valuation of schools. For instance, DD 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. A further important drawback to methods that use temporal shocks to school quality to derive WTP estimates is exposed 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 indeed 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.

For these reasons, and because we do not observe school boundaries, we build on a neighborhood fixed effects strategy to recover household valuation of a vector of neighborhood attributes. To do so, our strategy capitalizes on variation in house prices and school quality within very small neighborhoods (median size of 0.3 square mile) arising from discontinuous school catchment area within these clusters. Given the small geographic area that are spanned by these clusters, we are able to hold constant unobserved neighborhood amenities, such as air pollution, access to parks or healthcare. This is our first assumption, i.e that

unobserved housing and neighborhood characteristics do not vary systematically within the geographic clusters we use as fixed effects.

Furthermore, because the literature emphasises the importance of hyper local amenities and attributes, we add a host of neighborhood socio-demographics controls to capture the remaining potential characteristics that may matter to individuals and thus be correlated with school quality. Our second identifying assumption is that the measures of immediate neighborhood income, education, crime, marital status and origins included in the regression control fully for the attributes that households care about and that would be capitalized into housing prices.

4 Hedonic Price Regressions

4.1 Income Inequality and Segregation across Neighborhoods

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

To analyze the spatial decomposition of income inequality across neighborhoods in Denmark, we use the Theil's T Index. Figure A.5 of A.3 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 increase by a factor of 5 (to more than 26%) when we analyze the income inequality across small clusters.[3]

We also use the Gini Index measure to analyze the the income inequality at different neighborhood units in Denmark. Figure A.6 of Appendix A.4 plots the distribution of neighborhood Gini index for various neighborhood units, i.e. municipality, parish, large cluster, and small cluster. Figure A.6 of Appendix A.3 shows that the income inequality is dramati-

[3] Appendix A.3 discusses how we use the Theil's T Index to compute within- and between- neighborhood inequality.

cally lower at cluster levels, suggesting more homogeneity among individuals living close to each other in a our neighborhood unit.

We also analyze the segregation intensity over the income distribution by neighborhood unit. Figure A.7 of Appendix A.3 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 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 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. Appendix A.5 discusses the 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.[4]

4.2 Baseline and FE Estimates of School Quality Valuation

In this section, we describe our regression framework, which provides estimates of the marginal WTP for school quality, measured by test scores, and showcase the importance of controlling for unobserved neighborhood characteristics through our cluster fixed effects strategy as well as hyper local neighborhood attributes.

Based on Rosen's (1974) framework presented above, our main estimating equation relates house prices to a vector of housing and neighborhood characteristics, including school

[4]Results presented in Appendix A.5 focuses on income measures, but we observe a similar pattern when we look at other characteristics such as education level.

quality. We add a set of cluster fixed effects to control for unobserved neighborhood heterogeneity and estimate the following hedonic price regression:

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

where p_{imk} denotes logged 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 characteristics at the individual level denoted by X_{imk}. Finally, c_k, τ_t and ρ_{kt} are respectively neighborhood, time and time-by-neighborhood fixed effects.

We present a set of baseline estimates in columns (1) of Table 3. 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 18%, reflecting the important bias that would ensue from such a model, as pointed to in the literature. In our second OLS specification, we also add a vector of hyper local neighborhood characteristics, such as average income and education (presented in more details in the next section). This specification also includes a set of housing characteristics, including the number of floors, the number of units per building, the number of bedrooms, the living floor area as well as age of the building. The coefficient on test scores now decreases notably to 5.0%, 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 columns (2) of Table 3. 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.

Putting these two strategies together, the final column adds both the neighborhood fixed effects and the hyper local characteristics to recover the marginal WTP for school quality. Our estimate of 3.2% reflect that for the average housing, a one standard deviation increase

in average test score increase house prices by $\$10,000$. 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 (recall that the median size of these neighborhoods are around 0.3 squared miles), 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. The magnitude of the reduction in their estimate between the OLS and FE model is similar to ours. We note, however, that our estimate is about twice as large as the one they find, indicating greater mean MWTP for school quality in Denmark than in the Bay Area.

	(1) OLS		(2) FE	
Avg. Test Scores	.182***	.050***	.057***	.032***
Nbhd characteristics	No	Yes	No	Yes
Housing characteristics	No	Yes	No	Yes
Large cluster FE	No	No	Yes	Yes
Observations	233,629	205,167	231,008	202,725
R^2	. 11	.41	.35	.51

Note: Columns (1) show an OLS specification as a benchmark, while columns (2) show two different specifications with clusters as fixed effects. Sample includes all parents in Denmark whose children attend 9th grade 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 are clustered at the large cluster level. Neighborhood characteristics include parental gross income, and education as well as fraction married, intact family, crime, foreigners, and private schools, at neighborhood level. Housing characteristics include the number of floor, the number of units per building, the number of bedrooms, living floor area and age of property. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3: Regression results. Contrast between naive and FE estimates

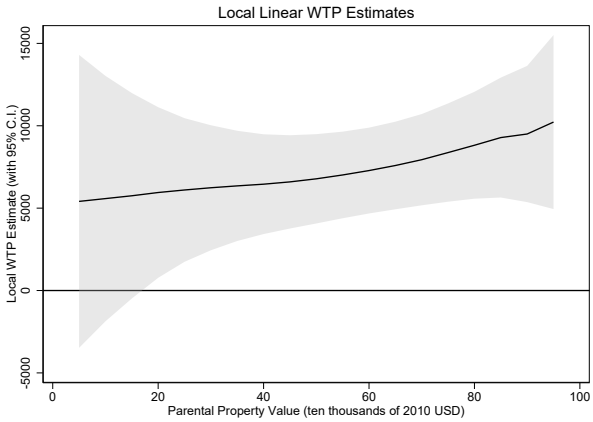
4.2.1 Heterogeneity

To explore the heterogeneity across individual characteristics, this section provides local linear estimates of the above WTP estimates, controlling for the same family and housing characteristics and using fixed effects at the small cluster level.

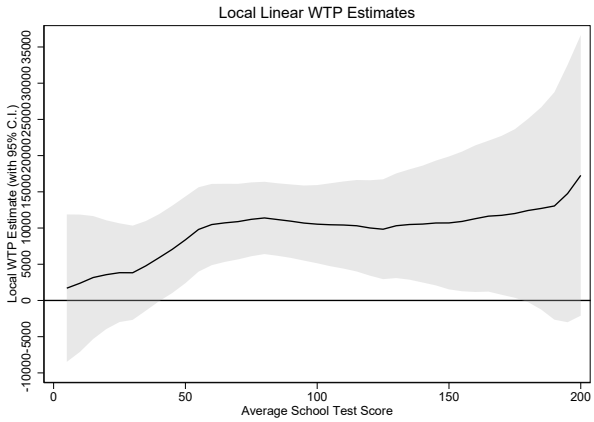
To do so, we run the same hedonic model as above, at different intervals along the distribution of specific individual-level characteristics. We look in turn at heterogeneity in property values, school quality, household income and parental education. Results presented in Figure 3.

From Figure 3, a general trend seems to be visible, whereby individuals with higher income, education as well as those with higher property values, tend to be willing to pay more, at the margin, for a one standard deviation increase in school quality. We also provide evidence that along the school quality distribution, the MWTP increases as well.

