VATT Working Papers 58

# Best education money can buy? Capitalization of school quality in Finland

*Oskari Harjunen*
*Mika Kortelainen*
*Tuukka Saarimaa*

GOVERNMENT 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 2014

We thank Essi Eerola, Elias Einiö, Antti Moisio, Tuomas Pekkarinen, Matti Sarvimäki, Janne Tukiainen and the seminar participants at FDPE, ERSA 2013, HECER, Jyväskylä and VATT for their useful comments.

The research is part of the Academy of Finland project “Modeling Housing Market Frictions – Tools for Policy Analysis” (grant number 255445).

Oskari Harjunen, Aalto University School of Economics and City of Helsinki Urban Facts. Email: oskari.harjunen@aalto.fi.

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

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

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

ISSN 1798-0291 (PDF)

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

Edita Prima Oy
Helsinki, December 2014

Cover design: Niilas Nordenswan

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

## Government Institute for Economic Research
## VATT Working Papers 58/2014

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: a one standard deviation increase in average test scores increases prices by roughly 2.5 percent, which is comparable to findings from the U.K and the U.S. This price premium is related to pupils' socio-economic background rather than school effectiveness.

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

JEL classification numbers: C21, H75, I20, R21

## Tiivistelmä

Tutkimuksessa tarkastellaan sitä, vaikuttavatko peruskoulujen väliset erot asuntojen hintoihin Helsingissä. Kansanvälisessä vertailussa suomalaisten oppilaiden oppimistulokset ovat hyviä ja samanaikaisesti koulujen väliset erot oppimistuloksissa pieniä. Suomalainen koulujärjestelmä eroaa useiden muiden maiden järjestelmistä myös siksi, että suomalaisia peruskouluja ei testata systemaattisesti standardoiduilla kokeilla, eikä tuloksia julkisteta niissä harvoissa tapauksissa, joissa kunnan kaikki koulut testataan. Näistä seikoista huolimatta havaitsemme, että koulujen väliset erot heijastuvat selvästi asuntojen hintoihin. Tulostemme mukaan yhden keskihajonnan nousu koulun keskimääräisessä standardoidussa koepistemäärässä nostaa asuntojen hintoja keskimäärin 2,5 prosenttia. Vaikutus on samaa suuruusluokkaa kuin Isossa-Britanniassa ja Yhdysvalloissa, joissa koulujen väliset erot ovat aiempien tutkimusten mukaan Suomea suurempia. Erityis- tai vieraskielisten oppilaiden osuudet sen sijaan eivät vaikuta asuntojen hintoihin. Lisätarkastelut viittaavat siihen, että asunnonostajille on tärkeää koulun oppilaiden sosioekonominen tausta, ei koulun vaikutus oppimiseen.

Asiasanat: Asuntojen hinnat, koulun laatu

JEL-luokittelu C21, H75, I20, R21

## 1. Introduction

Empirical evidence from a number of studies shows that differences in school quality, measured by school value-added, average 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, mainly 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. 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.[1] 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.[2]

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 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

[1] The Finnish basic education system has received a great deal of attention around the world, especially in the U.K. and the U.S. 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 16 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.

[2] 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).

only used internally by the schools or 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.[3] The 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 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.[4]

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. Vaattovaara and Kortteinen 2003; Dhallman and Vilkama 2009). 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.

With these institutional aspects in mind, it is interesting to find out whether parents really perceive that the Finnish schools are of equal quality. Anecdotal evidence suggests that parents in Helsinki go to some lengths in securing their child's place in a particular school.[5] In this paper, we ask whether this pattern is present more generally by studying whether homebuyers are 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.

---

[3] For example, using the U.K. data Hussain (2014) shows that inspection ratings can provide additional information on school quality, which is relevant for students and parents on top of other observable school characteristics.

[4] 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.

[5] Perhaps the most famous example is the story reported in Helsingin Sanomat (21.5.2011), the largest newspaper in Finland, about a couple who divorced on paper so that their child could gain residence within a particular catchment area.

