VATT Working Papers 58

Best Education Money Can Buy? Capitalization of School Quality in Finland

Oskari Harjunen
Mika Kortelainen
Tuukka Saarimaa

VATT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS

58

Best Education Money Can Buy? Capitalization of School Quality in Finland

Oskari Harjunen
Mika Kortelainen
Tuukka Saarimaa

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 2017

Oskari Harjunen, Aalto University School of Economics and City of Helsinki Executive Office – Urban Research and Statistics, Email: oskari.harjunen@aalto.fi

Mika Kortelainen, VATT Institute for Economic Research and University of Manchester, Email: mika.kortelainen@vatt.fi

Tuukka Saarimaa, VATT Institute for Economic Research, Email: tuukka.saarimaa@vatt.fi

We thank two anonymous referees, the editors of the special issue, Essi Eerola, Elias Einiö, Peter Fredriksson, Stephen Machin, Antti Moisio, Mattias Nordin, Tuomas Pekkarinen, Matti Sarvimäki, Olmo Silva, Janne Tukiainen and the seminar participants at CESifo, FDPE, ERSA 2013, HECER, Jyväskylä, RES and VATT for their useful comments. The authors thank the Academy of Finland for funding (grant number 255445).

This version (December 2017) replaces the December 2014 version.

ISBN 978-952-274-130-1 (PDF)

ISSN 1798-0291 (PDF)

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, December 2017

Best Education Money Can Buy? Capitalization of School Quality in Finland

Government Institute for Economic Research
VATT Working Papers 58/2017

Oskari Harjunen – Mika Kortelainen – Tuukka Saarimaa

Abstract

By international comparison, Finnish pupil achievement is high and school achievement differences small. The Finnish education system is unusual also because there are no national testing programs and information on school quality measures is not publicly disclosed. Is school quality capitalized into house prices in this environment? Using a boundary discontinuity research design and data from Helsinki, we find that it is: one standard deviation increase in average test scores increases prices on average by 3 percent, which is comparable to findings from the U.K and the U.S. We argue that this surprisingly large effect is at least partly explained by the inelasticity of housing supply, as we use data from a densely populated urban area. We also show that the effect depends on local land supply conditions within the city and is highest in areas with inelastic supply. Furthermore, the price premium seems to be related to pupils' socioeconomic background rather than school effectiveness.

Key words: Boundary discontinuity, house prices, school quality, spatial differencing

JEL classification numbers: C21, H75, I20, R21

1. Introduction

Empirical evidence from a number of studies shows that differences in school quality, measured by school value-added, test scores, school inputs or peer characteristics, are capitalized into house prices, thus revealing the valuation that homebuyers place on them (Black and Machin 2011; Nguyen-Hoang and Yinger 2011). However, current empirical evidence is tilted towards countries where school quality differences are considerable, mostly the U.S. and the U.K., and where residential location and school choice can potentially make a large difference in the education quality and life chances of children (e.g. Chetty et al. 2016).[1] It is unclear whether these results can be generalized to countries where overall school quality is high and differences in general small.

In this paper, we use hedonic regression techniques with a boundary discontinuity research design to study whether school quality differences are capitalized into house prices in Helsinki, the capital city of Finland. The Finnish case is of particular interest because in recent years the basic education system of Finland has been raised to something of a role model status in many countries.[2] The reason for the interest is that by international comparison Finnish pupil achievement is high and at the same time school level achievement differences are among the lowest in the world.[3]

The Finnish education system is quite distinct in other ways as well. The key features of the Finnish education policy for the purposes of this paper are that there is no central or nationwide testing program in comprehensive schools and standardized

[1] There are two papers that use Norwegian data, but these papers study lower (Fiva and Kirkebøen 2011) and upper secondary schools (Machin and Salvanes 2016), whereas we analyze primary schools.

[2] The Finnish basic education system has received a great deal of attention around the world. One example of this is that the book "Finnish Lessons: What can the world learn from educational change in Finland" written by Pasi Sahlberg and published in 2011, has been already translated to over 20 languages. One can also find a large number of stories praising the Finnish education system in newspapers, see e.g.

http://www.theatlantic.com/national/archive/2011/12/what-americans-keep-ignoring-about-finlands-school-success/250564/

http://www.telegraph.co.uk/news/worldnews/europe/finland/10489070/OECD-education-report-Finlands-no-inspections-no-league-tables-and-few-exams-approach.html.

[3] According to the findings of the OECD's Programme for International Student Assessment (PISA) implemented every third year since 2000, Finnish pupils are among the best performing students worldwide. Perhaps the most striking result from the PISA studies, however, is the extremely low between-school variance in student achievement in Finland. See OECD (2011 and 2013).

tests are not used in evaluating school accountability. Moreover, whenever pupils or schools are tested using standardized tests, the results are not publicly released, but are only used internally by the schools or by researchers for research purposes. Another relevant difference compared to many other countries is that there are no school inspections that could give additional information on school performance based on the subjective evaluation of inspectors.[4] The stated goal of these types of policies in Finland is to give children in all schools equal opportunities so that parents have no need for "school shopping" and would simply send their children to the closest available elementary school. This policy is reinforced by the high qualifications of Finnish teachers who, by and large, all have a Master's degree in education.[5]

School quality and residential-based access to schools also bear on the question of educational and residential segregation. The prevention of socio-economic and ethnic segregation is one of the main objectives of housing policy in Finland, and especially in Helsinki (e.g. Dhalmann and Vilkama 2009, Eerola and Saarimaa 2017). To further achieve this goal, the city of Helsinki practices positive discrimination in school finance so that schools with pupils from disadvantaged backgrounds or from ethnic minorities (based on mother tongue) receive more funding. Despite of these policies, residential segregation has grown in Helsinki during the past 20 years, which may also have affected the extent of school segregation through pupil composition (Bernelius and Vaattovaara and 2016).

With these institutional aspects in mind, it is interesting to study whether parents really perceive that the Finnish schools are of equal quality. Anecdotal evidence and parent interviews reported in e.g. Kosunen (2014) suggest that parents go to some lengths in securing their child's place in a particular school.[6] In this paper, we ask whether this pattern is present more generally by studying whether homebuyers are

[4] For example, using U.K. data Hussain (2015) shows that inspection ratings can provide additional and relevant information on school quality for students and parents.

[5] Moreover, teaching is a highly appreciated profession and education programs are among the most difficult programs to access in Finnish universities. See e.g. OECD (2013) for more details.

[6] Perhaps the most famous example is the story reported in Helsingin Sanomat (21 May 2011), the largest newspaper in Finland, about a couple who divorced on paper so that their child could gain residence within a particular catchment area. Moreover, Kosunen (2014) reports that several families she interviewed had moved to avoid particular schools that they found undesirable based on their knowledge, obtained from various social networks, about the schools.

willing to pay a house price premium in order to send their children to schools which they perceive to be of high quality.

We answer this question using house price and school quality data (standardized math test scores from the 6^{th} grade and the shares of pupils with special needs and foreign-language) for the city of Helsinki. The city is divided into school catchment areas so that parents can secure a place for their child in a particular school by buying a housing unit within the school's catchment area.

Our identification strategy makes use of these catchment area boundaries and is based on the now well-established spatial differencing method (Duranton et al. 2011; Fack and Grenet 2010; Gibbons et al. 2013). We match each transaction in our data near a catchment area boundary to the nearest transaction from the same building type (multi-storey or row house) that lies on the other side of that boundary and then estimate hedonic regression models using the differences between the matched transacted units. The discontinuity in school quality at the catchment area boundaries is fuzzy because pupils can apply to schools other than the one in their catchment area.[7] On the other hand, we use data only for multi-storey and row houses in a dense urban area, which means that the matched units are located very close to each other making this identification strategy particularly appealing. Moreover, it is easier to detect capitalization in dense areas, where housing supply is inelastic (e.g. Brasington 2002, Hilber and Mayer 2009). We also know from earlier research (see e.g. Oikarinen et al. 2015) that of all urban areas in Finland supply is particularly inelastic in Helsinki due to strict zoning regulations and geographical constraints.

Our results can be summarized as follows. The average standardized test scores are capitalized into house prices, while the share of pupils with special needs and the share of foreign-language pupils are not. A one standard deviation increase in the test scores increases prices by roughly 3 percent. This result is robust across a number of specifications. We also report results from a placebo test using fake boundaries and find that the school characteristics have no effect on prices at these fake boundaries. Since the 6^{th} grade is the final year of the first part of elementary schooling, the test score result may reflect either parents' demand for schools effectiveness (or value-added) or pupil composition or a mixture of both. Additional results based on proxies of parents'

[7] This may dilute the relationship between school quality and house prices as shown by Fack and Grenet (2010), Machin and Salvanes (2016), Brunner et al. (2012) and Schwartz et al. (2014) in other contexts.

characteristics suggest that the test score result is driven by parents demand for socio-economically favourable pupil composition, not for school effectiveness.