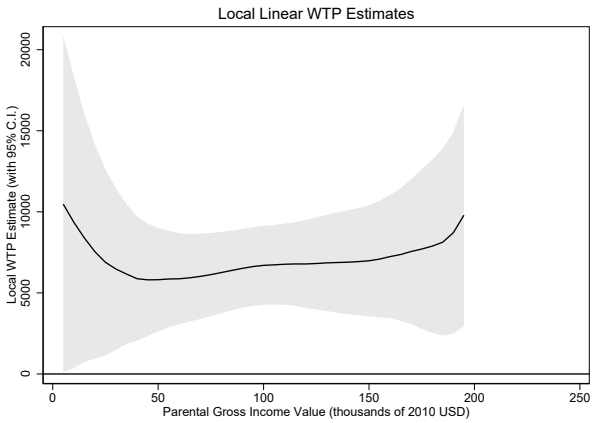
The non-linearity of these willingness to pay estimates speaks to the important role of "Matthews effect" in the provision of public goods in Denmark, whereby advantages from Denmark's public education system are reaped relatively more by the affluent rather than by the disadvantaged.



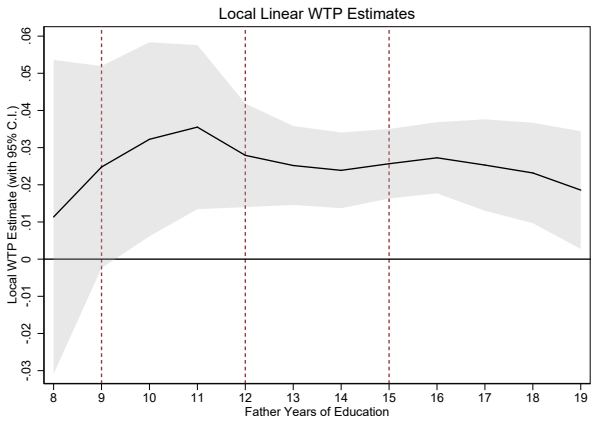
(a) Property Value



(b) School Quality



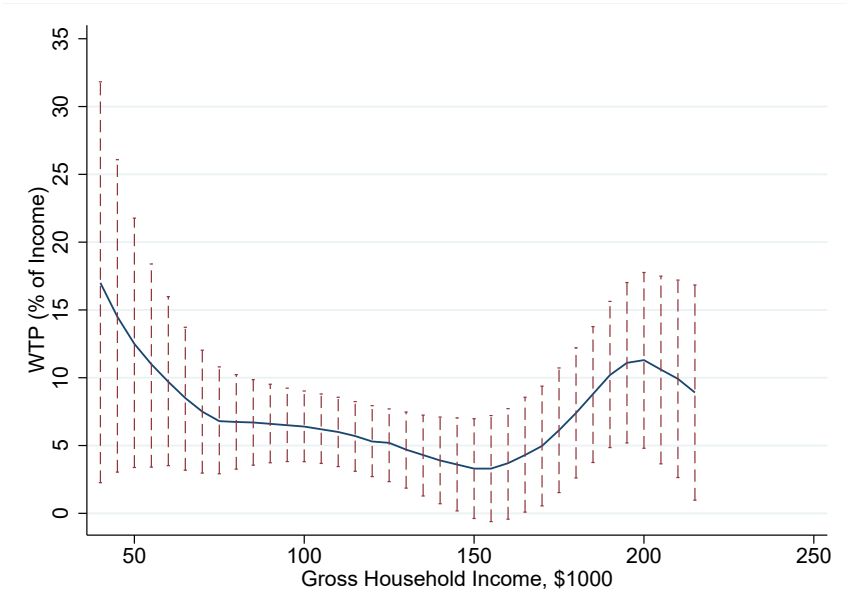
(c) Parental Gross Income



(d) Parental Education

Figure 3: WTP Heterogeneity: Local linear estimates of WTP showing heterogeneity of our estimates with respect to parental property value (Panel (a)), school quality as measured by average grade (Panel (b)), parental gross income (Panel (c)), and father's years of education (Panel (d)). Panel (a)-(c) show the WTP as a dollar amount. Panel (d) shows WTP as a percentage of housing values with three vertical lines referring to the three main school completion levels (primary school, high school and college).

Building on this, we propose here a simple characterization of the economic significance of these WTP estimates. Our goal here is to get a better understanding of the share of household income that would be spent on improving school quality. To do so, we consider our heterogeneous WTP estimates along the household income distribution presented above, but this time look at its importance as a share of income. The resulting U-shaped relationship provides evidence that households who spending most as a share of their household income on school quality are those lying at the bottom and those at the top of the income distribution.



Note: This figure shows the relationship between fraction of income spent for a one standard deviation increase in school quality, along the income distribution (from the bottom 5% to the top 95%). We use the following concepts. To compute the total amount spent on schooling, we use our local linear estimates of the WTP along the income distribution. We then divide by gross household income excluding transfers. School quality is measured by average test score at school level. We further note that the amount spent to get access to a better school corresponds to a one-off payment when purchasing a house, while income is measured on a yearly basis.

Figure 4: Share of income spent on schooling.

4.3 Valuation of Neighborhood Characteristics

We now turn to the valuation of neighborhood characteristics using the model presented above. We contrast two specifications, a standard hedonic price regression with 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 4 include 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, sometimes very notably. For instance, the coefficient on education is divided by six, whereas the coefficient on income is almost divided by three and becomes insignificant. Thus, controlling for unobserved heterogeneity, a one year increase in parental education at the neighborhood level, increases house prices by 4.8%, significantly less than the 26% estimate from the OLS specification. This clearly reflects that such attributes are positively correlated with unobserved 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 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 for households value neighborhood school quality above and beyond the school quality of their catchment area. The estimate, standing at 2.1% can be interpreted as reflecting households' valuation of neighborhood peers, beyond that of school peers.

A further set of estimates experience more radical changes through the introduction of unobserved heterogeneity. First, the coefficient on fraction of individuals who have committed a crime adjusts from 2.99 to -.093 and becomes insignificant at conventional levels. This

change in sign reflect may reflect the positive correlation between crime and economic activity, which we are able to capture through our neighborhood fixed effects strategy. Second, the coefficient on share of married household, goes from -1.1% to 3.9%. This is of a relatively small magnitude, but displays that households do indeed also care about household structure of their neighbors.

Finally, the coefficients on fraction of foreigners and fraction of non-western foreigners both become insignificant. It does not come as a surprise that in the Danish context, the coefficient on non-Western migrants is negative in an OLS specification. It is in fact the case that non-western migrants tend to be poorer and less educated, on average. Once we control for these characteristics at the neighborhood level, it becomes clear that non-western share is not negatively capitalized into housing values.

Neighborhood Attributes	(1) OLS	(2) FE
Average Test Scores	. 027***	. 021***
Income (log)	. 074***	. 022
Education (years)	. 265***	. 048***
Crime (Fraction)	2. 99***	-. 093
Married (Fraction)	-. 011***	. 039***
Non-intact HHs (Fraction)	. 023***	-. 013**
Foreigners (Fraction)	3. 51***	. 207
Non-western (Fraction)	-2. 21***	. 027
Private School (Fraction)	0. 73***	. 055***
Housing Characteristics	Yes	Yes
Observations	205,167	202,725
R^2	. 41	. 51

Note: Column (1) shows an OLS specification while column (2) adds cluster fixed effects. 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 a given neighborhood attribute. Housing characteristics are added as regressors and include the number of floor, the number of units per building, the number of bedrooms and living floor area. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4: Regression results. Valuation of Neighborhood Characteristics.

4.4 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, which we have thus far used to describe

school quality.

Furthermore, parents may also care about the quality of peers themselves at the school level, beyond those that are in their immediate neighborhood.