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. This may dilute the relationship between school quality and house prices as shown by Fack and Grenet (2010), Machin and Salvanes (2010) and Brunner et al. (2012) in other contexts. 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.[6]

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 2.5 percent. This result is robust across a number of specifications and the size of the effect is comparable to findings from the U.K. and the U.S. 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' 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 rather 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. 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 disclosed.

[6] In our baseline specification the maximum (mean) distance between matched units is 400 (235) meters. The results are robust to narrowing this maximum (mean) distance to 200 (134) meters.

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

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

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.[7]

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 (RDD), 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 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. This means that the probability of gaining access to the school jumps from something below unity to one. Second, there cannot be discontinuities in other neighbourhood characteristics exactly at the boundaries. These may arise due to exact residential sorting at the boundaries, as

[7] 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.

suggested by Bayer et al. (2007), or to other boundaries that coincide with catchment areas. 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 fundamentally 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 for a single municipality, which means that other policies, such as local tax rates, stay constant within the area. Second, 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. These controls are essential if households sort across catchment areas and the sorting prevails at the boundaries (Bayer et al. 2007).[8] Third, we can eliminate boundaries that coincide with major geographic obstacles that may induce discontinuities in neighbourhood quality, such as major roads, railways or waterways. Finally, we use transaction data only for 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. Hilber and Mayer, 2009).

While solving the major identification problem, spatial differencing introduces some problems for statistical inference because a particular housing unit may be the

[8] Bayer et al. (2007) also stress that 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 may 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 housing market (50 in our case). See also Bayer and McMillan (2008) for further discussion.

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 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.[10] 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).[11] 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 Finland, but these schools share the legislation of the public school system and are therefore very similar to public schools. 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

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

[10] There is also a one-year optional pre-school before primary school.

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

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 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).[12]

### 3.2. School quality measures

What exactly are the right school quality measures, and 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 average standardized math test scores from the 6$^{th}$ grade, the share of foreign-language pupils and the share of pupils with special needs. The school quality measures are for 2008.[13]

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 they need much extra attention and drain teachers' resources.[14]

[12] Homebuyers ought to be well informed about the school district that a house belongs to. For example, in Helsinki there is a free internet-based service offered by the city of Helsinki which assigns every address to a specific elementary school.

[13] We also obtained data on schools' total expenditure. We decided not to use it, because per-pupil expenditure is not comparable across schools due to reporting problems and measurement errors. These problems are mostly related to comparing schools that have only grades 1–6 to schools that have grades 1–9.

[14] 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 Lappish) as their

Unfortunately, we do not have data on other pupil or parent characteristics. However, in some additional models we use proxy measures for parents' characteristics. 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 standardized math test was 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 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.[15]

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 for the same year and building type on the opposite side of a catchment area boundary (see also Gibbons et al. 2013). In our baseline estimations, we use data where the maximum distance between matched units is 400 meters. In robustness analysis we vary this maximum distance from 400 to 200 meters.[16]

---

mother tongue. Pupils with special needs and foreign-language pupils receive their education partly in classes with regular pupils and partly in special groups.

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

[16] Gibbons et al. (2013) also use fake boundaries as placebo tests. We obviously cannot do this because there would be no variation in the school quality measure at a fake boundary within a catchment area.

Fig. 1 illustrates the catchment area boundaries in Helsinki. We use only the boundaries where access to grades 7–9 does not change. In Fig. 1, these boundaries are marked with a dashed 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). Fig. 1 shows that our data are spread over the whole of Helsinki. 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.


**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 that is assigned to them. 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 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 of 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.[18] We use only multi-storey and row houses in order to make sure that

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

[18] For confidentiality reasons, we are not allowed to map the transactions.

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.