The results are somewhat surprising and indicate that Finnish parents do perceive clear quality differences among elementary schools, even though school differences in student achievement are low by international comparison. We argue that the surprisingly large effect is at least partly due to inelasticity of housing supply, as we use data from a single and densely populated urban area. We also show that the effect depends on local land supply conditions within the city and is highest in areas with inelastic supply. The finding that parents value the quality of peer composition implies that perceived school quality differences together with a catchment area-based pupil intake can affect residential and school segregation patterns, even when school quality differences are relatively low and school quality information is not publicly disclosed.

The rest of the paper is organized as follows. In Section 2, we explain our empirical strategy. Section 3 describes the data and discusses the details of the school admission system in Helsinki and our school quality measures. Section 4 presents the econometric results. In this section we also discuss the likely mechanisms behind our findings and discuss the size of the effects. Section 5 concludes.

2. Empirical strategy

The starting point of our analysis is the hypothesis that spatial differences in the quality of local public goods are reflected in house prices, thus revealing households' marginal valuation of them (e.g. Black and Machin 2011). Elementary schools are a prominent example because the right to attend a particular school is often tied to residential location. The general problem in estimating the effects of school quality on house prices is that some neighbourhood variables that affect prices are unobservable and may also be correlated with school quality. This leads to endogeneity problems and biased estimates in a simple OLS regression model. However, if access to schools is spatially bounded based on catchment areas, there should be a discrete change (or discontinuity) in school quality at the catchment area boundaries, while other neighbourhood characteristics develop smoothly. In this case, a solution to the omitted variable problem is to concentrate on houses at school catchment area boundaries and use the discrete change in quality for identification.

To show this more formally, we follow Gibbons et al. (2013) and consider the following hedonic regression model:

$$(1) \qquad p = s(l)\beta + x(l)\gamma + g(l) + u,$$

where p refers to the (log) sale price of a housing unit and s to school quality, possibly a vector of school attributes, resources and effectiveness, that a homebuyer can access when residing in location l. The vector x includes observable housing unit attributes, whereas $g(l)$ refers to unobservable neighbourhood attributes (other than school quality) in location l. The last term u represents unobservable unit attributes and errors that we assume to be uncorrelated with s, x and l.

Our interest lies in β, which is the causal effect of school quality on housing prices. A simple OLS estimation of Eq. (1) will produce inconsistent estimates because in general $Cov[s(l), g(l)] \neq 0$. This problem can be solved by using spatial differencing and catchment area boundaries. The spatially differenced model for units in location i and j can be written as

$$(2) \qquad p_i - p_j = \left[s(l_i) - s(l_j)\right]\beta + \left[x(l_i) - x(l_j)\right]\gamma + \left[g(l_i) - g(l_j)\right] + \left[u_i - u_j\right].$$

As shown by Gibbons et al. (2013), choosing i and j to be geographically as close as possible and on opposite sides of a catchment area boundary eliminates the correlation between unobservable neighbourhood attributes and school quality, while maintaining variation in school quality, which identifies the causal effect of interest.[8]

In practice, of course, we do not have enough sales data exactly at the boundaries and we need to include observations from further away. As in the standard regression discontinuity design, this induces a bias-variance trade-off: including more observations from further away from a boundary increases the precision of the estimates, but at the

[8] Instead of spatial differencing, one could use boundary fixed effects (Black 1999, Bayer et al. 2007) where one subtracts the boundary level means from each observation. However, if boundaries are long with only a few observations at each boundary, the fixed effects may not be sufficient to wash away confounding unobservables. See Black and Machin (2011) for a discussion of the merits of different identification strategies.

same time increases the risk of bias because the assumption of constant neighbourhood quality is less and less likely to hold as distance increases (Lee and Lemieux 2010).

As Gibbons et al. (2013) point out, three assumptions need to hold for the spatial differencing strategy to work. First, there has to be variation or a discontinuity in school quality that homebuyers face at the catchment area boundaries. This discontinuity can be fuzzy so that there is a change in the expected school quality at the boundaries. This is the case in our data as a homebuyer can secure a place for his/her child in a particular school by buying a unit within the school's catchment area, but it is possible to apply to other schools as well. This means that the probability of gaining access to the school jumps from something below unity to one. Second, there cannot be discontinuities in neighbourhood quality unrelated to schools exactly at the boundaries. These may arise due to e.g. other boundaries that coincide with catchment areas or to major geographic obstacles, such as railways or waterways. Third, there should be no spatial trends in other neighbourhood characteristics or amenities across boundaries. Possible spatial trends become a problem because we never have enough data exactly at the boundary. Adding more data further away from the boundary increases the risk that neighbourhood amenities differ on average on different sides of a boundary. Discontinuities in other neighbourhood characteristics is a more severe problem than spatial trends because, at least in principle, adding more data near a boundary solves the spatial trend problem, but discontinuities can invalidate the design and bias the results with respect to willingness to pay for school quality regardless of data size. In this case, we could find a discontinuity in prices at the boundaries, but would mistakenly attribute it to variations in school quality.

We argue that we have a particularly good research design and data so that these assumptions are likely to hold at least approximately. First, we use data from a single municipality, which means that other policies, such as local tax rates, stay constant within the area. We also eliminate boundaries that coincide with major geographic obstacles that may induce discontinuities in neighbourhood quality, such as major roads, railways or waterways. Second, we include boundary specific distance to boundary polynomials and weight the observations based on the inverse distance between matched sales. Third, following Black (1999) and Gibbons et al. (2013) we can implement a falsification test based on fake catchment area boundaries. Fourth, we use transaction data only from multi-storey buildings and row houses. Thus, our data come from a dense urban area so that the average distance between matched transactions is short, which considerably mitigates the problem of confounding spatial trends. It should also be easier to detect capitalization in dense areas because housing supply is inelastic (e.g. Brasington, 2002, Hilber and Mayer, 2009, Black and Machin, 2011).

Finally, as noted by Bayer et al. (2007), household sorting may affect whose marginal willingness to pay can be identified from a hedonic regression when households have heterogeneous preferences for school quality. For example, households buying units on the "high quality" side of a boundary are likely to have a systematically higher willingness to pay for school quality than households buying on the "low quality" side. However, Bayer et al. (2007) do find that a hedonic regression produces a good approximation of the mean willingness to pay as long as there are a high number of different quality choices available, which is the case when there are many schools within in a local housing market (50 in our case). See also Bayer and McMillan (2008) for further discussion.

Another consequence of sorting is that it may be difficult to separate the willingness to pay for school quality the willingness to pay for neighbours' quality. This is because the relationship between school quality and house prices may work through its effect on neighbours' quality. Here we follow Gibbons et al. (2013) and use a control variables strategy. We can add an extensive set of close neighbourhood characteristics, measured at a 250 m x 250 m grid level, as control variables and show that including them does not affect our results. We also graphically and statistically test for the presence of discontinuities in neighbourhood characteristics.

While solving the major identification problem, spatial differencing introduces some problems for statistical inference because a particular housing unit may be the closest match for a number of units on the opposite side of the boundary. This induces correlation between all differenced observations that share a match. A simple solution to this problem would be to cluster standard errors at the boundary level, which allows for arbitrary correlation between all observations on either side of a given boundary.[9] However, in our baseline estimations we have only 33 clusters (boundaries) and in some

[9] A boundary, rather than a school, is the correct clustering level also for other reasons, as explained by Fack and Grenet (2010).

of our heterogeneity and robustness analyses the number of clusters is as low as 26. This may not be sufficient for the standard clustering procedure to work reliably (see Bertrand et al. 2004). Therefore, in addition to reporting clustered standard errors, we also report statistical significance based on the cluster generalization of the wild bootstrap suggested by Cameron et al. (2008). As shown by Cameron et al. (2008), the wild bootstrap procedure leads to improved inference when there are few clusters.

3. Finnish school system and data

3.1. School system

In Finland, local governments (municipalities) are responsible for providing primary education. The primary education system consists of a nine-year compulsory comprehensive school starting in the year the child turns seven. There is also a one-year optional pre-school before primary school. Comprehensive school is usually divided into a primary school with grades 1–6 and a lower secondary school with grades 7–9, but in some cases grades 1–9 are taught in the same school (joint comprehensive school).[10] Most of the comprehensive schools are public schools and children usually attend the school closest to where they live. Comprehensive schooling is completely free for the whole age group and includes daily lunches. There are also some private schools in Helsinki, but these schools share the legislation of the public school system and are therefore very similar to public schools.[11] Elite private schools charging sizable student fees do not exist in Finland.

Since the mid 1990's, school choice has, in principle, been free in Finland. However, in practice municipalities are still divided into catchment areas and the municipality guarantees each child living in a catchment area a place in the catchment area's school. Buying a unit within a catchment area of a particular school thus secures a place in that school.

Pupils are also allowed to attend other schools. To do this, parents need to apply for a place in another school and the school may accept the application if there is space

[10] This division is no longer used officially, but many schools still offer grades based on this division.