Our empirical strategy allows 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 5 reveal that this is not the case. The coefficient of .014 is around half of that of the coefficient on average test scores presented in Table 3. There are two potential reasons for such result. First, test scores are much more widely publicized and available to 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 5 we present estimates of the whole vector of school characteristics. This further comforts the two hypothesis made above. Parents do indeed seem 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 .17%. On the other hand, the coefficient on teacher quality becomes insignificant, while that on test scores is divided by almost three.

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.

School Attributes	(1)	(2)
Teacher quality	. 014***	. 004*
Test scores		. 011***
Income (log)		. 167***
Education (years)		. 011**
Married (Fraction)		. 098***
Non-intact HHs (Fraction)		-. 038***
Foreigners (Fraction)		. 001
Non-western (Fraction)		. 000
Housing Characteristics	Yes	Yes
Nbhd Characteristics	Yes	Yes
Large Cluster FE	Yes	Yes
Observations	202,725	202,691
R^2	. 51	. 51

Note: Column (1) shows a cluster fixed effects specification with teacher quality as only regressor at the school-level. Column (2) further adds a whole vector of school-level covariates. 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 a given attribute. Neighborhood and Housing characteristics are as above. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5: Regression results. Decomposing school characteristics.

5 Extensions and Sensitivity Analyses

5.1 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 and Pope (2014) present a theoretical framework showcasing 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 get modest fluctuation over time in the marginal WTP for school quality.

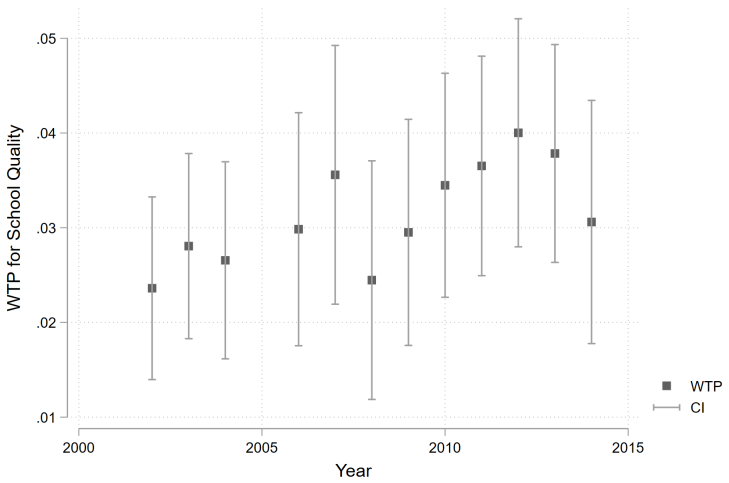


Figure 5: WTP for School Quality by Year

5.2 WTP over the life cycle

This section takes a life cycle approach to the household valuation of school quality. We look here at the heterogeneity in the estimates based on years before and after the arrival of a child.

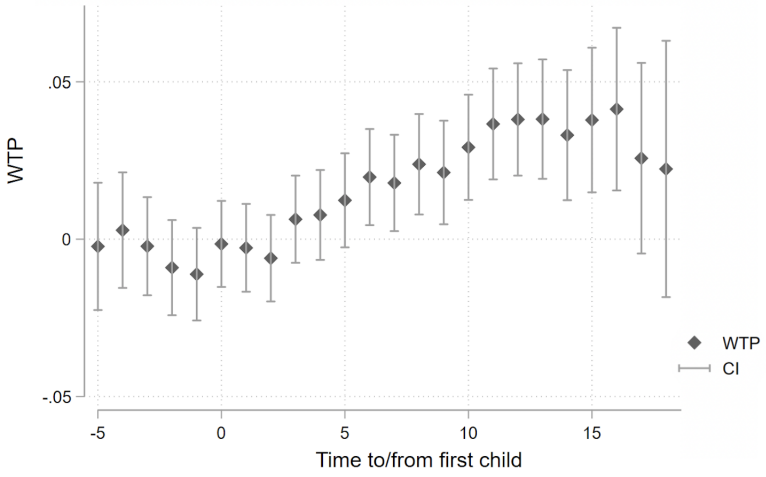


Figure 6: Heterogeneous WTP for School Quality along the Life Cycle

5.3 Further robustness checks

In this section, we report several further robustness checks of our results. We show that none of these sensitivity checks changes the nature of the estimated coefficient. Results are presented in Table 6.

First, and importantly, we have thus far not dealt with the anecdotal evidence that some households defy school catchment areas and find means to allocate their children to better schools to which they are not assigned. This issue, albeit serious, would bias our estimates downward, since the possibility for parents to defy boundaries would imply that in equilibrium school quality would not be capitalized into house prices. To exemplify, at the limit, if every parent can decide freely their child's school, then there is no reason to believe that we can find a positive relationship between house prices and school quality. Although we do not observe directly school boundaries, we can provide a robustness check to assess this issue. To do so, we consider all clusters satisfying two conditions: (1) more than 30% of students in the cluster attend the same school and (2) more than two schools are attended by at least 30% of students. Doing this, we are dropping clusters were students are spread around across too many schools, and where there may not be a clear boundary crossing. We are

left with a sample of approximately 30,000 students who lie in clusters where for the large majority two schools are attended thus increasing the possibility that there is a boundary crossing within the cluster. We distinguish two sets of results based on whether the cluster population is greater than 5 or 30 students, both pointing in the same direction. As expected, results from such specification increase and stand at 3.8%. This increases our confidence that we are able to estimate a lower bound for the willingness to pay for school quality.

Second, we look at whether using house values directly from sales data has an impact on the estimated coefficient. In fact, thus far we have used data that is based on prices from house transactions, but extended to the whole population of houses by Statistics Denmark. We show that our estimates are robust to this different measure of our dependent variable. We note that since only a fraction of houses are sold on the market every year, the number of observations drops for this analysis.

Third, we look at the impact of using past values of our school quality measures. A criticism pointed out earlier in the paper by Gibbons & Machin (2003) note 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. We do so for different lags and verify that our results are robust to this potential criticism. The estimates are however attenuated towards zero when considering significant lags, such as using school quality measures ten years prior. This is most likely due to attenuation bias from the measurement error we induce by using quality measures that are seriously outdated.

Finally, we drop students who attend private schools. Our estimates our robust to focusing only on students who attend public schools in Denmark. This should not come as a surprise since private schools are expensive and could only partially be considered as substitutes for the larger share of households.

	(1) Treating Defiers		(2) Transaction	(3) Lags of Quality			(4) Private Schools
	nbhd pop >5	pop >30		1st lag	5th	10th	
Avg. School Test Scores	.039***	.031***	.036***	.029***	.026***	.019***	.030***
adj. R^2	.56	.55	.44	.50	.49	.46	.50
N	27,460	7,707	46,214	161,093	103,303	27,960	190,090

Note: Table shows results from the various robustness checks presented. Column (1) shows our estimates from conducting a subsample analysis aimed at removing defiers, as explained in the text. Column (2) shows our estimate from using data on property transactions. Column (3) shows results from replacing our variable for school quality by lags of it (using average test score as measure of school quality). Finally, column (4) removes private school goers from the sample. 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 fixed effects are included. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 6: Regression results. Robustness Checks.

6 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. Dahl (2002) proposes a methodology to control for selection in a setting with polychotomous choice.

We consider the following selection model presented below. 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 parish 1, α_1 is a parish 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. 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:

$$\varepsilon_{i1} = \max_{j \neq 1}\{V_{ij} - V_{i1}\} = \max_{j \neq 1}\{z_i'\gamma_j + \eta_{ij} - z_i'\gamma_1 - \eta_{i1}\} \tag{2}$$

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

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

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 very large number of parameters, which rapidly makes it intractable for practical implementation. As a result, the 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 the next section.