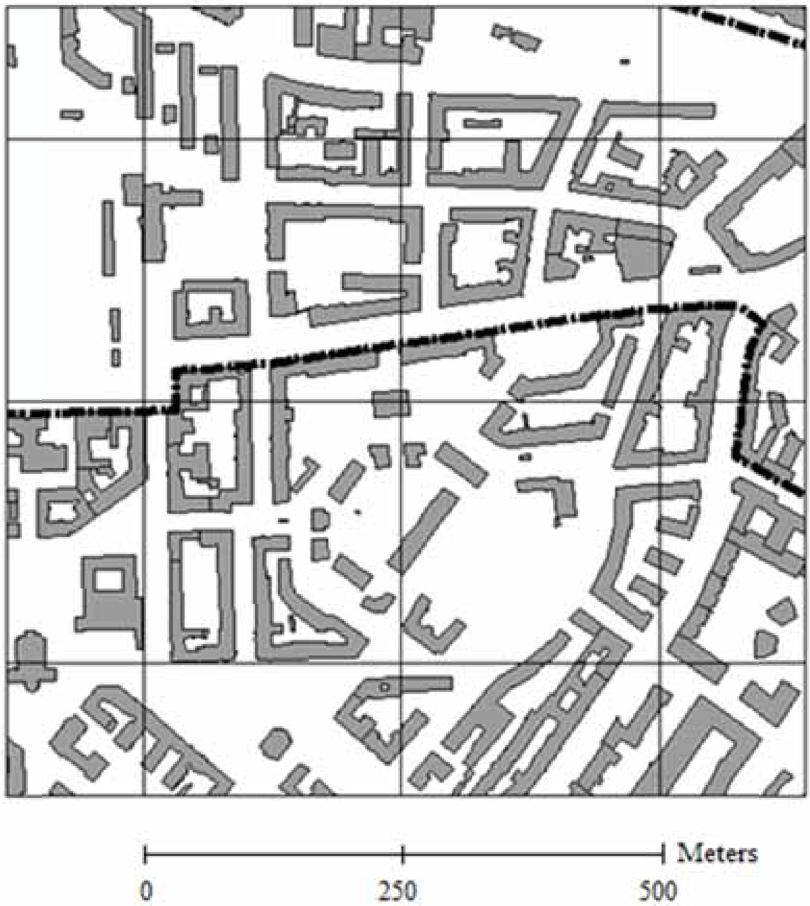


**Figure 2**. Buildings and close neighbourhood grids at a single catchment area boundary.

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.[19]

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

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 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 neighbour.[20] 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.

[20] 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 |
| % pension 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. 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.[21] First, we control for housing unit characteristics in all specifications to capture systematic differences between matched units. In addition to unit-level heterogeneity, we need to worry about household heterogeneity because sorting of heterogeneous households across catchment areas and specifically at the catchment area boundaries may lead to bias if households have preferences for their neighbours' characteristics (Bayer et al. 2007; Gibbons et al. 2013). In order to mitigate this issue, in column (2), we add close neighbourhood controls.[22] Finally, in column (3) we add a third-order polynomial of distance to the catchment area boundary in order capture any spatial trends in prices due to amenities. 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. A one standard deviation increase in test scores increases prices by roughly 2.5 percent. This result is statistically significant both when using standard clustering and also when using the wild bootstrap. Reassuringly, the result is robust both to adding close neighbourhood controls and distance to the boundary polynomials.[23]

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 close neighbourhood characteristics.

[21] The full results of the models in Panel B of Table 3 are presented in Table A2 in the Appendix.

[22] If there is sorting exactly at the boundary, some of the neighborhood variables may be endogenous because of unobservable neighborhood characteristics that are correlated with the observable characteristics (Bayer at al. 2007 and Gibbons et al. 2013). Fortunately, we are able to control for a large set of household characteristics, making this problem less of a concern than in most previous studies.

[23] In Fig. A1 in the Appendix, we also show the effect in Panel A of Table 3 graphically using a standard RDD approach with local linear regressions. However, RDD is not suited to cases where school quality changes in a number of dimensions and when the interest lies on these different dimensions.

