Stockholm School of Economics
Department of Economics
5350 Master's Thesis in Economics
Academic Year 2015/2016

Pricing Peers in Stockholm:

The Effect of School Grades on Housing Prices

Simon Vinge, 22028

Abstract: The relationship between school grades and house prices has not been properly studied in Sweden. Research is further warranted by the institutional setting of the capital municipality, Stockholm: school districts with the parallel rise of free, choice-elected, private schools. Using all detached house transactions (N=33,060) between 2001–2013, a hedonic model shows a strong positive correlation between lagged average city district final grades and house price premium, using fine grained locality controls. Further specifications draw upon the continuous distance to quality schools, interacted with the discontinuous school district endowment. The effect of grades is still clear, but the upward bias in the basic specification is very evident. At a mean school district size, the value added of the school also belonging to the district is positive and increasing for grades over 235 for a maximum house premium of 7.8 percent.
Keywords: House prices, School quality, Spatial analysis
JEL: R21, I28, C21

Supervisor: Örjan Sjöberg
External Supervisor: Dany Kessel
Date Submitted: May 16, 2016
Date Examined: May 23, 2016
Discussant: Sigita Zvirblyte
Examiner: Kerem Cosar

Contents

1 Introduction **4**

2 Theoretical Framework **5**
2.1 The Hedonic Regression . . . 5
2.1.1 Handling endogeneity . . . 6
2.1.2 What to Price? . . . 7
2.2 The Price of School Quality . . . 8
2.3 School Admission Policy in Sweden and Stockholm . . . 10

3 Research Question **13**
3.1 Theoretical Predictions . . . 13

4 Data **13**
4.1 Schools . . . 14

5 Empirical Strategy **17**
5.1 Econometric Specifications . . . 17
5.2 Choice of Variables . . . 19
5.3 Average School District Effect . . . 19
5.4 Distance or Endowment . . . 20
5.5 Hypothesis Specifications . . . 20

6 Results **21**
6.1 School District Estimates . . . 21
6.2 The Price of Endowment . . . 22

7 Discussion **26**
7.1 Conclusions . . . 27
7.2 Further Studies . . . 28

8 References **29**

A Terms Used **31**

B Tables and Figures **31**

List of Tables

1 **House Price Controls**, Stockholm 2000–2013 13
2 **Sales per House Block** 14
3 **Number of House Sales**, by Year 15
4 **House Resales** 15
5 **House Sale**, by Construction Decade 16
6 **Grades and Number of Municipal Schools**, by Stadsdel 16
7 **District Average Specification**, Estimation Results 24
8 **Distance Specification**, Estimation Results 25
9 **Prices by Year and Stadsdel**, TSEK 32
10 **Average Income and Population**, by Year and Stadsdel 33
11 **District Average Specification**, Estimation Results Full Controls 34

List of Figures

1 **Proportion of Stockholm Municipality Students in Private or Non-Municipal Pre- and Primary Schools**, by Year 11
2 **Approximate Rendering of School Districts Within Bromma City District** 12
3 **Border Distance and Estimation Probability** 18

Acknowledgements

This thesis has been written alongside a research project studying school choice through a number of aspects at Stockholm University, led by Elisabet Olme and Dany Kessel. The research group has been a frequent and invaluable help, and I hope that some parts of my writings will become of future use for this research. I am also in debt to: Patrik Tingvall at Ratio institute, and his spatial analysis skills; Nils Westling without whose help I could not even project a map in the geographic software package I used; Ingela Brimberg at Stockholm municipality who with great patience answered my countless school questions; Marianne Ekengren at Äppelviksskolan helping me to bring clarity to all peculiarities of the Stockholm school choice system; Dany Kessel for assuming an invaluable role as informal supervisor—and for letting me occupy his computer for calculations that my laptop and no SSE computer could handle; Elisabet Olme and Erik Lindgren for adding immensely to my abysmal coding skills; and last but not least Thorun Wahlström and her tolerance towards the strange working hours of thesis writing when the needs of a six month old is actually much more important; among many others.

1 Introduction

For the individual family, school choice is often perceived as one of the most important parental decisions. However, the degree of choice expressed co-varies strongly with socioeconomic status. This makes residential sorting and segregation an obvious consequence. Since some studies point towards peer quality being significant for individual outcomes (Söderström and Uusitalo, 2010) it is highly relevant to clarify the link between house pricing, school quality and school choice. This is especially important in a Swedish context, being a country that has seen much reform but little research on the economics of school choice.

Choice is always present. The family school preference can be expressed through a voucher system as famously argued by Friedman (1955). Another common allocation mechanism is through the geographic distribution of students, which makes school choice a second order effect of resident location. Partly for this reason, finding *how much* school endowment is actually valued, has been a frequently asked research question. It has been difficult to answer, however.