6.1 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 to neighborhoods in the same quantile of neighborhood income in which they live.

Then, for each neighborhood in Denmark, we compute the WTP in two ways. First, we compute corrected coefficients using the above methodology proposed by Dahl (2002). Second, we run a set of neighborhood-level OLS regressions controlling for the same individual and housing characteristics, as in the corrected estimates.

Figure 7, presents a scatter plot of these estimates for 190 neighborhoods. We see that both corrected and uncorrected estimates are very close to each other.

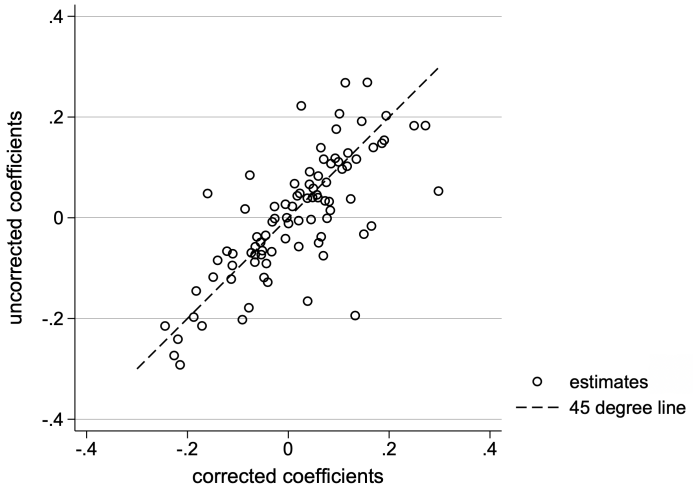


Figure 7: Scatter Plot of Corrected vs. Uncorrected Estimates

To better assess how these two sets of estimates differ, Figure 8 presents the densities of both the corrected and uncorrected estimates. We see that the density of the corrected coefficients 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.

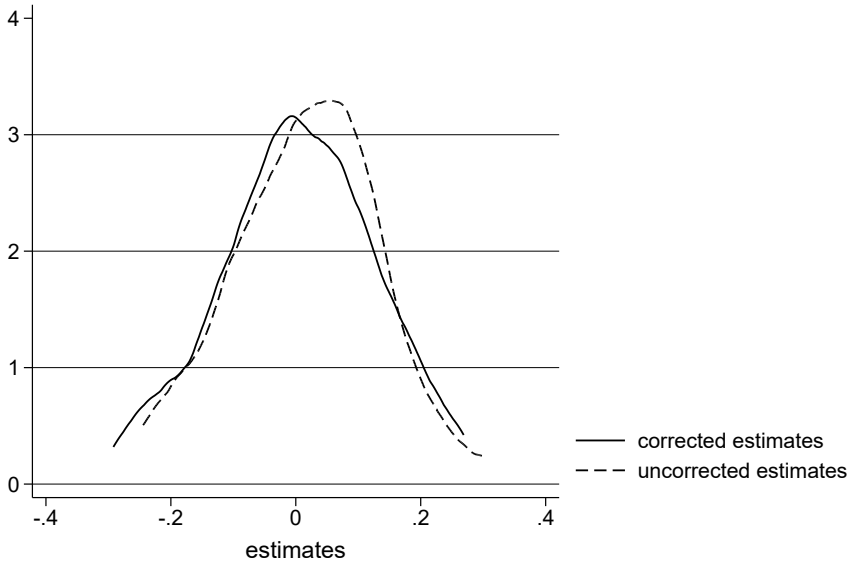


Figure 8: Distribution of neighborhood-level corrected vs. uncorrected estimates.

7 School Quality & Later Life Outcomes

In this section, we turn our focus to assessing the effect of school quality on later life outcomes. We look at two main outcomes of interest, namely college completion and earnings. Beyond households' willingness to pay for better school, a key question of interest is whether better schools actually lead to better latter life outcomes. Such analysis can shed 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, for our earliest cohort. Data availability on school quality based on test scores commences in 2002. We restrict our data to cohorts who complete 9th grade between 2002 and 2005. This is to avoid capturing individual outcomes too early in their life cycles. For these cohorts, we can observe college completion in 2020, when they reach the ages of around 30 to 33. On the other hand, the

latest data on earnings is available in 2019. Since taking a single year would provide a noisy measure of earnings, we average over wages between 2018 and 2019. This means that for the later cohorts, incomes are measured between ages 28 and 30, while they are measured between ages 31 and 33 for the earliest cohorts. The measure of earnings we focus on here is annual gross income excluding transfers.

Estimating the effect of school quality on later life outcomes requires to account for several important endogeneity issues, as discussed in the literature (see for instance Dearden *et al.*, 2002, 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 model 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., 2020).

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 indeed 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 household fixed effects model. More specifically, we look at within variation in school quality amongst households with different 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 household 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 qualiy on later life outcomes. We have here provided evidence that school finances are very homogeneous across Danish neighborhoods and schools. Moreover, our focus within neighborhood variation allows us to fully address this issue.

Our model of the education production function is as follows:

$$y_{imjkt} = \alpha + \beta S_{mjkt} + \gamma X_{jkt} + f_k + \tau_t + \rho_{kt} + \varepsilon_{imjkt}, \tag{3}$$

where y_{imjkt} denotes college completion of individual i who attends school m in small cluster j within large cluster k in time period t. It also corresponds to earnings in our second specification. S_{mjkt} denotes our measure of school quality. X_{jkt} denotes our vector of housing, neighborhood and individual-level controls. f_k denotes our large neighborhood fixed effects. Finally, τ_t and ρ_{kt} denote respectively time and time-by-neighborhood fixed effects.

As mentioned above, we also add household fixed effects, in some specifications. Moreover, to control for the fact that individuals may locate in different labor markets later in life, we add municipality fixed effects to our model looking at the effect of school quality on earnings.

In the household fixed effects model, our identifying assumption is as follows. Within households, individuals who are the same in terms of observable but who attended schools

with varying levels of quality, do not differ on average, in terms of unobservables.

7.1 Results

We present our results on college completion in Table 7. The outcome variable being dichotomous, we run the above specification using a logit model. As discussed above, we find that the OLS estimates are significantly upward biased. These would imply that a one standard deviation increase in school quality leads to 37% greater probability of completing college. In contrast, the neighborhood fixed effects regression which also includes hyper local neighborhood controls provide an estimate of 10%. This difference is driven by parents who willingly sort across neighborhoods and thus leads to correlate neighborhood attributes with school quality. When controlling for household fixed effects, in column (3), we actually find that the coefficient increases to 17%. However, here again, controlling for observed and unobserved neighborhood, reduces the coefficient down to 12%. This result contrasts with the slightly downward biased neighborhood fixed effects specification, which stood at 10%. This difference may reflect the fact that parents who choose better schools also substitute away from parental home investment, as discussed above.

This evidence shows that parents not only are willing to pay significant shares of their income for better schools, and importantly this then results into better outcomes for their offspring.

School quality	1. 37***	1. 10**	1. 17***	1. 12***
Individual characteristics	No	Yes	No	Yes
Nbhd characteristics	No	Yes	No	Yes
Housing characteristics	No	Yes	No	Yes
Large cluster FE	No	Yes	No	Yes
Household FE	No	No	Yes	Yes
Municip. FE (in adulthood)	No	No	No	No
Observations	155,342	138,180	8,561	8,042

Note: Column (1) shows a Logit specification as a benchmark, while column (2) shows a logit specifications with large clusters fixed effects. Column (3) shows a household fixed effects specification, without controls, while the last column adds controls. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2015 and own a property. Housing and neighborhood characteristics as above. Individual characteristics include income, education, crime, distance to work, marital status, birth order and gender of the child. Coefficients can be interpreted as odds ratios.