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

| | Panel A: Test score | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Math test score | 0.034*** | 0.028*** | 0.028*** |
| | [0.010] | [0.007] | [0.007] |
| | (0.080) | (0.004) | (0.006) |
| N | 3852 | 3852 | 3852 |
| $R^2$ | 0.84 | 0.85 | 0.85 |
| | Panel B: Test score and pupil composition | | |
| | (4) | (5) | (6) |
| Math test score | 0.032*** | 0.027*** | 0.026*** |
| | [0.009] | [0.008] | [0.008] |
| | (0.016) | (0.022) | (0.024) |
| % special needs pupils | 0.077 | 0.027 | 0.023 |
| | [0.146] | [0.144] | [0.146] |
| | (0.706) | (0.897) | (0.915) |
| % foreign-language pupils | -0.183** | -0.110 | -0.105 |
| | [0.076] | [0.075] | [0.075] |
| | (0.058) | (0.248) | (0.272) |
| N | 3852 | 3852 | 3852 |
| $R^2$ | 0.84 | 0.85 | 0.85 |
| Unit characteristics | yes | yes | yes |
| Close neighbourhood characteristics | no | yes | yes |
| Distance to boundary polynomials | no | no | yes |

Notes: 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 a quarter-year of sale dummies.

### 4.2. Robustness using alternative samples

The main identifying assumption of our econometric approach is that other neighbourhood characteristics develop smoothly across catchment area boundaries. Since we omit boundaries that coincide with major geographic obstacles and use a sample where the maximum distance between matched units is less than 400 meters, this assumption is likely to hold. Nonetheless, we have run a number of robustness checks, which are reported in Table 4.

The first issue is that some catchment area boundaries may coincide with a 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 (1) 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.

Second, in columns (2) and (3) of Table 4, we present results using our main specification 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.

Finally, we concentrate on boundaries where it is difficult to send a child to the school on the opposite side of the boundary.[24] Here we use the share of children living within a catchment area that attend the school in their catchment area. If this share is large for a particular school, it might be more difficult for children from other catchment areas to attend that school. We label the schools for which this share is above the median of the schools in our data as high-intake schools. In column (4) of Table 4, we use only those observations that are on the "wrong" side of a high-intake school's catchment area boundary. Also in this subsample, test scores are capitalized and the effect is of a similar magnitude as in the full sample.

Interestingly, at these boundaries the share of foreign language pupils also obtains a significant coefficient when inference is based on clustered standard errors. However, the number of clusters in this sample is only 26, so we are reluctant to rely on these standard errors. In fact, the wild bootstrap procedure produces a *p*-value of 0.126, meaning that we cannot reject the hypothesis of no effect.

The final robustness check that we conducted is related to the timing of sales. In Tables 3 and 4, we used transactions for 2008 to 2012, but the school quality measures

[24] Gibbons et al. (2013) do a similar exercise. They concentrate on boundaries where pupils rarely cross. We do not have this information available to us. We only know the share of pupils who chose the school in their own catchment area.

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 A3 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.

**Table 4**. Robustness checks using alternative samples.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Same zip-code | Maximum distance < 300m | Maximum distance < 200m | Next to a high-intake area |
| Math test score | 0.024*** | 0.035*** | 0.038** | 0.028*** |
| | [0.005] | [0.012] | [0.016] | [0.008] |
| | (0.000) | (0.052) | (0.056) | (0.044) |
| % special needs pupils | -0.027 | 0.188 | 0.03 | 0.28 |
| | [0.161] | [0.185] | [0.219] | [0.321] |
| | (0.971) | (0.561) | (0.873) | (0.511) |
| % foreign language pupils | -0.080 | -0.082 | -0.038 | -0.348** |
| | [0.075] | [0.076] | [0.095] | [0.159] |
| | (0.410) | (0.396) | (0.805) | (0.126) |
| N | 2725 | 2770 | 1515 | 1804 |
| $R^2$ | 0.84 | 0.86 | 0.87 | 0.89 |
| Unit characteristics | yes | yes | yes | yes |
| Close neighbourhood characteristics | yes | yes | yes | yes |
| Distance to boundary polynomials | yes | yes | yes | yes |

Notes: 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 (1) 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 and close neighbourhood characteristics are reported in Table 2. All the models include a quarter-year of sale dummies.