The literature points towards estimates ranging between the most important cause of variation to somewhere between 1.4–20 percent of house value—but the most commonly cited method, as used by Black (1999), generally explains 5–10 percent of house value variation. However, even the best estimates are probably very biased due to the deeply complex and dynamic nature of resident location.

The potential house price school premium is obviously dependent on institutional specifics regarding school system, grading, demographic structure, the housing market, etc., which warrants research in the particulars of institutional settings. To this date, no proper study into this premium has been conducted within Stockholm or Sweden—despite excellent data, and despite large recent institutional change and continuous reform undertaken.

However the different country estimates are often remarkably similar in magnitude. For example when comparing the US and the UK, the school funding is very different. In the US the local tax base is the main source of school funds which increases the effect of the neighborhood channel, whereas the UK, just like Sweden, has centralized the funding base to the national level. This gives some specific dynamics in the US, such as the price effects of charter schools which bypasses the local tax funding, but it does not seem to make any large differences for the general school district price effects of regular state schools.

So while it might be relevant to explore the specific context of Sweden, I argue that the institutional setting of schooling of capital municipality Stockholm will also give insights that are applicable to the more general relationship between schools and house prices.

The institutional setting is that of a school district that mimics the UK and US quite well: each address is tied to a particular municipality run school, constant over time. Added to this is a voucher school system that exists parallel to the municipality schools, where only queue time is relevant for application. This ‘free school’-system receives the same funding as municipality schools and is, just like all other schools, not allowed to charge any fees.

Previous literature show that a similar penetration of non-fee, private, voucher schools in France significantly reduced the state school house premium. Alongside this finding, it is common in the UK studies to use the private school fees as an estimation cap of school premiums since this category of schools is seen as a substitute for state run schools.

This study can thus add to previous studies in several ways. As well as the productive institutional setting of Stockholm municipality referred to above, I can utilize the high quality data for both housing stock and incomes available in Sweden. This will allow me to sort out some of the most obvious upward biasing of previous studies. Adding to this, I can utilize the rich geo-data to address geographic endogeneity through isolating the discontinuous school district endowment effect from a continuous school distance effect.

Within Stockholm, during 2000–2013, I show that there is an economically and statistically significant positive effect of endowment to schools with higher grades. When improving the standard hedonic model estimates through more fine grained geo-coding though, I can show how the more naive measurements overstate the true effect of school quality on housing prices to a large degree.

I can also confirm the high quality of housing stock data through using a subsample of resales where estimates are very similar to the basic model. This indicates that bias is due to residential sorting rather than uncontrolled for housing variables or income.

I also show that despite its massive growth during the period (roughly doubles in proportion of students), the penetration of costless private voucher schools has not seemed to affect the school district house price premium. This shows that non-district schools must be carefully considered within their context, and probably only affect the school price premiums if their degree of substitution is very high.

2 Theoretical Framework

2.1 The Hedonic Regression

Finding the housing price premium assigned to school district quality enter a more general tradition of hedonic literature, aiming to find the value assigned to various characteristics of real estate. Herath and Maier (2010) gives a review of its history where, in its present form, Rosen (1974) is cited as seminal. From his 1974 study follows a literature of finding the price effect of various variables on real estate—including such various applications as the (negative) effects of crime rates upon apartment prices (Ceccato and Wilhelmsson, 2011) or the (unclear) value of nearby wildlife (Maslianskaia-Pautrel and Baumont, 2016) to just mention a few of countless empiric applications.

The research on school quality effects upon housing prices is very large in a UK and US context. This is in part explained by the common policy of rigid geographic school districts for state run schools, whose quality variation can be substantial. Expensive private schooling is often the only channel for opting out of assigned state run school, which makes the choice of address very important for school assignment. Indeed the attempts to measure this school district price effect on housing generally renders large estimates. Certain studies even suggests that "school quality is the most important cause of the variation in constant quality house prices"(Haurin and Brasington, 1996, p. 363). Sahlgren and Jordahl (2016) proposed in a recent literature review of 24 articles that "the link between school quality and house prices is one of the most robust connections that has been studied within the field of empirical economics"[1] (Sahlgren and Jordahl, 2016, p. 128).

Despite this robust link, estimates vary and seem sensitive to specifications and endogeneity issues, both of which I will return to shortly. Schooling is steeped into local peculiarities of what matter for perceived quality, be it test scores, grades, pass rates, etc.—and any quality variable is therefore correlated with most independent variables.