Table 7: Regression results. College Completion.

We now turn to presenting results on the effect of school quality on income in Table 8. Showing the same set of specification across columns, the effect on later life income ranges from 2.9% to 9.7%, depending on the model.

Income	0. 089***	0. 022***	0. 029***	0. 107**	0. 095*
	(0. 007)	(0. 011)	(0. 008)	(0. 050)	(0. 054)
Individual characteristics	No	Yes	Yes	No	Yes
Nbhd characteristics	No	Yes	Yes	No	Yes
Housing characteristics	No	Yes	Yes	No	Yes
Large cluster FE	No	No	Yes	No	Yes
Household FE	No	No	No	Yes	Yes
Municip. FE (in adulthood)	No	No	Yes	Yes	Yes
Observations	78,570	74,334	73,175	5,507	5,205

Note: Column (1) and (2) shows an OLS specification as a benchmark with and without controls, while column (3) shows a specifications with large clusters fixed effects. Column (4) shows a household fixed effects specification, without controls, while the last column adds controls. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2015 and own a property. Housing and neighborhood characteristics as above. Individual characteristics include income, education, crime, distance to work, marital status, birth order and gender of the child. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 8: Regression results. Later Life Income.

We should note, however, that the 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 examine the impact of school quality on adulthood income while we do not control for the educational attainment in equation 3. Table 9 presents results on the effect of school quality on income in adulthood when we do not control for the college attainment, i.e. the total effect on income from attending a better school quality. Our estimates provide evidence that a one standard deviation increase in school quality leads from 5.3% to 10.9% greater yearly income, from attending better schools - an economically significant impact.

Income	0. 036***	0. 053***	0. 107**
	(0. 008)	(0. 011)	(0. 054)
Individual characteristics	Yes	Yes	Yes
Nbhd characteristics	Yes	Yes	Yes
Housing characteristics	Yes	Yes	Yes
Large cluster FE	No	Yes	Yes
Household FE	No	No	Yes
Municip. FE (in adulthood)	No	Yes	Yes
Observations	74,334	73,175	5,205

Note: Column (1) shows an OLS specification with controls, as a benchmark, while column (2) shows a specifications with large clusters fixed effects. Column (3) shows a household fixed effects specification, with controls. Sample includes all parents in Denmark whose children attend 9th grade between 2002 and 2015 and own a property. Housing and neighborhood characteristics as above. Individual characteristics include income, education, crime, distance to work, marital status, birth order and gender of the child. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 9: Regression results. Later Life Income Without Controlling for College Completion.

8 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 developed a novel empirical strategy to estimate the marginal willingness to pay (WTP) for schools in Denmark, where public schools are free. Our main results reveals 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. Moreover, we link the school quality to later life outcomes of children who attend those schools. Exploiting the variation in school quality among siblings, we found that a one standard deviation increase in school quality leads to around 10% greater probability of completing college, and around 5%-10% increase in annual income at around age 30, in addition to the impact of school quality on college completion.

References

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.

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

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.

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

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. The Construction of Neighbourhoods and its Relevance for the Measurement of Social and Ethnic Segregation: Evidence from Denmark. *Ssrn*.

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

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.

Gensowski, Miriam, Landersø, Rasmus, Bleses, Dorthe, Dale, Philip, Højen, Anders, & Justice, Laura M. 2020. *Public and parental investments and children's skill formation*. Rockwool Foundation Research Unit.

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.

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.

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

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

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

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.

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.

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. 1971. A note on the measurement of racial integration of schools by means of informational concepts.

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

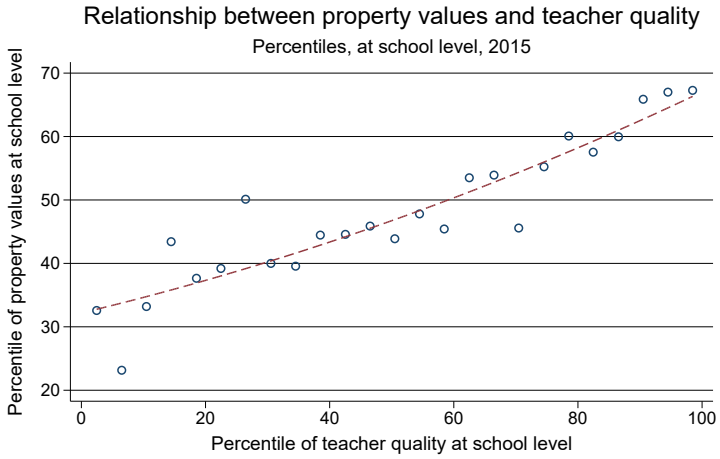
Tiebout, Charles M. 1956. A pure theory of local expenditures. *Journal of political economy*, **64**(5), 416–424.

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

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

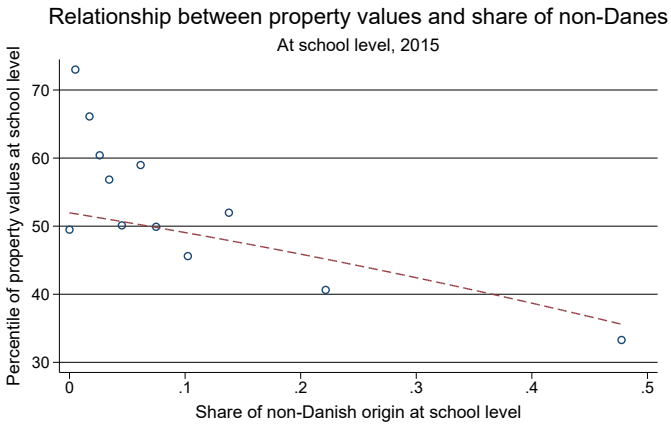
A Appendix

A.1 Sorting



Note: 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 footnote **??**.

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



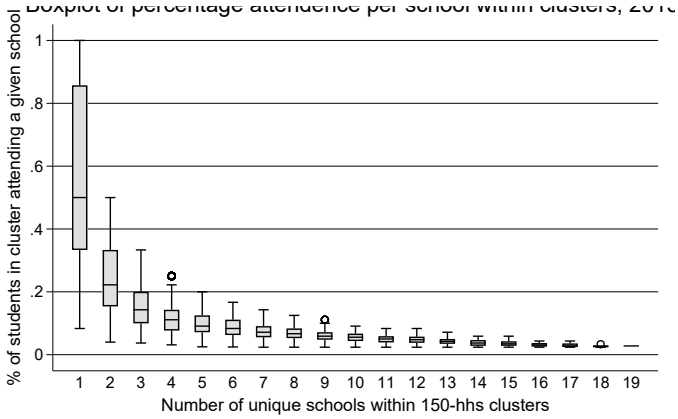
Note: The figure shows a binned scatterplot, with a quadratic fit, of property values percentiles on share of students with non-Danish origins at the school level in 2015.

Figure A.2: Origin of students and property values

A.2 Summary Statistics

A prerequisite for our empirical strategy to work is that sufficient students attend 9th grade in 2015 within each cluster, that these students attend different schools and that the quality of schools attended (as proxied by grades) also varies within clusters. We provide evidence of the former two aspects below in Figure A.3 and the third in Figure A.10 in the Appendix, in subsection A.6. We also provide evidence on the extent to which property values vary within clusters in Figure A.11 of the Appendix A.6.