### 4.3. Discussion of likely mechanisms

The overall picture that emerges from Tables 3 and 4 is that school quality (broadly defined) is capitalized into house prices in Helsinki. However, it is not clear whether Finnish 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 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). We can follow these studies, although only imperfectly. To do so, we calculated average 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 5 presents results where these two variables are added into our main model specification. In column (1), we add only mean income, in column (2) education level, and finally, in column (3) both. We learn two things from Table 5. 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.[25] Second, and more importantly, once we add either or both of the proxies for parental background, the test score coefficient diminishes considerably. This suggests that a large part of the test score effect can be explained by pupil composition.

[25] 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 very careful with the interpretation of the coefficients of income and education in column (3). First, the partial effect of education is generally different when we condition and when we do not condition on income. Second, since both proxy variables are measured with error, it is difficult to know how measurement error affects the estimates in this case.

**Table 5**. Effects of school quality and catchment area measures.

| | (1) | (2) | (3) |
|---|---|---|---|
| Math test score | 0.011 | 0.018** | 0.012** |
| | [0.008] | [0.008] | [0.006] |
| | (0.156) | (0.052) | (0.030) |
| % special needs pupils | -0.016 | 0.045 | -0.163 |
| | [0.125] | [0.143] | [0.123] |
| | (0.949) | (0.861) | (0.380) |
| % foreign language pupils | -0.022 | -0.031 | 0.003 |
| | [0.080] | [0.077] | [0.058] |
| | (0.815) | (0.675) | (0.957) |
| Catchment area income | 0.048*** | | 0.066*** |
| | [0.012] | | [0.015] |
| | (0.002) | | (0.001) |
| Catchment area education | | 0.301* | -0.455** |
| | | [0.175] | [0.186] |
| | | (0.220) | (0.024) |
| N | 3852 | 3852 | 3852 |
| $R^2$ | 0.84 | 0.85 | 0.85 |
| Unit characteristics | yes | yes | yes |
| Close neighbourhood characteristics | yes | yes | yes |
| Distance to boundary polynomials | yes | yes | yes |

Notes: 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 and close neighbourhood characteristics are reported in Table 2. All the models include a 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, no school-level value-added measures have been estimated for either primary or lower secondary schools in Finland. Moreover, previous research shows that without publicly released value-added measures, it might be either challenging for parents to identify schools with superior value-added or they might simply use test scores as an imperfect proxy for school effectiveness (see e.g. Rothstein 2006, Zhang 2013). Second, several studies have found that the most selective schools or schools with favourable student composition are not the top schools in terms of value-added. For example, Abdulkadiroglu et al. (2014) show that exam schools, which are very popular and selective lower secondary schools, have little effect on student achievement (or value-added) in Boston and New York. Similarly, Zhang (2013) finds that in China the most sought-after elite middle schools are those with the highest student achievement level, rather than those with the largest value-added effect on test scores.

Finally, it is unlikely that differences in value-added would be 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. In fact, we know from recent research that even in very selective Finnish upper secondary schools, value-added estimates are very similar or indistinguishable from most other schools (Kortelainen et al. 2014). Given our additional results and the discussion above, we think that it is likely that most of the 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 2.5 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. 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


Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica* 82(1): 137–196.

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." Andrea Baranzini, José Ramirez Caroline Schaerer and Philippe Thalmann (ed.s) Hedonic Methods in Housing Markets. Springer.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The 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*, Volume 3. North-Holland: Amsterdam.

Brasington, David and Donald R. Haurin. 2006. "Educational Outcomes and House Values: A Test of the Value Added Approach." *Journal of Regional Science* 46(2): 245–268.

Brasington, David and Donald R. Haurin. 2009. "Parents, Peers, or School Inputs: Which Components of School Outcomes are Capitalized into House Prices." *Journal of Regional Science* 39: 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.

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.

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.

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.

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. 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. 2014. "Subjective Performance Evaluation in the Public Sector: Evidence From School Inspections." Forthcoming *Journal of Human Resources*.