It is also striking that the connection has not been addressed in Sweden, and the closest studies in the Sahlgren and Jordahl (2016) meta study are from Norway and Finland, respectively; and

[1] "Sambandet mellan skolkvalitet och huspriser är ett av de mest robusta sambanden som har analyserats inom empiriska nationalekonomiska studier," my translation.

this despite an institutional setting of Stockholm municipality that makes its housing market an excellent research object, as of why I will return to.

2.1.1 Handling endogeneity

As noted, the key issue for a hedonic model is endogeneity. For something so complex as the valuation of school qualities observed through house transactions, a first fear is naturally unobserved variables. As is well known, correlation between quality variables such as grades or test scores, and our dependent value, can bias our estimates in any direction. Many of these are probably neighborhood characteristics where no set of control variables is exhaustive. A second issue is that of reverse causality, where the quality of interest can be driven by the immigrating residents through local engagement, say, rather than the school quality attracting them in the first place. This is related to residential sorting, where we risk capture the characteristics of residents sorted for entirely different reasons instead of the causal effect of school quality (Kane et al., 2006).

School "qualities" are indeed closely correlated with many relevant attributes. Determinants of school grades are to a certain degree generated by the school itself—but probably to a much higher degree through *peer effects*. The term is here used in its most general sense, as suggested by among others Rothstein (2006), to attribute everything that is not the specific value added by the school: "[Peer effects include] the effects of individual student characteristics on their own test scores, any direct peer group effects, and any indirect effects of a school's composition on the quality of its instruction. If wealthy schools attract better teachers or more parental involvement, this is for my purposes a peer effect" (Rothstein, 2006, p. 1333).

This can be handled in four ways following Chiodo et al. (2010), which is in practice rather just in two ways according Cheshire and Sheppard (2004). Namely through a natural experiment—or through "variants of hedonic regressions" (Cheshire and Sheppard, 2004, p. F398).

Using natural experiments—that is, finding an event that will randomly distribute independents of interest—is both difficult and by its nature opportunistic: we have to look where the experiment takes place. In the hedonic school literature, one such rare occurence is in Shanghai where Li (2012) argues that recent property market reform makes price changes occur after school quality was known and public, and that thus capitalization can be interpreted causally. Other attempts have been made using district boundary changes, such as by Bogarth and Cromwell (2000), but it is generally difficult to infer such changes are neither unanticipated nor that district changes are random and made without neighborhood considerations, and therefore not strictly exogenous.

What is referred to above as a second category contains variants of hedonic regressions can be broadly summarized into a few categories; namely non-random discontinuities, time variance, spatial and instrumental methods.

Through finding non-random but nonetheless discontinuities, both Black (1999) and Gibbons and Machin (2003), argue for a causal interpretation of their estimates through identifying discrete changes at school district borders. Given an otherwise continuous change between areas, including the assumption of smooth residential sorting, they argue that observing a discrete difference between two districts can be assigned to school quality difference. The approach of Black (1999) has for this reason been reproduced by a number of authors among them Harjunen et al. (2014).

A weaker exogeneity assumption used by Bogarth and Cromwell (2000) is that with good enough time variation, e.g. variance in measured school quality over time, causal inference can be made with respect to house prices. It is not entirely clear however, why quality measures would not co-variate with sorting or omitted variables.

An important component of endogeneity is related to sales not being spatially independent, given that prices are generally correlated with nearby prices—a correlation assumed in these models to diminish linearly over distance. Through econometric methods, a number of authors corrects for this spatial autocorrelation through a spatial autocorrelation matrix method. One of the authors include Brasington (1999) who find that applying this method had little effect upon estimates, which could indicate that the conventional spatial controls at their disposal captured most of the spatial influence—or that the spatial approach did not.

The issue with all of these three approaches is obviously that there is no way to test their exogeneity assumption. Cheshire and Sheppard (2004), among others, criticize the (Black, 1999) approach in that it does not fully account for variation in localities that we do not have data for: e.g. such things as view, noise, local amenities.

A fourth distinct method for reducing bias is through using instrumental variables, but it is evident from browsing the literature that it is difficult to find robust instruments that holds through the required assumptions. Gibbons and Machin (2003) argue that through instrumenting with church schools that have existed for a long time, the bias of reverse causality is reduced—since the in-flow of new parents will not, is assumed, affect the conservative structure of the church schools as compared to more responsive and recent non-religious schools. I will leave to the reader to determine the strength of that assumption and whether all bias can be assumed absorbed, but note a stronger instrument utilized by Rosenthal (2003). Through sampling newly Ofstedt inspected schools, they relate the common surge in grades as is common after a random visit by the UK school inspection agency to improved school quality.