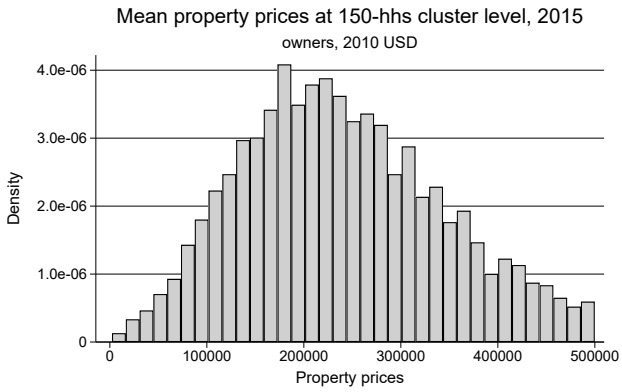
Figure A.3 shows the percentage of students attending each individual school within a cluster.



Note: Each box plot shows the distribution of percentage attendance across clusters for the nth unique school in each small clusters. The horizontal line inside the box represents the mean, while the lower and upper hinge respectively represent the 25th and 75th percentiles. The adjacent values are the most extreme values within 1.5 interquartile range of the nearer quartile.

Figure A.3: Boxplots of percentage attendance of each unique schools across clusters

Figure A.4 shows the distribution of property values of parents who are homeowners in 2015 within clusters.



Note: The figure shows the distribution of parental property values below $\$500,000$ (in 2010 USD) for parents who have children attending 9th grade in 2015.

Figure A.4: Distribution of property values, 2010 USD

A.3 Spatial Decomposition of Inequality

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:

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

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

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

$GE(a)$ 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.

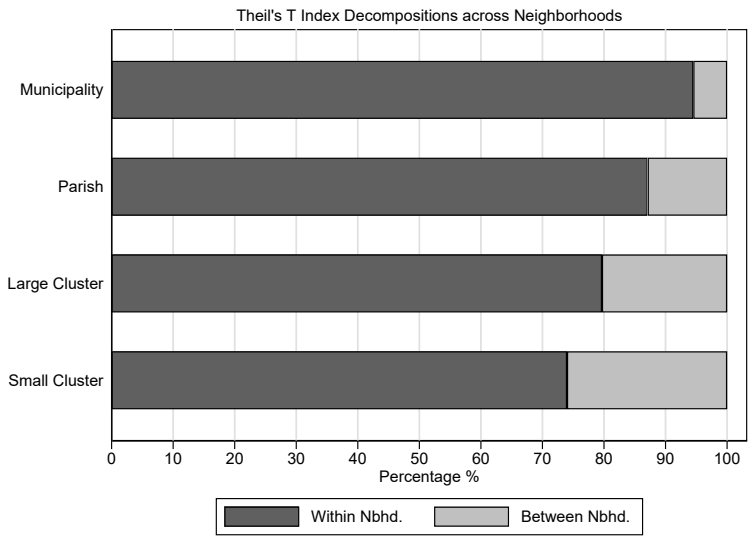
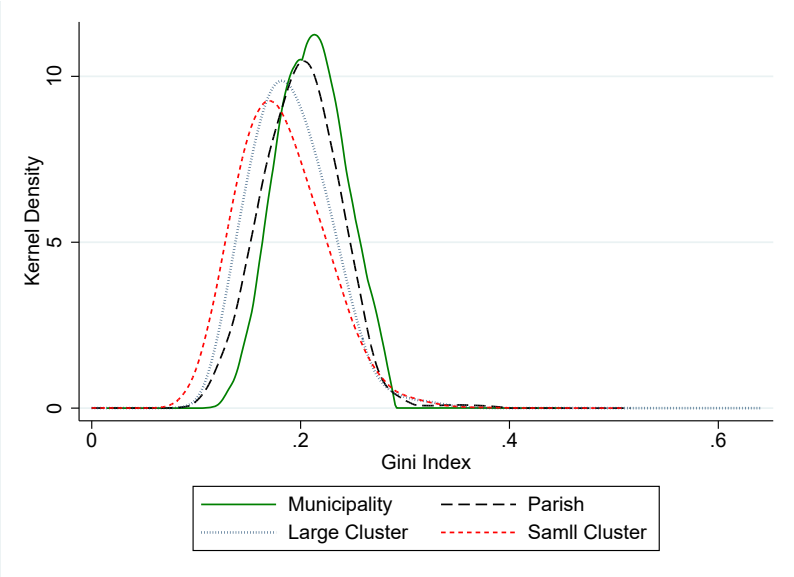


Figure A.5: Variance Decomposition across Neighborhoods

A.4 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.



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.

Figure A.6: Gini Index Kernel Density across Neighborhoods

A.5 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.

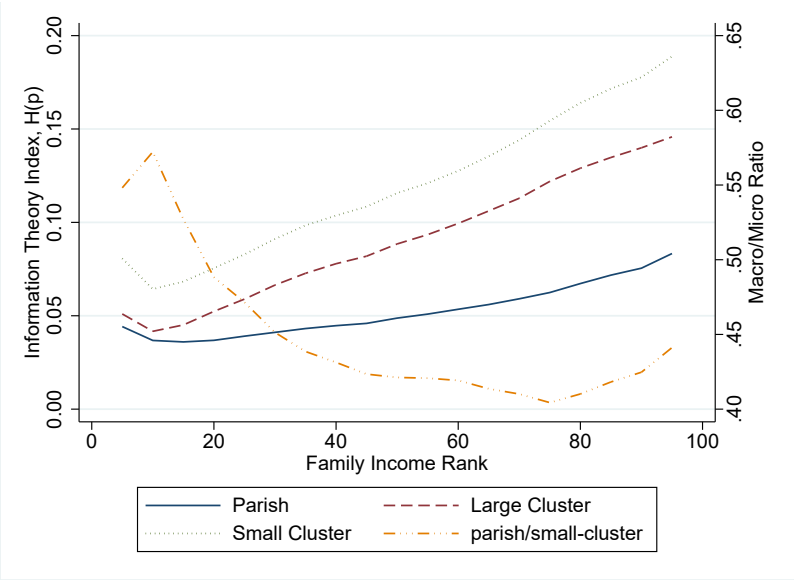
Reardon *et al.* (2006) describe the rank-order information theory index in detail. First,let p denote income percentile ranks (scaled to range from zero to one) in a given income distribution (i.e., $p = F(Y)$) where Y measures income and F is the cumulative income density function). Now,for any given 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, 1971; 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 (1971); Theil (1972); Pielou (1977)). That is,

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

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

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.

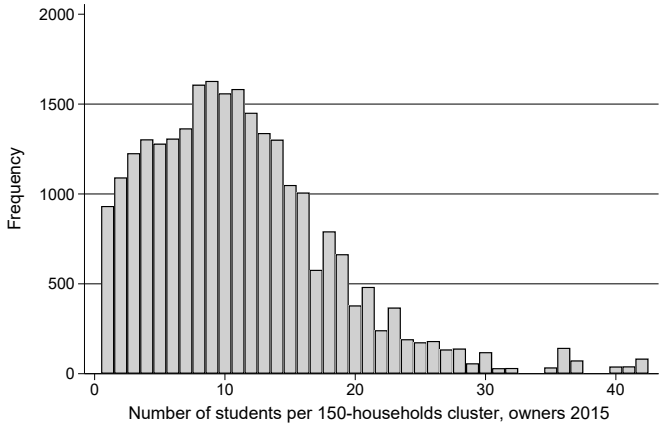
Figure A.7 shows family income segregation for the municipality of Copenhagen in 2015. The figure indicates estimated between-parish (cluster) segregation (as measured by the information theory index, H) between families with incomes above and at or below each percentile of the municipality-wide family 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 separation 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).