[11] In Helsinki there are 10 private schools providing education for pupils of grades 1–6. All of these schools are either religious schools or language/international schools. This implies that these schools have specific requirements (e.g. certain religion, language skills) for the choice of pupils.

to accommodate pupils from other catchment areas. When there are limited places available, acceptance depends on whether siblings attend the school, travel time, aptitude tests or in some cases a lottery. This means that in our setup the discontinuity in school quality at the boundaries is fuzzy so that there is a discrete jump in expected school quality. Linking housing sales to school data is straightforward because in Helsinki each housing unit is assigned to one elementary school (grades 1–6). Homebuyers ought to be well informed about the catchment area that a given building belongs to because In Helsinki there is a free internet-based service offered by the city of Helsinki which assigns every address to a specific elementary school.[12]

3.2. School quality measures

What exactly are the right school quality measures or what school characteristics homebuyers are willing to pay for are questions that still remain unanswered (e.g. Rothstein 2006, Black and Machin 2011). In this study, we follow the existing literature and use a number of different measures. More specifically, we use the average standardized math test scores from the 6th grade, the share of foreign-language pupils and the share of pupils with special needs.[13] The school quality measures are from 2008.[14]

The latter two school quality indicators are measured over grades 1–6 and they both aim to capture extreme aspects of pupil composition. Having pupils with special needs or a foreign language in ordinary classes may have negative effects on the learning of others if these pupils need much extra attention and drain teachers' resources.[15] Unfortunately, we do not have data on other pupil or parent characteristics. However, in some additional models we use proxy measures for parents' characteristics.

---

[12] The service can be found here: http://www.hel.fi/palvelukartta/.

[13] This is the only test score that is available to us from the 6th grade.

[14] We also obtained data on school total expenditure. We decided against using these data, because per-pupil expenditures are not comparable across schools that have only grades 1–6 and schools that have grades 1–9.

[15] Those pupils that do not satisfy the learning objectives are considered pupils with special needs. The reasons for special needs usually stem from different problems in the evolution of physical and mental abilities. Foreign-language pupils are pupils that do not have Finnish or Swedish, or an indigenous minority language, e.g., Sami language as their mother tongue. Pupils with special needs and foreign-language pupils receive their education partly in classes with regular pupils and partly in special groups.

As these variables are measured with error and not used in the main models, we discuss them in more detail in Section 4.3.

In our setting, probably the most interesting school quality measure is the average standardized math test score from the 6th grade. The tests were organized by the Finnish National Board of Education, which has monitored the learning results of comprehensive school pupils with the help of national standardized tests since 1998. About 20 percent of schools take part in these tests, but participating schools generally differ across tests. However, in the city of Helsinki all public schools have taken part in the tests. Despite the fact that all public schools participate in the standardized tests, the results are not publicly disclosed in Helsinki or in any cities in Finland, unlike many other countries.[16]

Since the test scores are for the final phase of primary school, they reflect differences in both schools' effectiveness (or value-added) and pupil composition (see e.g. Gibbons et al. 2013). Unfortunately, we do not have any information about the pupils' prior achievement which could be used to construct a value-added measure. We return to this issue after we present our main results.

3.3. Matching across boundaries

In order to estimate the spatially differenced model, we match housing sales on opposite sides of catchment area boundaries based on sale year and building type. For each observation, we find the closest sale from the same year and building type on the opposite side of a catchment area boundary, i.e. we match a row house unit to a row house unit and multi-storey unit to a multi-storey unit (see also Gibbons et al. 2013). In our baseline estimations, the maximum distance between matched units is 400 meters. In robustness analysis, we vary this distance from 400 to 200 meters.

Fig. 1 illustrates the catchment area boundaries in Helsinki. We use only the boundaries where access to grades 7–9 does not change. In addition, we eliminated those boundaries that coincide with major geographic obstacles, such as major roads, railways or waterways, which may cause a discontinuity in neighbourhood quality that is not related to school quality. In Fig. 1, these boundaries are marked with a dashed

[16] Therefore we are indirectly testing if homebuyers have 'unofficial' information about the differences in school quality in Helsinki.

line. The solid lines represent boundaries where access changes both for grades 1–6 and 7–9 or boundaries with major geographic obstacles. Since we do not have school quality information for higher grades, including transactions from these boundaries might lead to biased results (Nguyen-Hoang and Yinger 2011). We return to this issue when we discuss the mechanisms behind our results in Section 4.3. Fig. 1 shows that our data are spread over the whole of Helsinki. The catchment areas are also quite stable in time as only two boundaries were changed slightly from 2008 to 2014.

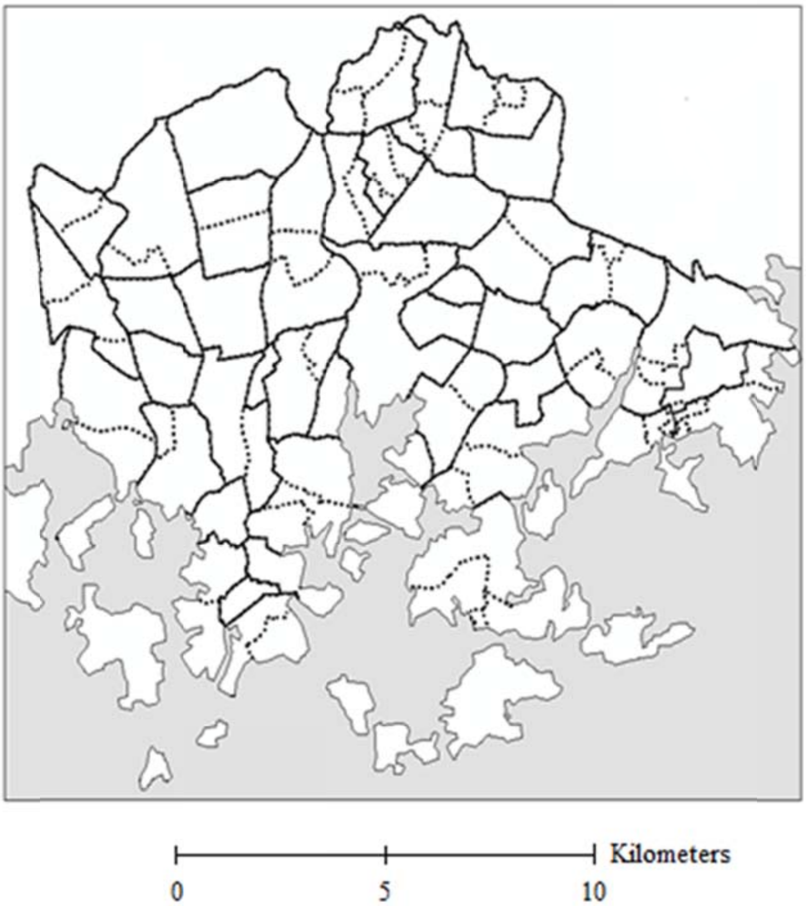

Figure 1. Catchment area boundaries in Helsinki.

Notes: The solid lines represent catchment area boundaries where access to both grades 1–6 and 7–9 changes or boundaries that coincide with geographic obstacles. The dashed lines represent boundaries where access changes only for grades 1–6. The boundaries were obtained from the city of Helsinki.

3.4. Data and descriptive statistics

Our data come from three sources. First, the school-specific characteristics were obtained from the Education Department of the city of Helsinki. Table 1 presents the descriptive statistics of these school characteristics as well as the share of pupils that go to the school in their catchment area. The somewhat large variation in the share of pupils with special needs and foreign language students suggest that the schools are segregated. However, this variation does not seem to translate into particularly large variation in math test scores.[17]

Table 1. Descriptive statistics for schools characteristics, $N = 50$.

	Mean	Std. Dev.	Min	Max
Math test score	32.1	2.92	21	38
% pupils with special needs	0.09	0.07	0	0.36
% foreign-language pupils	0.11	0.10	0	0.44
% pupils going to the school in their catchment area	0.71	0.13	0.36	0.91

Second, we use unit level transaction price data. These data are voluntarily collected by a consortium of Finnish real estate brokers and the dataset is refined and maintained by the VTT Technical Research Centre of Finland. As not all real estate agencies participate in the consortium, the dataset represents a sample (albeit rather large) of the total volume of transactions. The transactions in our data took place in 2008–2012.

Finally, we use Statistics Finland's grid database for 2008. This database is based on 250 x 250 meters grids which contain, in addition to grid coordinates, information on the population structure, education, main type of activity and income, households' stage in life, type of buildings and number of jobs. We use these close neighbourhood data as control variables to assess the robustness of our results.

Fig. 2 illustrates the detail of our close neighbourhood data for a small stretch of a single boundary. For confidentiality reasons, we are not allowed to map the individual transactions. We use only multi-storey and row houses in order to make sure that our matched units of the same building type are close to each other so that we can treat them as if they were in the same neighbourhood.

[17] The maximum score in the math tests was 54 points. The exercises were in algebra, geometry and statistics, which were defined as important on the grounds of the National Curriculum in 2004.