The issue with the last category is of course finding a proper instrument, and then arguing that it holds for all instrument assumptions. As of this date, the cases put forth have not been totally convincing.

2.1.2 What to Price?

A central question is what school quality we want to estimate the implicit value of. A choice can broadly be made between *inputs* such as teacher wages or general costs, *value added* (VA) calculated through various more or less complex methods, and last through *outputs* such as final grades or test scores.

Brasington (1999) does a thorough overview of these measures and concludes that inputs such as teacher wages are sensitive to estimation technique, but that spending per pupil and pupil/teacher ratio are consistently capitalized into prices. He also shows that VA-measures, such as teacher experience and education, in the sample do not show any consistent correlation to housing prices and argues it should be avoided as quality measure. In a more recent study further into value added methods, they find little support and concludes that "regardless of how much better value added may be as a measure of the performance of school inputs, only the levels of test scores and expenditures are capitalized." (Brasington and Hauring, 2006, p. 247).

This also seems to be the sentiment of the majority of research performed, which most often use such output measures as test scores and pass grades. It is also plausible that it is these measures that are priced, since they are generally the most observed and easily interpreted. The measures also hint at the quality Rothstein (2006) argue is the most important: namely that, more explicitly, of peers. The importance of peers indicates, according to Rothstein (2006), choice is based upon pupils as much as other factors—of which grades is the key signal. The Harjunen et al. (2014) study from Finland also confirms this claim, finding that the house price premium is related to pupils' socioeconomic background rather than school quality.

This links to the discussion regarding controls as raised by most authors and notably Cheshire and Sheppard (2004) for insight regarding socioeconomic controls. They conclude that socio-economic controls risk reduce the fit of the model, and find a much lower effect when they drop the socio-economic measure, which is an argument in agreement with choosing peers. If peers are actively selected, it makes little sense to try holding the peers constant through socio-economic controls if we want to find the price premium.

2.2 The Price of School Quality

Despite the various specifications and methods, the effect of school quality is almost always strongly positively correlated with house prices, when housing is linked to a school endowment. I will here refer to a number of estimates from the literature. An often comparable measure is that of crossing a border between a low and a high quality school district, and the implicated price premium on an average sample house. I will go through some often cited studies from US cities whose estimates are quite coherent, and then expand to a UK and closer institutional contexts as well as taking into account other specifications and methods.

Through utilizing observable border discontinuities, Black (1999) estimates that a 5 percent increase in scores corresponds to a 2.5 percent increase in house prices when estimating the effects of school districts in Massachusetts Boston suburbs between 1993–1995 (corresponding to $5,452 at mean house value). The paper is often cited for its method of finding similar house pairs at each side of a school district border, and it makes an argument for a Difference-in-Difference approach reduces the bias from neighborhood characteristics. The argument is that neighborhood effects and residential sorting follow a continuous function of distance, whereas school district effects are discrete—thus the finding of a discrete jump implicates the finding of a school district effect.

The estimate of Brasington and Hauring (2006) is similar, using average district proficiency test scores and expenditures to analyze the VA for a year 2000 cross section of seven urban areas in Ohio, spanning 77,578 houses within 310 school districts. They find that an increase in test scores by one standard deviation, ceteris paribus, raise house prices by 7.1 percent.

Studying data from Mecklenburg, North Carolina, between 1994–2001, Kane et al. (2006) also produces estimates comparable to that of Black (1999) when they observe similar discontinuous boundary effects: a 10 percent house premium attached with one SD difference in school mean test score. They conclude, however, with stating that the effect on prices seems to be largely indirect, through the residential sorting.

Haurin and Brasington (1996) find even larger premiums when comparing their intracity estimate to a reference district. Their data includes 134 districts within six US metropolitan areas. When they compare two identical houses, the premium for a 2 SD increase above sample mean in their test score variable, is an 18 percent house value increase ($13,815).

Chiodo et al. (2010) expands the hedonic house pricing model with a non-linear school term and finds that the top schools commands the highest premium and therefore argues that a quadratic model gives a closer fit. From cross-sectional data of 38,656 single family homes sold between 1998–2001 in St Louis, Missouri, they generate a range of estimates both through pairing houses at school boundaries as well as using the full sample. For comparison their estimated premium for a 1 SD increase in school quality as measured through test scores results in a premium of 10.87 percent (about $16,000) at the mean house price.

Cheshire and Sheppard (2004) studies Reading, situated 35 miles west of London, during

1999/2000, and find that quality of both secondary and primary school quality is substantially capitalized into house prices. They estimate that moving from the worst to the best possible secondary school would increase average house value by 18.7 percent (£23,763) and ditto for primary by 33.5 percent (£42,541).