Kortelainen, Mika, Heikki Pursiainen and Jenni Pääkkönen. 2014. "Value-Added Rankings of Finnish Upper-Secondary Schools." [In Finnish] VATT Research Reports 179.

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. 2010. "Valuing School Quality via a School Choice Reform." IZA Discussion Paper No. 4719.

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.

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.

Vaattovaara, Mari and Matti Kortteinen 2003. ”Beyond Polarisation versus Professionalisation? A Case Study of the Development of the Helsinki Region, Finland.” *Urban Studies* 40(11): 2127–2145.

Zhang, Honliang. 2013. The Mirage of Elite Schools: Evidence from Lottery-based School Admissions in China. Mimeo.

**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 |
| % pension 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**.Full results for the main model specifications.

| | Coeff. | S.D. | Coeff. | S.D. | Coeff. | S.D. |
|---|---|---|---|---|---|---|
| Constant | 0.002 | 0.009 | -0.003 | 0.011 | 0.043 | 0.025 |
| Math test score | 0.032*** | 0.009 | 0.027*** | 0.008 | 0.026*** | 0.008 |
| % special needs pupils | 0.077 | 0.146 | 0.027 | 0.144 | 0.023 | 0.146 |
| % foreign language pupils | -0.183** | 0.076 | -0.110 | 0.075 | -0.105 | 0.075 |
| log(floor area) | 0.892*** | 0.060 | 0.894*** | 0.058 | 0.893*** | 0.057 |
| log(age) | -0.051**** | 0.008 | -0.060**** | 0.011 | -0.059**** | 0.010 |
| Good (0/1) | 0.168*** | 0.020 | 0.153*** | 0.020 | 0.151*** | 0.019 |
| Satisfactory (0/1) | 0.073*** | 0.021 | 0.059*** | 0.017 | 0.059*** | 0.017 |
| Own lot (0/1) | 0.065** | 0.025 | 0.080** | 0.022 | 0.084*** | 0.022 |
| Elevator (0/1) | -0.003 | 0.023 | -0.007 | 0.022 | -0.008 | 0.021 |
| Floor level | 0.020*** | 0.003 | 0.019*** | 0.003 | 0.019*** | 0.003 |
| Total number of floors | -0.011**** | 0.002 | -0.010**** | 0.002 | -0.010**** | 0.002 |
| Maintenance charge | -0.007 | 0.005 | -0.009 | 0.005 | -0.009 | 0.005 |
| Road distance to CBD (km) | 0.127* | 0.070 | 0.159*** | 0.050 | 0.165*** | 0.048 |
| Homeownership rate | | | -0.121 | 0.063 | -0.123* | 0.063 |
| Mean income (€) | | | 0.106*** | 0.027 | 0.110*** | 0.030 |
| % college degree adults | | | 0.302 | 0.205 | 0.311 | 0.207 |
| Unemployment rate | | | -0.372 | 0.303 | -0.371 | 0.324 |
| % pension households | | | -0.135 | 0.155 | -0.118 | 0.153 |
| % households with children | | | -0.142 | 0.159 | -0.127 | 0.157 |
| Number of service jobs per capita | | | -0.003 | 0.005 | -0.002 | 0.005 |
| Number of buildings | | | 0.002** | 0.001 | 0.002** | 0.001 |
| Mean floor area of units ($m^2$) | | | -0.162** | 0.074 | -0.165** | 0.078 |
| Population | | | -0.000036* | 0.000 | -0.000035* | 0.000 |
| % foreign language residents | | | -0.051 | 0.203 | -0.030 | 0.197 |
| N | 3852 | | 3852 | | 3852 | |
| $R^2$ | 0.84 | | 0.85 | | 0.85 | |
| Distance to boundary polynomials | no | | no | | yes | |

Notes: The table reports the full results for spatially differenced models, corresponding to Table 3 in the main text. 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. All the models include a quarter-year of sale dummies.