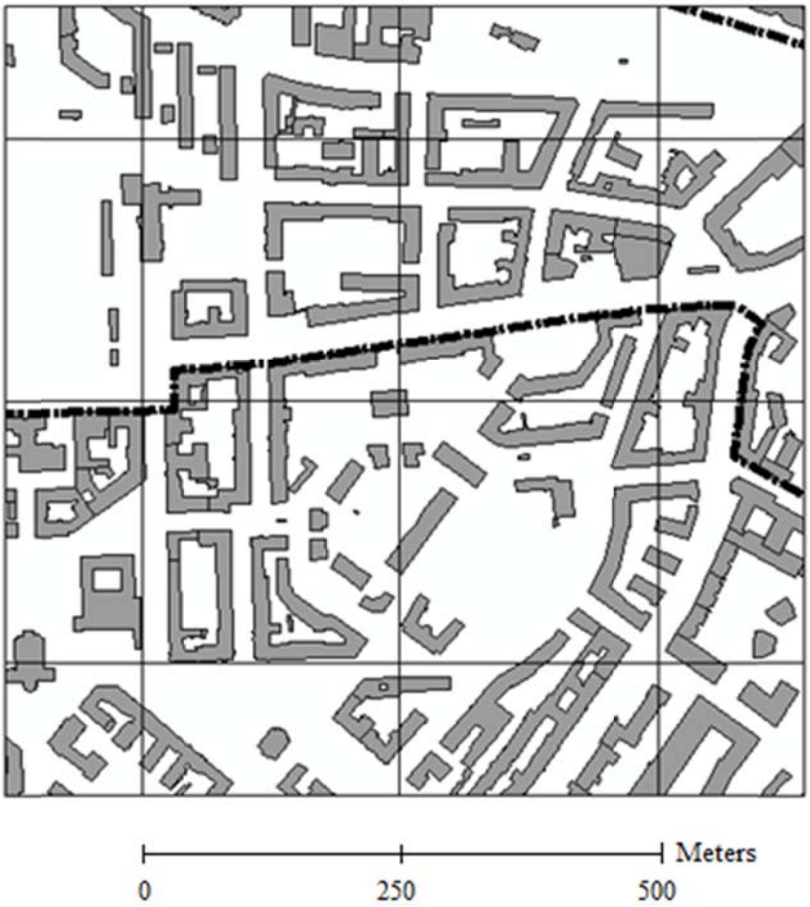

Figure 2. Buildings and close neighbourhood grids at a single catchment area boundary.

Note: The grey polygons represent buildings, the solid lines the 250 m x 250 m grids and the dashed line the catchment area boundary.

The descriptive statistics of the housing units and close neighbourhoods are reported in Table 2. We report the full sample statistics as well as the statistics of the matched sample. In our final matched sample, we focus only on units for which the distance to a similar building type neighbour, located on the other side of a catchment area boundary, is less than 400 meters. The mean distance between matched pairs in this sample is 235 meters. As can be seen from Table 2, the full sample and matched sample are comparable both in terms of dwelling and close neighbourhood characteristics.[18]

The housing characteristics included in the data are the transaction price of the unit, floor area, age, broker's estimate of the unit's condition (used internally by the agency), building type, indicator that the building is situated on a freehold lot (rather

[18] We report the same descriptive statistics for different subsamples according to maximum match distance in Table A1 in the Appendix.

than a city leasehold lot), indicator whether there is an elevator in the building, floor, total number of floors in the building, maintenance charge, distance to CBD and distance to nearest match.[19] In the analysis, we use only units that have at least two rooms (in addition to a kitchen) because smaller units are not suitable for families with children.

[19] In Finland owner-occupied units in multi-unit buildings are part of cooperatives that are incorporated as limited liability companies. Buying a housing unit from a building means that one buys the shares of the company on the open market. The company owns all the common facilities (and usually the lot) and charges shareholders a maintenance charge for common costs.

Table 2. Descriptive statistics on dwellings and close neighbourhoods.

	Full sample		Matched sample (< 400 m)	
	Mean	Std. Dev.	Mean	Std. Dev.
Number of observations	14,061		3,852	
Housing unit:				
Price (€)	251,244	152,787	255,211	162,428
Floor area (m^2)	68.6	27.4	67.5	25.1
Age (years)	45.0	33.2	44.1	31.1
Condition (broker estimate):				
Good (0/1)	0.65	0.48	0.65	0.48
Satisfactory (0/1)	0.32	0.47	0.31	0.46
Poor (0/1)	0.03	0.18	0.04	0.18
Building type:				
Row (0/1)	0.09	0.29	0.06	0.25
Multi-storey (0/1)	0.91	0.29	0.94	0.25
Own lot (0/1)	0.71	0.45	0.80	0.40
Elevator (0/1)	0.56	0.50	0.61	0.49
Floor level	2.95	1.78	3.10	1.78
Total number of floors	4.57	2.33	4.83	2.22
Maintenance charge (€/m^2/month)	3.26	1.07	3.28	1.20
Road distance to CBD (km)	5.94	3.97	6.28	4.55
Distance to match (km)	0.45	0.26	0.23	0.09
Close neighbourhood (250 m x 250 m):				
Home ownership rate	0.51	0.20	0.51	0.19
Mean income (€)	32,231	12,061	32,583	15,183
% college degree adults	0.30	0.12	0.28	0.11
Unemployment rate	0.06	0.04	0.06	0.03
% retired households	0.22	0.10	0.21	0.10
% households with children	0.16	0.10	0.15	0.10
Number of service jobs per capita	0.44	1.20	0.50	1.34
Number of buildings	21.0	12.8	23.5	14.8
Mean floor area of units (m^2)	61.6	17.2	59.3	16.2
Population	734	550	905	673
% foreign language residents	0.09	0.06	0.09	0.05

4. Results

4.1. Main results

This section presents our main results. We start by reporting graphical evidence concerning discontinuities in test scores and prices in Fig. 3. We follow the regression approach used in Gibbons et al. (2013) so that the figures are obtained as predictions from regressions of the cross-boundary difference in the relevant variable on distance-quintile dummies from the boundary on each side (negative and positive). The maximum distance to boundary on either side is 726 meters and each outcome is standardized within this sample. The graphs are restricted to 500 meters on each side. The dashed lines represent the 95% confidence intervals based on clustering at the boundary level. We also report the *p*-value of an *F*-test, which tests whether the differences are equal at the boundary.

The left panel of Fig. 3 shows that there is sizable jump in test scores at the catchment area boundaries on average. Of course, this jump in the figure occurs by construction because we have defined the negative and positive sides of Fig. 3 based on the test scores. The discontinuity in test scores simply illustrates where our identification arises from and that there is identifying variation at the boundaries. From the right panel of Fig. 3, we see that there is a discontinuity also in prices at the boundary. This discontinuity occurs in the raw data as we are only controlling for the date of sale (quarter-year), not any other unit characteristic when constructing this figure. In Fig. A1 in the Appendix, we report discontinuity graphs for a number of close neighbourhood characteristics that measure neighbours' quality. It seems that households are not systematically sorted across catchment area boundaries with respect to the attributes we can observe.

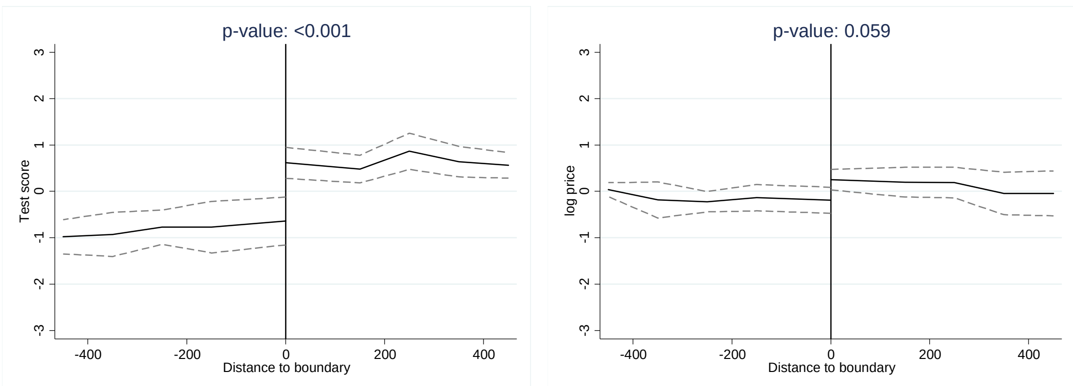

Figure 3. Discontinuities in test scores and log prices.

Notes: Distance to boundary is measured in meters and negative (positive) distance indicates the side of the boundary with a lower (higher) test score. The *p*-values refer to *F*-tests testing whether the differences are equal at the boundary. The confidence intervals and the *F*-test account for clustering at boundary level.

Next we turn to regression results where we can fully exploit the spatial differencing technique, which allows for more than one school characteristic at a time. Table 3 presents six model specifications. In Panel A, we include only the standardized test score, as this is the most often used measure of school quality in prior literature. In Panel B, we add the share of pupils with special needs and the share of foreign language pupils.

In both panels a richer set of control variables is added as we move across the columns. First, we only control for quarter-year sale dummies. Second, we control for housing unit characteristics to capture systematic differences between matched units and to increase the precision of our estimates. In addition to unit-level heterogeneity, we need to worry about possible spatial trends in prices due to local amenities and possible household sorting. Thus, in the third column we add boundary specific distance to boundary polynomials (cubics) and we also weight the observations using inverse distance weights. These weights are based on distance between the matched sales so that sales closer to each other receive more weight in the estimation. Finally, in the fourth column, we add close neighbourhood controls.[20] We report both standard errors clustered at the boundary level and, due to the small number of clusters (33), also *p*-values from a wild bootstrap procedure.