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.

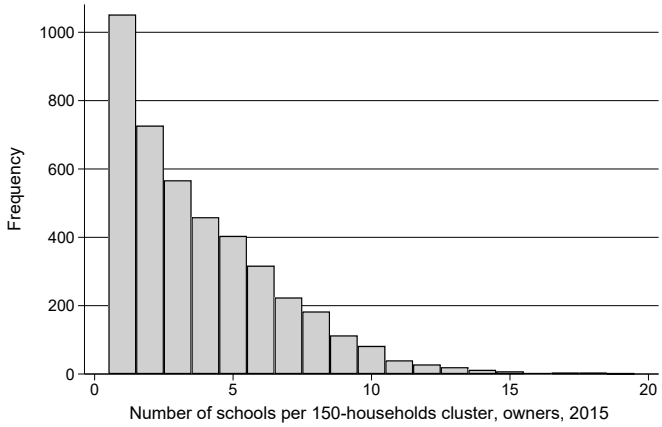
Figure A.7: Segregation Index across Neighborhoods

A.6 Within small cluster variation



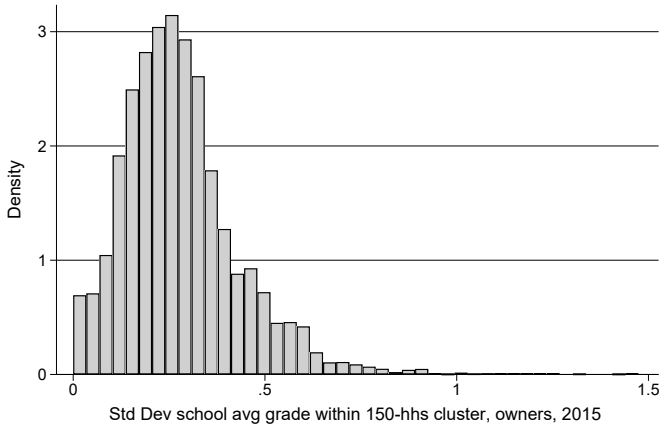
Note: The figure shows the distribution of students within each small cluster who attend 9th grade in 2015.

Figure A.8: Distribution of number of students within clusters



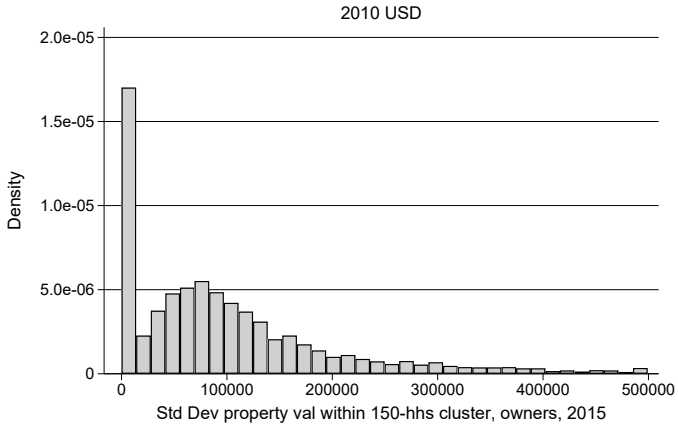
Note: The figure shows the distribution of schools attended within each of these clusters.

Figure A.9: Distribution of schools attended within clusters



Note: The figure shows the variation in the standard deviation of average school grades within clusters.

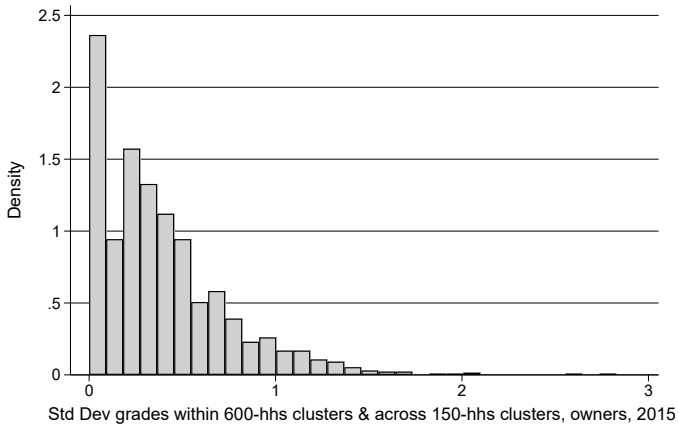
Figure A.10: Distribution of standard deviation of average grades



Note: The figure shows the variation in the standard deviation of property values within clusters.

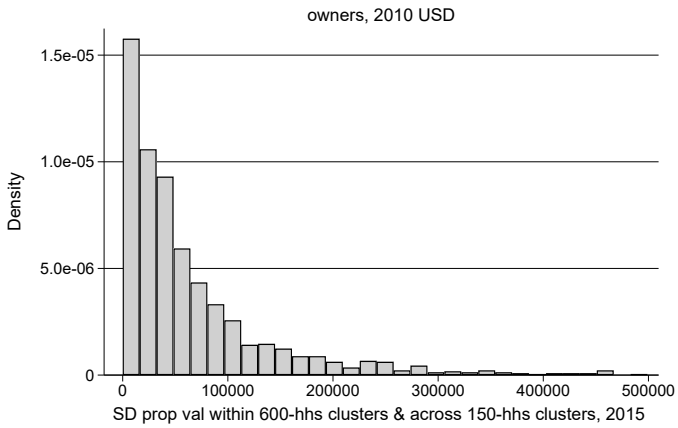
Figure A.11: Distribution of standard deviation of property values

A.7 Within large and across small cluster variation



Note: The figure shows the variation in standard deviation of average grades within large and across small clusters.

Figure A.12: Distribution of standard deviation of average grades



Note: The figure shows the variation in standard deviation of property values within large and across small clusters.

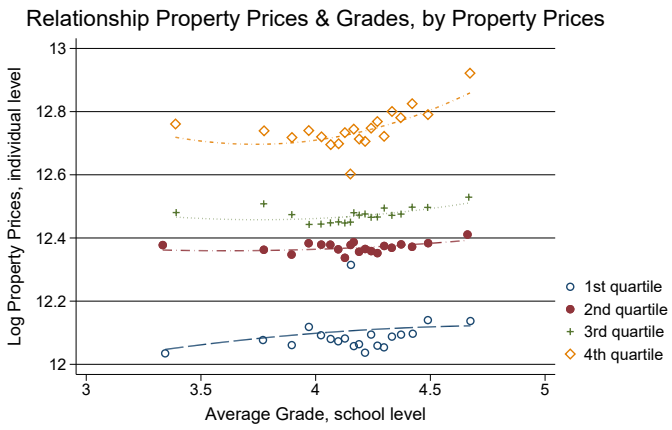
Figure A.13: Distribution of standard deviation of property values

A.8 Heterogeneity of WTP estimates

A.8.1 Within small cluster variation

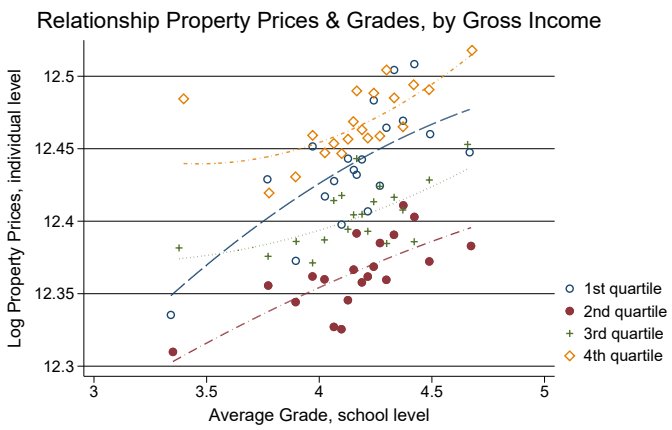
Note: This figure shows a binned scatterplot, with quadratic fit, of the relationship between property prices and school quality measured by average grades at school level. The slope of the fit represents the WTP at any given value of school quality. small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.14: Relationship between Property Prices and Grades