Gibbons and Machin (2003) use cross-sectional data from England and Wales and through utilizing some methods from Black (1999) above finds similar order effects and that a 10 percent increase in pupils reaching a certain target level, is linked to a property premium of 3.6 percent (which equals between £4,500–£13,500).

Rosenthal (2003) uses data of nearly 150,000 house sales in England from the three school years 1995/6, 1996/7 and 1997/8. The instrumental approach of using assumed exogenous grade increases following random Ofstedt checks, as mentioned above, generates an estimate of dwelling purchase price elasticity, with respect to secondary school exam pass rates in England, to be around .05. This is a substantial reduction when compared to the naïve and presumed more biased non-instrumented estimate.

Through the natural experiment resulting from suppressed capitalization of house values until liberalization in 2000 in Shanghai, the main result by Li (2012) is that prices on average increase 41.9 percent more in the top-tier school districts over the following 8 years after reforms. It is not entirely clear how they argue that this effect does not co-variate with other kinds of residential sorting.

Most papers implicitly assume the effect is symmetric around zero, and this seems to be confirmed by Bogarth and Cromwell (2000) who show that the disruption of school districts 1987 in Ohio Shaker Heights reduces home values by 9.9 percent ($5,738) at mean house value, an estimate robust to alternative specifications and repeat sales analysis.

When studies are expanded to similar legislatures outside of the English speaking, a smaller estimated effect of school quality is calculated for Paris, through analyzing almost 200,000 house sales during the period 1997–2004. When increasing public school performance by one SD, an increase in house prices by 1.4–2.4 percent follows (Fack and Grenet, 2010). The study is further interesting for a Swedish context, as the school system for upper secondary in Paris is similar to that of Sweden, in mixing state schooling assigned through districts with a private school system lacking geographical assignment. They subsequently show how the saturation of private schools within an area reduces the school price premium effect.

In the neighbor capital of Helsinki, Finland, Harjunen et al. (2014) has taken an approach much similar to that of Black and find accordingly similar estimates of 2.5 percent premium following a one standard deviation change in 6^{th}-grade test scores between 2008–2012. They argue however that they mostly capture the preference for peers rather than school quality per se.

For Norway, Fiva and Kirkebøen (2008) find a temporary effect of around two percent through utilizing a value added measure fixed effects model through an identification strategy taking into account difference in information sets studying 79,322 observations from Oslo between 2003–2006.

The studies referred to identify school district effects relevant to a Swedish context. I therefore argue that, for example, the US charter schools fall outside of this study. Indeed their effect seem more ambiguous, and they can have a range of different price effects within the US system that are not transferable. A working paper by Imberman et al. (2015) studying sales within Los Angeles County between 2008–2011 show that increased quality or penetration of charter schools can have zero or even negative effects on house prices.

2.3 School Admission Policy in Sweden and Stockholm

School is mandatory for Swedish children from the age of 7, running through the primary school grades 1–9. Admission to municipal primary schools in Sweden is allocated on the municipal level according to state level established general principles. During the period of interest for this study, the specifics of this allocation was delegated to the municipalities. Alongside the municipally run primary schools, a parallel system of voucher funded private 'free schools' exists that generally operate through a queue on first-come-first-serve basis. This system was introduced in the 90ies and have gradually increased its proportion of the student body, and the approximate rate of free school penetration for Stockholm municipality over time can be seen in Figure 1. The levels are generally above those of other municipalities, and the proportion of municipality students in private schools have more than doubled during the period from a low at 10 percent, to closer to 25 percent at the end.

The Stockholm family primary school choice therefore stands between municipality schools, where the choice of residence is very important for application—and the free school system, operating through open queues with no consideration for family resident location.