According to Table 3, the average standardized math test score of a school has a sizable positive effect on prices, both when included alone and when other quality measures are included as well. In our preferred specifications with unit characteristics, a one standard deviation increase in test scores increases prices by roughly 3 or 3.5 percent. This result is statistically significant both when using clustered standard errors and also when using the wild bootstrap. Reassuringly, the result is robust both to adding distance to the boundary polynomials and weighting, and to close neighbourhood controls.

Table 3 also shows that the other school quality measures are not capitalized into prices. The share of foreign language pupils obtains a negative coefficient, as one might expect based on earlier literature, but it is borderline statistically significant in only one specification (with wild bootstrap), where we do not control for distance to boundary or

[20] Note that some of the neighborhood variables may be endogenous (Bayer at al. 2007 and Gibbons et al. 2013), but nonetheless it is of value to see whether the results are robust to their inclusion.

close neighbourhood characteristics. The results for these variables remain the same when we omit the test score from the regressions.

Table 3. The effect of school quality measures using cross-boundary differences.

	Panel A: Test score			
	(1)	(2)	(3)	(4)
Math test score	0.077**	0.034***	0.036***	0.030***
	[0.031]	[0.010]	[0.008]	[0.006]
	(0.076)	(0.080)	(0.002)	(0.012)
N	3852	3852	3852	3852
R^2	0.09	0.84	0.86	0.87
	Panel B: Test score and pupil composition			
	(5)	(6)	(7)	(8)
Math test score	0.079**	0.032***	0.041***	0.033***
	[0.037]	[0.009]	[0.013]	[0.011]
	(0.228)	(0.016)	(0.030)	(0.090)
% special needs pupils	0.143	0.077	0.230	0.121
	[0.429]	[0.146]	[0.176]	[0.167]
	(0.830)	(0.707)	(0.424)	(0.633)
% foreign language pupils	-0.100	-0.183**	-0.166	-0.095
	[0.295]	[0.076]	[0.108]	[0.084]
	(0.729)	(0.058)	(0.214)	(0.358)
N	3852	3852	3852	3852
R^2	0.09	0.84	0.86	0.87
Unit characteristics	no	yes	yes	yes
Inverse distance weights and boundary distance cubics	no	no	yes	yes
Close neighbourhood characteristics	no	no	no	yes

Notes: The dependent variable is the log sale price of the housing unit. The table reports results for spatially differenced models. The data include only observations with two or more rooms and for boundaries where access to grades 7–9 does not change. The maximum distance between matched units is 400 meters. The standard errors are clustered at the school boundary level and are reported in brackets. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level, respectively, based on the clustered standard errors. *P*-values based on a wild bootstrap procedure with 999 repetitions are reported in parentheses. Unit and close neighbourhood characteristics are reported in Table 2. All the models include quarter-year of sale dummies.

4.2. Validity checks

The main identifying assumption of our econometric approach is that other neighbourhood characteristics develop smoothly across catchment area boundaries. In addition to the evidence presented in Fig. A1, we have run a number of robustness and validity checks, which are reported in Table 4.

First, in columns (1) and (2) of Table 4, we present results using our main specification, but where we further narrow the maximum distance between matched pairs first to 300 and then to 200 meters. The mean distances between matched units in these samples are 185 and 134 meters, respectively. Again, the results are roughly the same as when using the sample that allows for a longer maximum distance. In fact, the effect is slightly larger, but at the same time the estimates become more imprecise as the standard errors increase due to smaller sample size.

The second issue is that some catchment area boundaries may coincide with some other well-known residential area division, such as zip codes. If some zip codes are particularly prominent, households may value such addresses, which may even be reflected as a discontinuity in prices at these zip code boundaries. In column (3) of Table 4, we present the results of a regression where we omitted catchment area boundaries that coincide with zip code boundaries. The results remain the same for this sub-sample, indicating that the results are not driven by changes in these residential area boundaries.

Finally, we implement a falsification test based on fake catchment area boundaries following Black (1999) and Gibbons et al. (2013). We do this by estimating a spatially differenced model between sales that are actually within the same catchment area. We achieve this by geographically translating the actual boundaries in a south-west direction and maintaining the original school characteristics for these fake catchment areas.[21] As Gibbons et al. (2013) point out, finding of a positive association between school quality and housing prices along the fake boundaries would falsify the claim that price effects are causally linked to cross-boundary school quality discontinuities. The fake boundary results are reported in panel (4) of Table 4. Reassuringly, the school characteristics have no effect on prices at these fake boundaries as all the coefficients are very close to zero. This is strong evidence in favor of a causal interpretation of our results using actual catchment area boundaries.

[21] The boundaries were moved by approximately 1 km. Fig. A3 in the Appendix illustrates the original and the fake boundaries. We include all boundaries in the fake boundary analysis.

Table 4. Validity checks: varying match distance, excluding zip code boundaries and using fake boundaries.

	(1) Maximum distance < 300m	(2) Maximum distance < 200m	(3) Within same zip-code sample	(4) Fake boundary sample
Math test score	0.043**	0.081***	0.036***	0.004
	[0.018]	[0.024]	[0.011]	[0.008]
	(0.026)	(<0.001)	(<0.001)	(0.713)
% special needs pupils	0.185	0.550*	0.139	0.051
	[0.199]	[0.298]	[0.223]	[0.116]
	(0.517)	(0.112)	(0.639)	(0.783)
% foreign language pupils	-0.129	-0.018	-0.097	0.024
	[0.114]	[0.127]	[0.086]	[0.090]
	(0.412)	(0.913)	(0.376)	(0.865)
N	2770	1515	2725	4717
R^2	0.87	0.89	0.86	0.85
Unit characteristics	yes	yes	yes	yes
Inverse distance weights and boundary distance cubics	yes	yes	yes	yes

Notes: The dependent variable is the log sale price of the housing unit. The table reports results for spatially differenced models. The data include only observations with two or more rooms and for boundaries where access to grades 7–9 does not change. In columns (3) and (4) the maximum distance between matched units is 400 meters. The standard errors are clustered at the school boundary level and are reported in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level, respectively, based on the clustered standard errors. *P*-values based on a wild bootstrap procedure with 999 repetitions are reported in parentheses. The unit characteristics are reported in Table 2. All the models include quarter-year of sale dummies.

We also checked whether the results are robust with respect to the timing of sales. In Tables 3 and 4, we used transactions for 2008 to 2012, but the school quality measures are for 2008. Of course, it is plausible to assume that school quality differences change slowly and using data for later years is not a major problem. Nonetheless, in Table A2 in the Appendix we report results where we narrow this time window. Again, the results are roughly the same as those obtained using more data.

In addition to validity checks of the research design, the size of the effect merits further discussion. I.e., why do we find surprisingly large capitalization effects for primary schools in Helsinki? There are two issues that are of particular importance. First, as Black and Machin (2011) emphasize in their review, the size of the capitalization effect depends on the housing supply elasticity so that there is an inverse relationship between housing supply elasticity and capitalization effects. Brasington

(2002) and Hilber and Mayer (2009) among others provide empirical evidence in favour of this argument by showing that capitalization of is weaker in areas where housing supply elasticity is greater (Hilber 2017 provides a recent survey).

In contrast to many previous studies (e.g. Black 1999, Gibbons et al. 2013), we concentrate on a dense urban area using data mostly from multi-story buildings, where housing supply is inelastic. For example, Oikarinen et al. (2015) find that of all urban areas in Finland supply is particularly inelastic in Helsinki. They attribute this both to the fact that Helsinki has strict zoning regulations and that there are geographic obstacles to supply because downtown Helsinki is situated at a peninsula as can also be seen from Fig. 1 (see also Saiz 2010). Since many previous studies use transaction data at the boundaries of urban areas (boundaries of local authorities in the case of Gibbons et al. 2013), the estimates from these studies are not directly comparable to studies that use data only within a single urban area, like ours.

We can also indirectly test this argument as land availability varies also across neighbourhoods. In Table 5, we present results for subsamples where supply elasticity varies. Unfortunately, we do not have direct measures of developable land so we need to resort to proxies. First, we have divided our data into 3 equal-sized groups based the units' distance to CBD. As is the case with most cities, the centre parts of the city are the oldest and it is often difficult build more housing in these old parts of the city. Also as the city centre and neighbourhoods close to it are surrounded by sea (Fig. 1), there are strong reasons to expect supply elasticity to be lowest close to the centre. The units situated closest to the CBD are in the first tertile and so on. Our second proxy for supply elasticity is the units' age. In this case, the oldest units are in the first tertile and the most recently build units are in the third tertile. The idea is that housing supply elasticity should be higher in the sample with more recently build units, whereas it should be quite inelastic in the sample consisting of old units. After all, building has not taken place recently within this sample of older units.