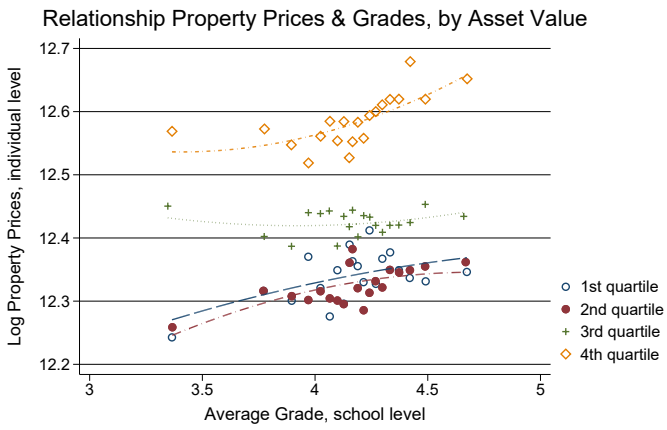
Note: This figure shows a binned scatterplot, with quadratic fits, across quartiles of parental property prices. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.15: WTP Heterogeneity with respect to property prices



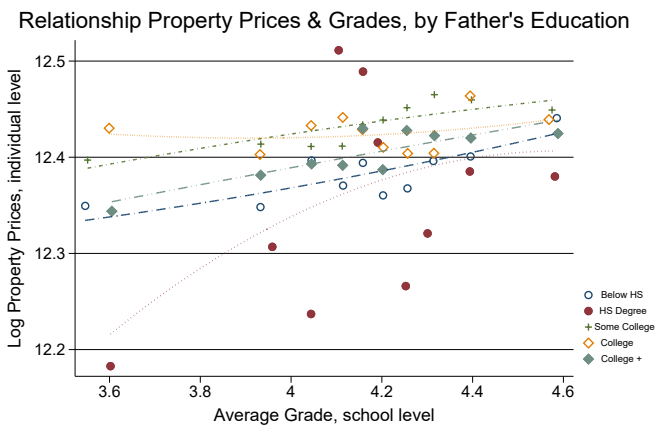
Note: This figure shows a binned scatterplot, with quadratic fits, across quartiles of parental gross income. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.16: WTP Heterogeneity with respect to gross income



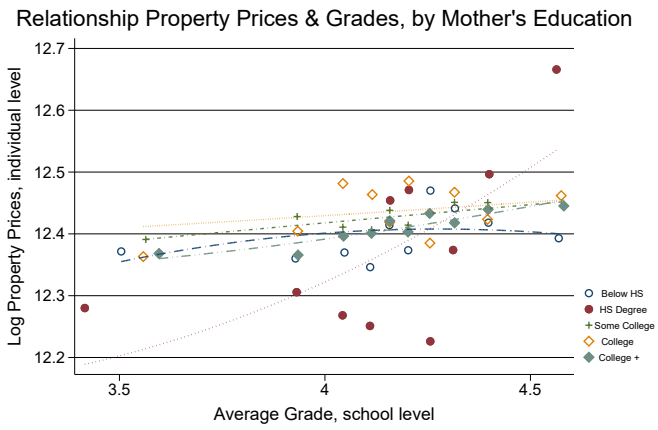
Note: This figure shows a binned scatterplot, with quadratic fits, across quartiles of parental asset value. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.17: WTP Heterogeneity with respect to asset value



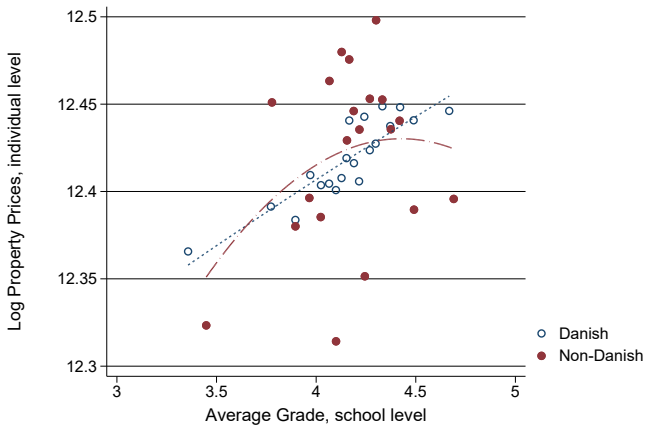
Note: This figure shows a binned scatterplot, with quadratic fits, across quartiles of father's years of education. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.18: WTP Heterogeneity with respect to father's education



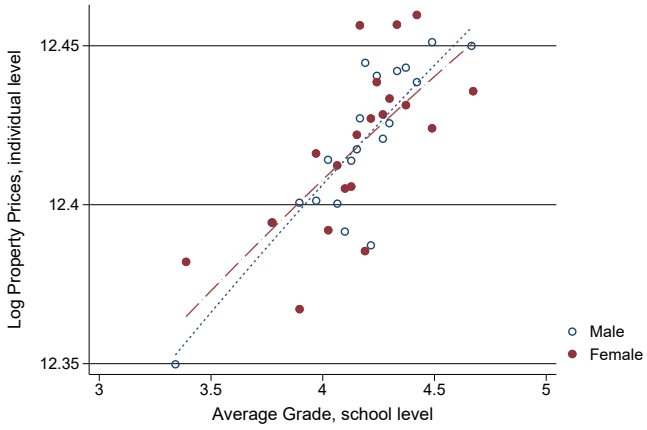
Note: This figure shows a binned scatterplot, with quadratic fits, across quartiles of mother's years of education. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.19: WTP Heterogeneity with respect to mother's education



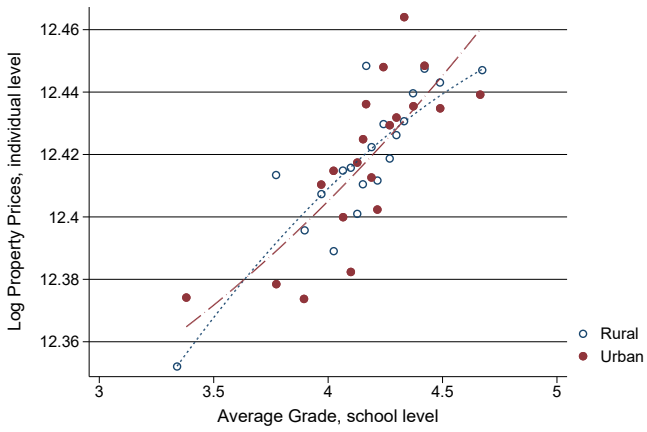
Note: This figure shows a binned scatterplot, with quadratic fits, across origins of children. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.20: WTP Heterogeneity with respect origins



Note: This figure shows a binned scatterplot, with quadratic fits, for both male and female. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.21: WTP Heterogeneity with respect to gender



Note: This figure shows a binned scatterplot, with quadratic fits, for both rural and urban areas. The slope of the fit represents the WTP at any given value of school quality. Small cluster fixed-effects are used as well as the usual controls for family and housing characteristics.

Figure A.22: WTP Heterogeneity with respect to urban or rural area