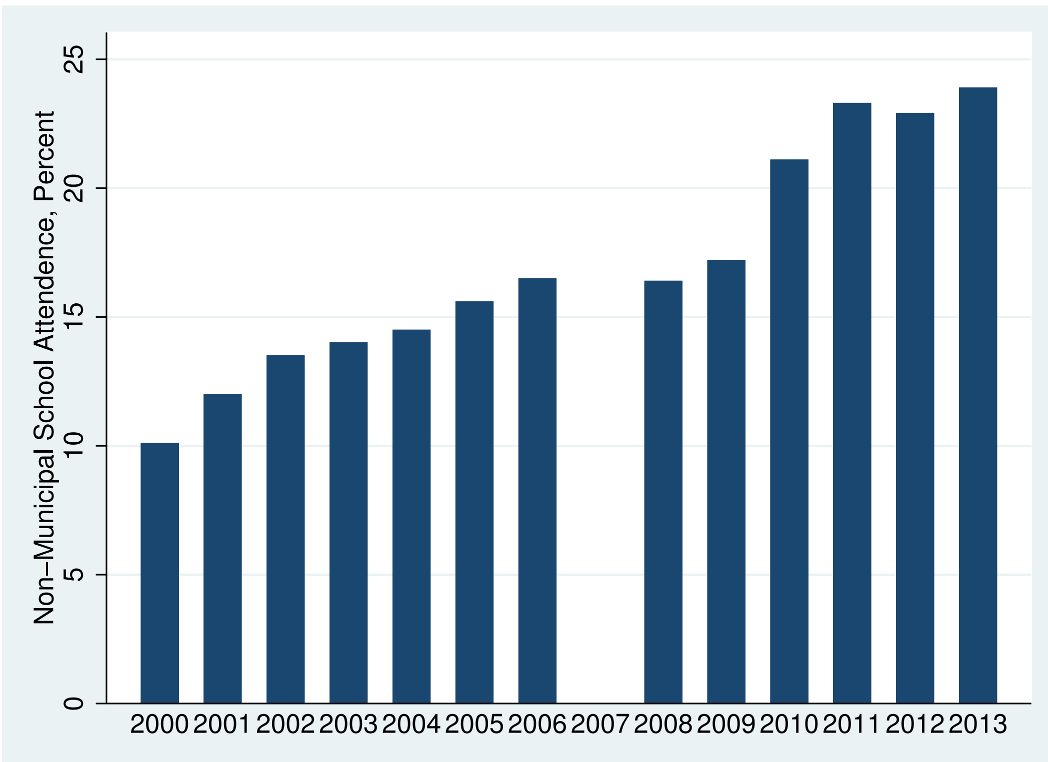
The Stockholm municipality specifics are of delegation one further level, to the level of city district (*stadsdel*). Each *stadsdel* was responsible for all students within the district, dividing it in turn further into *skolpliktsområden*, since all districts contain several schools. Each skolpliktsområde, hereby referred to as school district, links a number of addresses in a defined geographical area to a specific primary school. From this follows that each city area is divided into school districts that do not overlap city area borders, giving that all city area administrative borders coincide with school district borders. Please refer to Figure 2 for a schematic geocode rendering of school districts for one of the city districts of Stockholm (Bromma).

The school district does not, in principle, grant its residents an endowment to that particular school. However, it gives that particular school responsibility for the (compulsory) schooling of all its residents—which often in practice amounts to accepting the students even during oversubscription or demographic shocks. Only in third and fourth hand can other city district students, or even other municipalities', be considered. From interviews with headmasters and Stockholm municipality civil servants responsible, the general perception among residents have been that of endowment to a particular school. The consequence, I argue, has been that through the period covering at least 2000–2013, Stockholm municipality has run what is in principle a rather strict school district system similar to that of the UK or US.

This is surprising in two senses. First that the system has been generally more rigid than most have considered. A study performed by Stockholm municipality in 2012 could show that for the districts with the highest graded students, two thirds of the students (66 percent in Bromma)[2] study in their assigned school district municipal school (USK, 2012). Second, the UK system for example whose rigidity is taken as a given for many of the studies conducted is not hermetic and Cheshire and Sheppard (2004) shows that during 1997–2000, the success rate of appeals against school allocation was 23 percent, and for certain regions more than 50 percent. These two facts, I argue, make the systems comparable.