In panel A of Table 5, we report the results for the subsamples based on distance to CBD and in Panel B for the subsamples based on age. We use only the test scores in these regressions, because the number of boundaries is quite low in some of the subsamples. For example, in the first (the closest subsample to CBD) distance tertile there are observations only from three different boundaries (see Fig. 1). Including all three school variables would result in major multicollinearity issues.[22]

In both panels the results are as expected. The capitalization effect is highest in the subsample where we expect the supply to be most inelastic. This is especially evident when using the subsample division based on distance. In the subsample of units closest to the CBD, the test score effect is 4.5 percent (column (1)), whereas it is fairly close to zero in the subsample farthest away from CBD where it is easier to build new housing (column (3)). The same pattern, although not as clearly, can be seen with respect to age in columns (4)–(6).[23]

[22] We also estimated models where we created the subsamples so that each subsample would include the same number of boundaries. The results from these models are quite similar to those reported in Table 5.
[23] We also divided the data into two and four-equal sized subsamples and the same pattern was evident using these subsamples.

Table 5. Effects of school quality according to supply elasticity.

	Panel A: Subsamples based on distance to CBD		
	(1)	(2)	(3)
	I tertile	II tertile	III tertile
Math test score	0.045***	0.014**	0.008
	[0.004]	[0.005]	[0.012]
	(<0.001)	(<0.001)	(0.498)
N	1298	1277	1277
R^2	0.87	0.83	0.90
	Panel B: Subsamples based on age		
	(4)	(5)	(6)
	I tertile	II tertile	III tertile
Math test score	0.058***	0.020**	0.034***
	[0.011]	[0.008]	[0.007]
	(<0.001)	(0.080)	(<0.001)
N	1282	1281	1289
R^2	0.85	0.83	0.92
Unit characteristics	yes	yes	yes
Inverse distance weights and boundary distance cubics	yes	yes	yes

Notes: The dependent variable is the log sale price of the housing unit. The table reports results for spatially differenced models. The data include only observations with two or more rooms and for boundaries where access to grades 7–9 does not change. The maximum distance between matched units is 400 meters. The standard errors are clustered at the school boundary level and are reported in brackets. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level, respectively, based on the clustered standard errors. *P*-values based on a wild bootstrap procedure with 999 repetitions are reported in parentheses. Unit characteristics are reported in Table 2. All the models include quarter-year of sale dummies.

The second issue related to the size of the effect is how parents gather information on school quality when quality measures, such as test scores, are not disclosed to the public. Of course in general, capitalization of school quality does not require that parents would have exact information on test scores or schools quality differences. Moreover, several previous studies based on the boundary discontinuity design have found significant capitalization effects, even when parents had no or only limited information on test scores or other school performance measures (e.g. Black 1999, Davidoff and Leigh 2008, Machin and Salvanes 2016).

Furthermore, a recent qualitative study by Kosunen (2014) investigated parental preferences in school choices in Finland, concentrating on families in the Helsinki metropolitan area (the city of Espoo). The study used data from parental interviews (n =

96) that asked how parents select schools for their children and what kind of information they use in the school choice process. Kosunen (2014) reports that several parents admitted that they had moved to another neighbourhood in order to avoid particular undesirable schools. According to the interview results, parents used various social networks to obtain and share information about the reputation of different schools and about the strategies to avoid certain undesirable schools. Of course the sample size in this study is small, but it nevertheless points to the same direction as our econometric results.

4.3. Discussion of likely mechanisms

The overall picture that emerges from Tables 3–5 is that school quality (broadly defined) is capitalized into house prices in Helsinki. However, it is not clear whether Helsinki parents value school effectiveness in producing value-added or whether they value different aspects of pupil composition, such as good peers. The reason why we cannot make this distinction is that we do not have data on value-added, and the standardized tests that we use were taken at the end of the first stage of elementary schooling. This means that the test scores are a mixture of both school effectiveness and pupil composition.

Of course, we did control for the percentage of pupils with special needs and foreign language, which reflect pupil composition to a certain extent. However, these measures may be more related to extreme situations like classroom disturbance and how much these types of pupils drain teachers' time and other school resources. They do not measure the average or overall pupil composition that parents expect their child to be exposed to when attending a given school.

Previous studies have used parental background to capture composition effects (e.g. Downes and Zabel 2002, Clapp et al. 2008, Brasington and Haurin 2009, Fack and Grenet 2010, Gibbons et al. 2013). In order to follow these studies, we calculated mean income and the share of highly educated residents (% with at least a college degree) within each school catchment area using the grid database. Unfortunately, these are biased measures of parents' characteristics, because they are based on all residents within a catchment area, not just those with school-aged children and because some parents send their children to schools in other catchment areas (see Table 1). For these reasons, these variables should be seen as proxies and one should be careful when interpreting the magnitude of the coefficients of these variables.

Table 6 presents results where these two variables are added into our main model specification. In column (1), we add only the mean income, in column (2) the education level, and finally, in column (3) both. We learn two things from Table 6. First, the coefficients for the two proxy variables are as expected, when added individually into the model indicating that they are meaningful proxies for parental background.[24] Second, and more importantly, once we add the mean income in the catchment area the test score coefficient diminishes considerably. This suggests that a large part of the test score effect can be explained by pupil composition.[25]

[24] In column (3), where we control both mean income and education level, the estimate for the latter is negative and significant. Even though this is somewhat surprising, one should be careful with the interpretation of the coefficients of income and education in column (3). The partial effect of education is generally different when we condition and when we do not condition on income.

[25] We also estimated these models using the boundaries where access changes both for grades 1–6 and 7–9. The results for these models were similar, but in some specifications the test score effect was even closer to zero. The results for alternative models including all boundaries are available from the authors by request.

Table 6. Effects of school quality and catchment area measures.

	(1)	(2)	(3)
Math test score	0.017	0.034**	0.026*
	[0.014]	[0.015]	[0.014]
	(0.184)	(0.086)	(0.048)
% special needs pupils	0.054	0.22	0.033
	[0.169]	[0.169]	[0.186]
	(0.865)	(0.426)	(0.917)
% foreign language pupils	-0.019	-0.114	-0.091
	[0.111]	[0.131]	[0.121]
	(0.917)	(0.480)	(0.563)
Catchment area income	0.039***		0.049**
	[0.014]		[0.019]
	(0.010)		(0.042)
Catchment area education		0.189	-0.392
		[0.295]	[0.303]
		(0.597)	(0.254)
N	3852	3852	3852
R^2	0.86	0.86	0.86
Unit characteristics	yes	yes	yes
Inverse distance weights and boundary distance cubics	yes	yes	yes

Notes: The dependent variable is the log sale price of the housing unit. The table reports results for spatially differenced models. The data include only observations with two or more rooms and for boundaries where access to grades 7 through 9 does not change. The maximum distance between matched units is 400 meters. The standard errors are clustered at the school boundary level and are reported in brackets. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level, respectively, based on the clustered standard errors. *P*-values based on a wild bootstrap procedure with 999 repetitions are reported in parentheses. Unit characteristics are reported in Table 2. All the models include quarter-year of sale dummies.

Of course, the above results do not completely rule out the possibility that parents are also paying for school effectiveness. However, there are some additional reasons which make this mechanism unlikely in the Finnish case. First, it is very difficult for parents to obtain information on value-added or school effectiveness. To our knowledge, school-level value-added measures have not even been estimated for either primary or lower secondary schools in Finland. It can also be argued that it is unlikely that the differences in value-added are considerable among primary schools in Helsinki. Given that the share of qualified teachers (i.e. teachers with a Master's degree in education) is very high in Finnish schools overall, and especially in the Helsinki metropolitan area, one would expect differences in effectiveness to be much smaller than in many other countries. Given small differences between schools in Finland, this means that it is very challenging if not impossible for parents to identify the most effective schools.

Second, the survey evidence in Kosunen (2014) also points to this direction. Kosunen (2014) reports that several parents considered peer-group composition important and, in particular, the children from middle- and upper-class families were considered a desirable peer group. Given our additional results and the discussion above, we think that it is likely that most of the detected price response to test scores is related to pupil composition.

5. Conclusions

In this paper, we use hedonic regression techniques with a boundary discontinuity research design to study whether school quality differences are capitalized into house prices in Helsinki, the capital city of Finland. The Finnish case is of particular interest because by international comparison Finnish pupil achievement is high and school quality differences are among the lowest in the world.

We find that, even in this environment, school quality differences are capitalized into house prices. More precisely, a one standard deviation increase in standardized test scores increases prices by roughly 3 percent. The magnitude of this effect is comparable to countries where school quality differences are much larger, such as the U.K. and the U.S. We argue that this surprisingly large effect is at least partly explained by the inelasticity of housing supply, as we use data from a densely populated urban area, where adding new buildings is difficult. We also show that the effect depends on local land supply conditions within the city and is highest in areas with inelastic supply. Additional results based on proxies of parents' characteristics suggest that this result is driven by parents demand for socio-economically favourable pupil composition, not for school effectiveness.

The results indicate that Finnish parents do perceive clear quality differences among elementary schools, even though school differences in student achievement are low by international comparison. This also suggests that residential-based school assignment can lead to residential segregation even in an environment where differences in school effectiveness are low and overall performance high, and where school quality measures are not publicly disclosed.

References

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

Bayer, Patrick and Robert McMillan. 2008. “Distinguishing Racial Preferences in the Housing Market: Theory and Evidence.” In Andrea Baranzini, José Ramirez Caroline Schaerer and Philippe Thalmann (eds.) *Hedonic Methods in Housing Markets*. Springer.