**Table A3**. 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 | 0.026*** | 0.027*** | 0.028*** | 0.037*** |
| | [0.008] | [0.008] | [0.009] | [0.009] |
| | (0.024) | (0.016) | (0.022) | (0.006) |
| % special needs pupils | 0.023 | 0.049 | 0.061 | 0.037 |
| | [0.146] | [0.175] | [0.193] | [0.202] |
| | (0.915) | (0.853) | (0.833) | (0.889) |
| % foreign language pupils | -0.105 | -0.150 | -0.130 | -0.133 |
| | [0.075] | [0.093] | [0.089] | [0.093] |
| | (0.272) | (0.238) | (0.280) | (0.282) |
| N | 3,852 | 3,100 | 2,273 | 1,403 |
| $R^2$ | 0.85 | 0.86 | 0.87 | 0.87 |
| Unit characteristics | yes | yes | yes | yes |
| Close neighbourhood characteristics | yes | yes | yes | yes |
| Distance to boundary polynomials | yes | yes | yes | yes |

Notes: 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 and close neighbourhood characteristics are reported in Table 2.

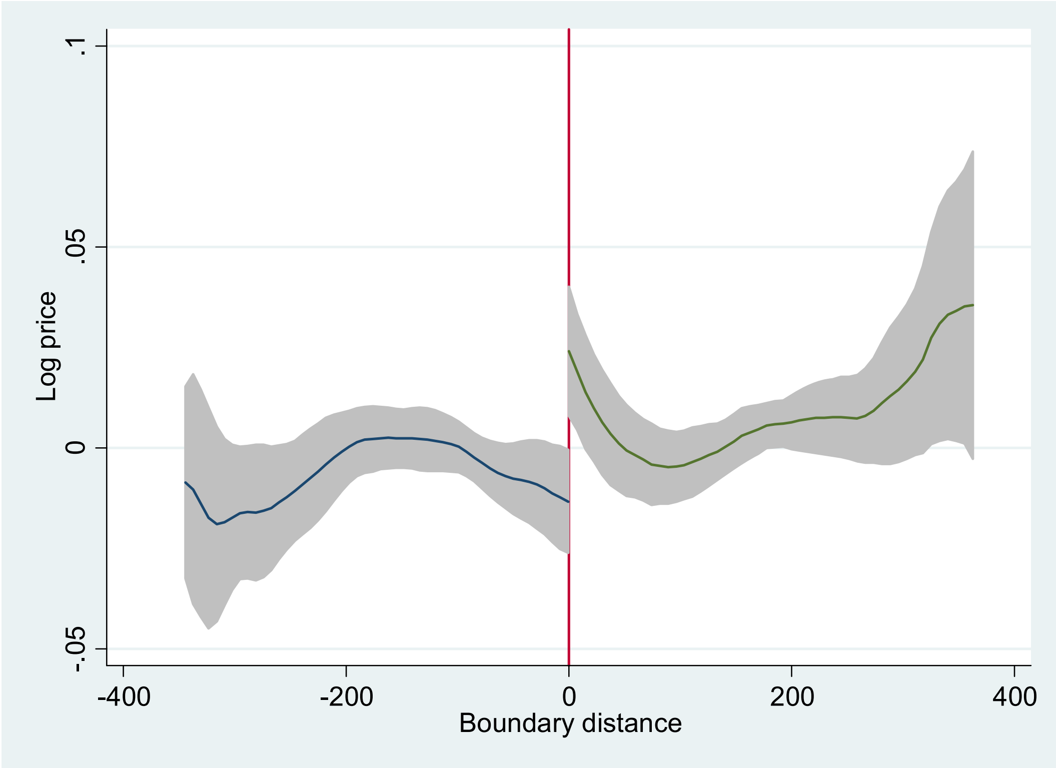


**Figure A1**. Discontinuity in test scores and prices.

Notes: The figure depicts fits and confidence intervals for local linear regressions with a triangular kernel. The left-hand-side variable is the residual from a hedonic price regression where we control for unit and close neighbourhood characteristics. The bandwidth in the local linear regression is 150. The confidence interval does not account for clustering.