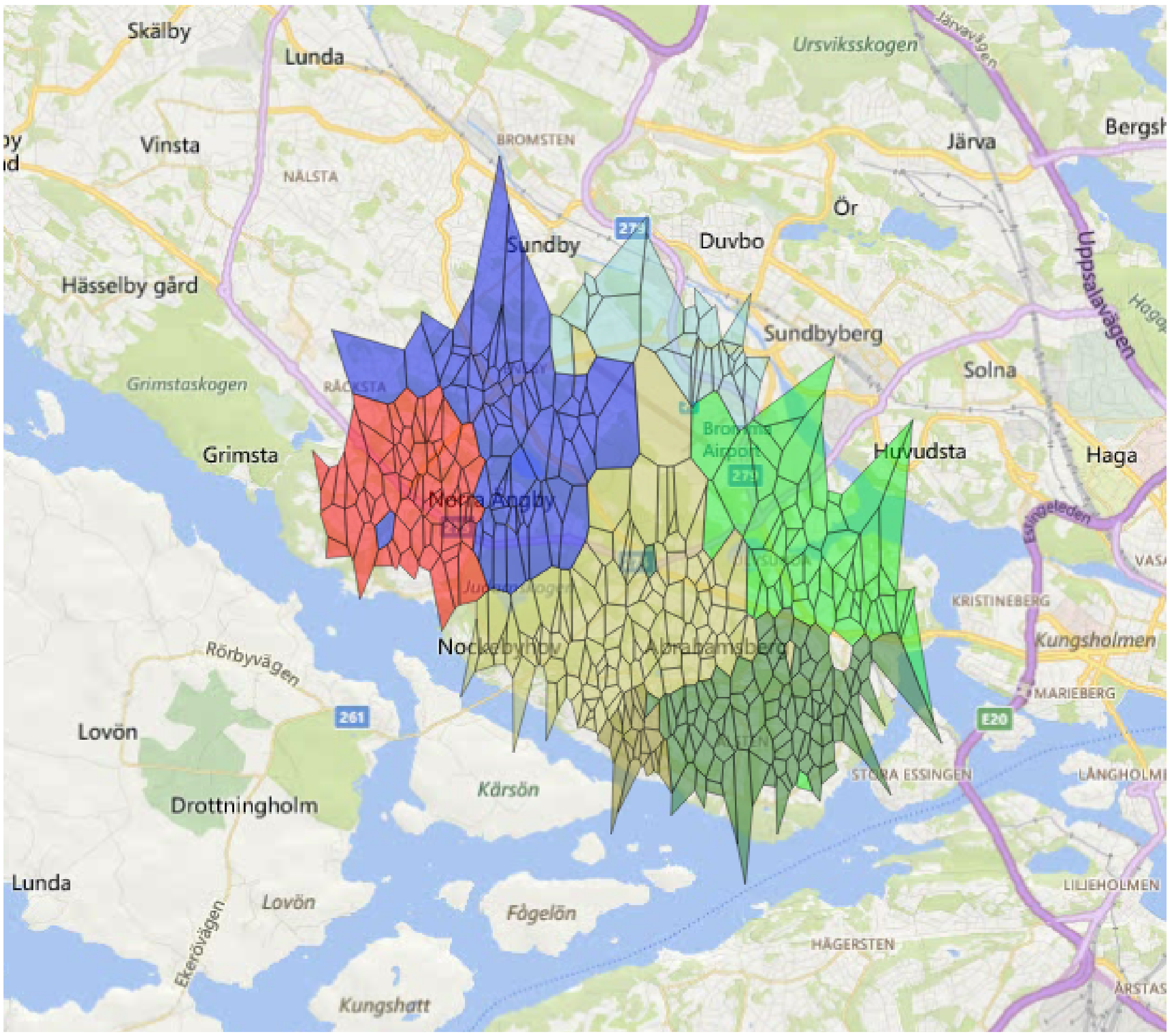
[2]Consider also the average free school penetration at this time.

Figure 1: **Proportion of Stockholm Municipality Students in Private or Non-Municipal Pre- and Primary Schools**, by Year



Data compiled from publicly available Skolverket data fetched in 2016 and rendered by author. Note missing value for 2007. Figures includes primary schools and pupils studying in other municipalities, due to lack of quality aggregate statistics. This does not have any large effect on the general trend since the number of primary school pupils are much lower, the number of extra-municipal students is negligible, and both display the same proportions and trend.

Figure 2: **Approximate Rendering of School Districts Within Bromma City District**



Districts rendered by author using Qgis geographic software package, from addresses for the school year 2011/2012. (Stockholm Stad 2016) Each polygon represents one address cluster from the school district ledger, color coded by school, imposed on Google map (acquired 2016).

3 Research Question

3.1 Theoretical Predictions

1. **School quality affect house prices in Stockholm municipality through a district endowment effect that generate premium positively correlated with quality or peers.**

   I expect that the general finding from the literature above will also be found within Swedish data, and that this result will persist through different identification techniques. In Section 2.3 I argue as to the similarities between the UK system and Stockholm municipality, giving that similar estimates would not be unreasonable to expect.

2. **The premium is diluted during the period through free schools and reduced during the period.**

   Following the results of Fack and Grenet (2010) regarding private school penetration, and the reasoning of several UK authors regarding that the tuition fees of private schools should be cap for school district premium (Cheshire and Sheppard, 2004), I expect the district premium effect to become smaller over the period studied.

Please refer to Section 5.5 for hypothesizes following model specifications.