Bernelius , Venla and Mari Vaattovaara. 2016. ”Choice and Segregation in the “Most Egalitarian” Schools: Cumulative Decline in the Urban Schools and Neighbourhoods of Helsinki, Finland.” *Urban Studies* 53(15): 3155–3171.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. “How much should we trust differences-in-differences estimates?” *Quarterly Journal of Economics* 119(1): 249–275.

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

Black, Sandra E. and Stephen Machin. 2011. “Housing Valuations of School Performance” in Eric A. Hanushek, Stephen Machin and Ludger Woessmann (eds.) *Handbook of the Economics of Education*, Vol. 3. North-Holland: Amsterdam.

Brasington, David. 2002. “Edge Versus Center: Finding Common Ground in the Capitalization Debate.” *Journal of Urban Economics* 52(3): 524–541.

Brasington, David and Donald R. Haurin. 2009. “Parents, Peers, or School Inputs: Which Components of School Outcomes are Capitalized into House Prices.” *Regional Science and Urban Economics* 39(5): 523–529.

Brunner, Eric J., Cho, Sung-Woo and Reback, Randall 2012. “Mobility, Housing Markets, and Schools: Estimating the Effects of Inter-District Choice Programs.” *Journal of Public Economics* 96(7): 604–614.

Cameron, A. Colin, Jonah G. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3): 414–427.

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

Clapp, John M., Anupam Nanda and Stephen L. Ross. 2008. “Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values.” *Journal of Urban Economics* 63(2): 451–466.

Davidoff, Ian and Andrew Leigh. 2008. “How Much do Public Schools Really Cost? Estimating the Relationship between House Prices and School Quality” *Economic Record* 84(265), 193–206.

Dhallman, Hanna and Katja Vilkama. 2009. “Housing policy and the ethnic mix in Helsinki, Finland: perceptions of city officials and Somali immigrants.” *Journal of Housing and the Built Environment* 24(4): 423–439.

Downes, Thomas A. and Jeffrey E. Zabel. 2002. “The Impact of School Characteristics on House Prices: Chicago 1987–1991.” *Journal of Urban Economics* 52(1): 1–25.

Duranton, Gilles; Laurent Gobillon and Henry G. Overman. 2011. “Assessing the Effects of Local Taxation using Microgeographic Data.” *Economic Journal* 121(555): 1017–1046.

Eerola, Essi and Tuukka Saarimaa. 2017. ”Delivering Affordable Housing and Neighborhood Quality: A Comparison of Place- and Tenant-Based Programs.” *Journal of Housing Economics* forthcoming.

Fack, Gabrielle and Julien Grenet. 2010. “When do Better Schools Raise Housing Prices? Evidence from Paris Public and Private Schools.” *Journal of Public Economics* 94(1–2): 59–77.

Fiva, Jon H. and Lars J. Kirkebøen. 2011. “Information Shocks and the Dynamics of the Housing Market.” *Scandinavian Journal of Economics* 113(3): 525–552.

Gibbons, Stephen; Stephen Machin and Olmo Silva. 2013. “Valuing School Quality Using Boundary Discontinuities.” *Journal of Urban Economics* 75(1): 15–28.

Hilber, Christian A.L. 2017. “The economic implications of house price capitalization: A synthesis.” *Real Estate Economics* 45(2): 301–339.

Hilber, Christian A.L. and Christopher Mayer. 2009 "Why Do Households without Children Support Local Public Schools? Linking House Price Capitalization to School Spending." *Journal of Urban Economics* 65(1): 74–90.

Hussain, Iftikhar. 2015. "Subjective Performance Evaluation in the Public Sector: Evidence From School Inspections." *Journal of Human Resources* 50(1): 189–221.

Kosunen, Sonja 2014. "Reputation and Parental Logics of Action in Local School Choice Space in Finland." *Journal of Education Policy* 29(4): 443–466.

Lee. David S. and Thomas Lemieux 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281–355.

Machin, Stephen and Kjell G. Salvanes. 2016. "Valuing School Quality via a School Choice Reform." *Scandinavian Journal of Economics* 118(1), 3–24.

Nguyen-Hoang, Phuong and John Yinger 2011. "The Capitalization of School Quality into House Values: A Review." *Journal of Housing Economics* 20(1): 20–48.

OECD. 2011. "Strong Performers and Successful Reformers in Education – Lessons from PISA for the United States." OECD Publishing. http://dx.doi.org/10.1787/9789264096660-en

OECD. 2013. "Education Policy Outlook: Finland." OECD Publishing.'

Oikarinen, Elias, Risto Peltola and Eero Valtonen. 2015. "Regional Variation in the Elasticity of Supply of Housing, and Its Determinants: The Case of a Small Sparsely Populated Country." *Regional Science and Urban Economics* 50: 18–30

Rothstein, Jesse M. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review* 96(4): 1333–1349.

Sahlberg, Pasi. 2011. *Finnish Lessons*. Columbia University: New York.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125 (3): 1253–1296.

Schwartz, Amy Ellen, Ioan Voicu and Keren Mertens Horn. 2014. "Do Choice Schools Break the Link Between Public Schools and Property Values? Evidence from House Prices in New York City." *Regional Science and Urban Economics* 49: 1–10.

Appendix. Additional figures and tables.

Table A1. Descriptive statistics for dwelling and close neighbourhood characteristics, alternative sub-samples based on maximum match distance.

	Matched sample (< 300 m)		Matched sample (< 200 m)	
	Mean	Std. Dev.	Mean	Std. Dev.
Number of observations	2,770		1,515	
Housing unit:				
Price (€)	258,200	162,961	257,108	154,744
Floor area (m^2)	67.1	25.3	66.2	25.2
Age (years)	41.8	31.7	41.0	31.1
Condition (broker estimate):				
Good (0/1)	0.67	0.47	0.68	0.47
Satisfactory (0/1)	0.30	0.46	0.29	0.45
Poor (0/1)	0.04	0.19	0.04	0.19
Building type:				
Row (0/1)	0.06	0.24	0.05	0.21
Multi-storey (0/1)	0.94	0.24	0.95	0.21
Own lot (0/1)	0.82	0.39	0.84	0.37
Elevator (0/1)	0.66	0.47	0.72	0.45
Floor level	3.18	1.81	3.29	1.89
Total number of floors	4.92	2.29	5.07	2.33
Maintenance charge (€/m2/month)	3.31	1.27	3.38	1.47
Road distance to CBD (km)	6.18	4.62	5.83	4.60
Distance to match (km)	0.18	0.07	0.13	0.04
Close neighbourhood (250 m x 250 m):				
Home ownership rate	0.50	0.20	0.48	0.20
Mean income (€)	32,199	15,906	31,555	15,443
% college degree adults	0.28	0.11	0.28	0.10
Unemployment rate	0.06	0.04	0.07	0.03
% retired households	0.21	0.10	0.19	0.09
% households with children	0.14	0.10	0.14	0.09
Number of service jobs per capita	0.52	1.25	0.38	0.65
Number of buildings	23.2	14.4	24.2	12.4
Mean floor area of units (m^2)	58.5	16.0	56.3	14.9
Population	940	683	1051	701
% foreign language residents	0.10	0.05	0.09	0.05

Table A2. Additional robustness checks with respect to sale year.

	(1)	(2)	(3)	(4)
	Baseline	sale year < 2012	sale year < 2011	sale year < 2010
Math test score	0.041***	0.036***	0.038***	0.043***
	[0.013]	[0.013]	[0.013]	[0.015]
	(0.030)	(0.022)	(<0.001)	(<0.001)
% special needs pupils	0.230	0.189	0.159	0.154
	[0.176]	[0.196]	[0.194]	[0.232]
	(0.424)	(0.507)	(0.577)	(0.579)
% foreign language pupils	-0.166	-0.216	-0.192	-0.183
	[0.108]	[0.134]	[0.129]	[0.148]
	(0.214)	(0.228)	(0.258)	(0.360)
N	3852	3100	2273	1403
R^2	0.86	0.87	0.87	0.87
Unit characteristics	yes	yes	yes	yes
Inverse distance weights and boundary distance cubics	yes	yes	yes	yes

Notes: The dependent variable is the log sale price of the housing unit. The table reports results for spatially differenced models. The data include only observations with two or more rooms and for boundaries where access to grades 7–9 does not change. The maximum distance between matched units is 400 meters. The standard errors are clustered at the school boundary level and are reported in brackets. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level, respectively, based on the clustered standard errors. *P*-values based on a wild bootstrap procedure with 999 repetitions are reported in parentheses. The unit characteristics are reported in Table 2.

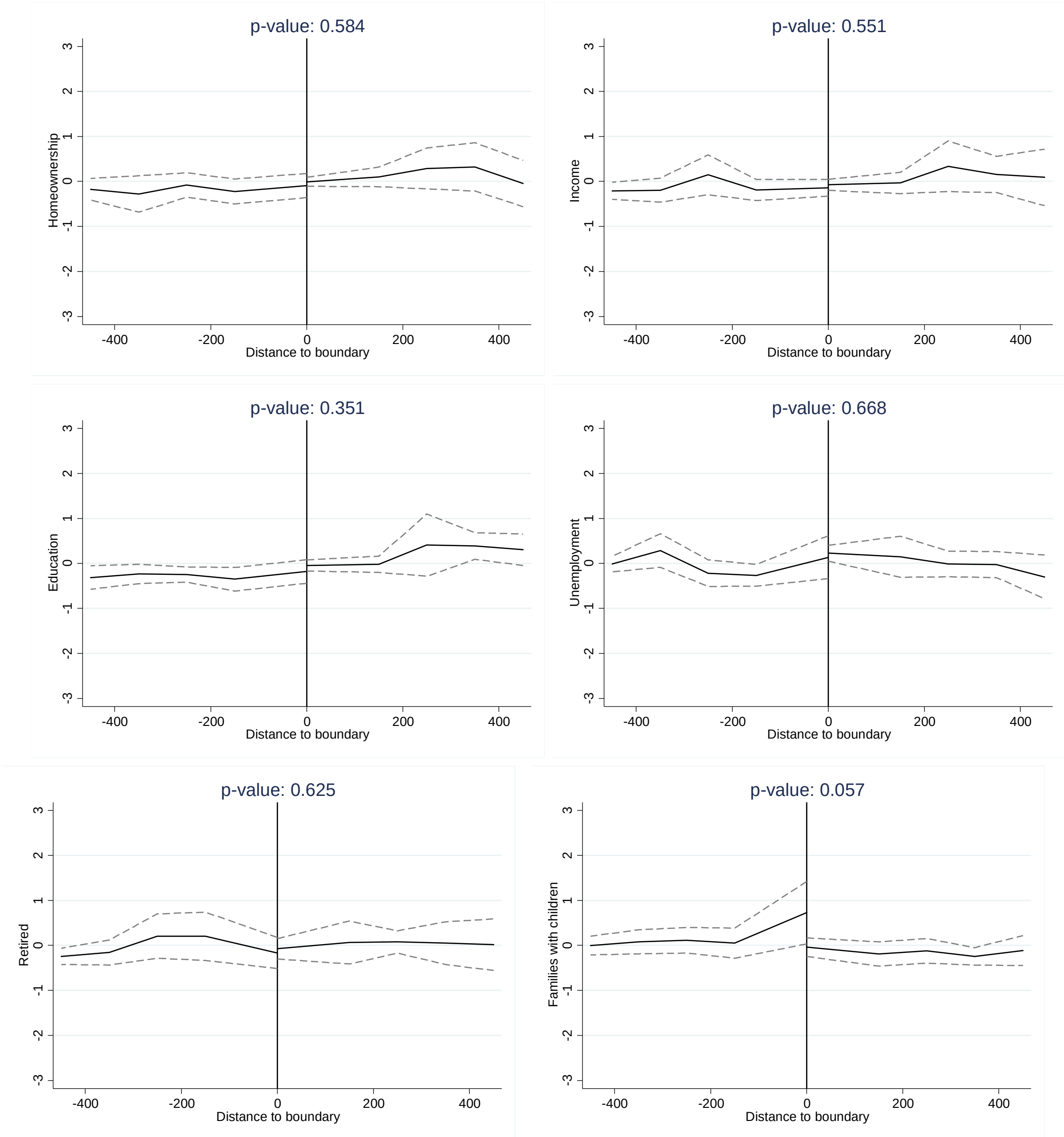

Figure A1. Discontinuities in close neighbourhood characteristics.

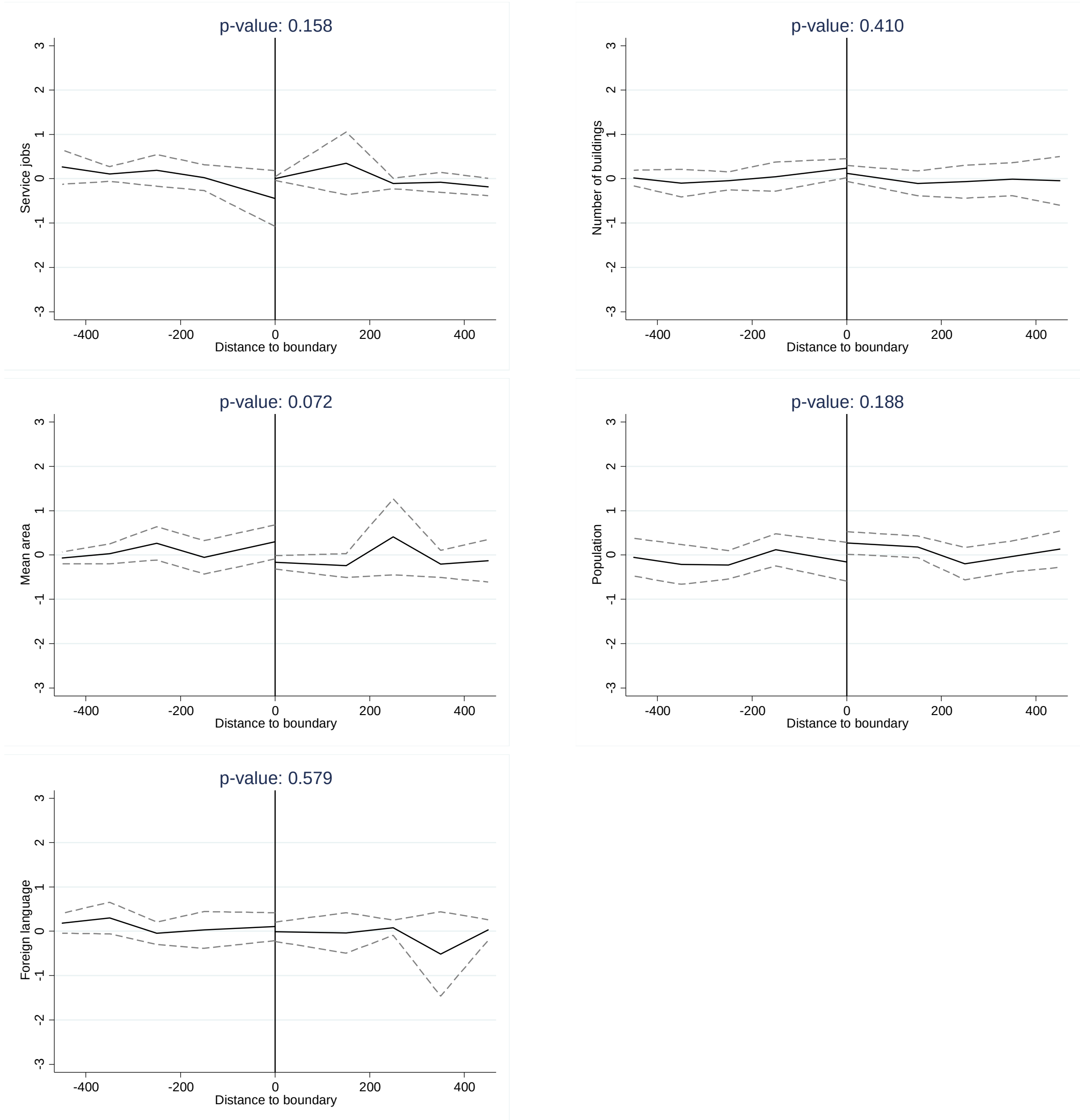

Figure A1. Continued.

Notes: Distance to boundary is measured in meters and negative (positive) distance indicates the side of the boundary with a lower (higher) test score. The *p*-values refer to *F*-tests testing whether the differences are equal at the boundary. The confidence intervals and the *F*-test account for clustering at boundary level.

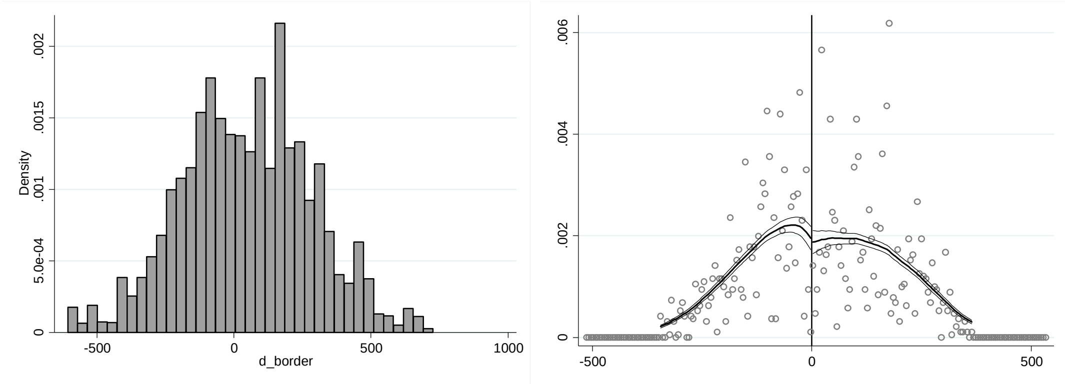

Figure A2. Histogram of number of units according to distance to boundary and the respective McCrary test.

Figure A3. Catchment area boundaries in Helsinki.

Notes: The solid lines represent the true catchment area boundaries and the dashed lines the fake boundaries.