4 Data

Housing price data from 2001–2013 has been purchased from the government agency Lantmäteriet and covers all transactions of detached housing during the period. This excludes /multi family homes/ apartments and certain detached houses in the *bostadsrätt* category. The data includes controls as listed in Table 1. Price is expressed in thousand SEK and the *standard sum* is a value reflecting the quality of the house. The observations are assigned coordinates and string block-names that will construct a locality control, as seen in Table 2. Last, the building year is assigned dummies for decade as seen in Table 5. For number of observations each year please refer to Table 3. For number of times a unique estate is re-sold during the period please refer to Table 4. Mean prices by stadsdel and year can be find in Appendix Table 9. The data has been filtered for extreme and erroneous values such as symbolic transaction at 1 SEK as well as a handful of multi-hundred MSEK transactions within the city core that are obviously outside of the focus for this thesis. The price data will, last, be month demeaned for all specifications to reduce noise.

Table 1: **House Price Controls**, Stockholm 2000–2013

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Price	2,505.12	(2,164.99)	45	14,875	33,381
Property area in m^2	536.85	(297.21)	9	6,940	33,381
Standard sum	27.72	(4.59)	6	52	33,060
Living space in m^2	115.8	(37.87)	10	732	33,060
Total areal in m^2	536.49	(291.44)	42	4,256	33,381
Nr of estates (N for > 1)	N=53	–	1	2	33,381

Lantmäteriet 2014, calculations by author.

4.1 Schools

School data is freely available data from the Swedish school government agency, Skolverket, and covers the same period 2000–2013 (since I will lag the grades one year). For grades by stadsdel and number of schools in each stadsdel, please refer to Table 6. The grades are final level grade, and a school average of a possible range 0–320 where 320 is the theoretical max if all students receive top grades. The grades are published for each school and I will calculate an unweighted mean [3] for each district, under the premise that a) schools are similar in size and the assumption b) parents observe school grades, but do not apply calculus on size given that a) holds.

To the school data I map certain other publicly available data: the coordinates for each school in the sample and the mean income for each stadsdel (please refer to Table 10) over time available freely from any map and from Statistics Sweden, respectively.

Table 2: **Sales per House Block**

Block Sale Statistic	**Observations**
Total Block N	2,729
Avg. Sales per Block	51
Median Sales per Block	41
Min. Sales per Block	1
Max. Sales per Block	216

Lantmäteriet 2014, calculations by author.

Regarding grade inflation, which is something that has been readily discussed in Sweden, I can see no signs of this during the sample period. The grade data set has the same high and low levels, and the same span between the top and the bottom city districts during the entire period. The maximum year difference in grades is 63.8 points.

I last use primary schools within the entire study since the policy has been most consistent and transparent over time, when compared to secondary schools that during this period has seen much more different policies. Arguably primary schools also has a much more clear cut price effect since it arguably affects families decisions to a much higher degree due to longer time span and potentially multiple children, than the schooling between 15–18 years of age.

The data on school district I do not judge being of sufficient quality to include though. Partly this can be seen in Figure 2.3 for a city district subset during a particular year. The primary issue is not the amount of work required to clean and check the data properly (it is not obvious whether the contrasting islands inside other districts are actually district discontinuities or mis-specifications of the district ledger or the geo-algorithm I have used to extract the coordinates, for example). Rather it is due to the sparse availability of this data during the period studied. The only school district ledgers I have been able to find locally[4] covers one or two years during the period studied. In part, it is also that interviews reveal they have reportedly have been applied with some discretionary flexibility by the schools, during parts of the period, to cope with variation from demographic shocks and changes. Last it is not clear how well known or available the information has been. The consequence, I argue, is that the city districts are the more appropriate borders to use despite a

[3] $yearaverage_{jt} = \frac{\sum_1^n grade_{ts}}{n_{jt}}$ where $grade_i$ is the published final grades for schools s in city district j with N schools at time t.

[4] Centrally the ledgers exist only for the most recent year.

Table 3: **Number of House Sales**, by Year

Year	**Observations**
2001	7580
2002	3190
2003	2359
2004	2585
2005	2622
2006	2572
2007	2220
2008	1942
2009	2437
2010	2189
2011	1652
2012	1736
2013	297

Lantmäteriet 2014, calculations by author.

Table 4: **House Resales**

Resales	**Observations**
0	15,247
1	12,276
2	4,698
3	1,040
4	120

Lantmäteriet 2014, calculations by author.

Table 5: **House Sale**, by Construction Decade

Decade	Observations
pre-1900	17
1900–1909	523
1910–1919	733
1920–1929	4,266
1930–1939	6,026
1940–1949	4,356
1950–1959	3,927
1960–1969	5,716
1970–1979	5,127
1980–1989	1,390
1990–	683
Missing	617
Total	33,381

Lantmäteriet 2014, calculations by author.

Table 6: **Grades and Number of Municipal Schools**, by Stadsdel

Stadsdel	Average grade over period	Number of schools
Bromma	237	5–6
Hägersten-Liljeholmen	214	4–6
Spånga-Tensta	178	4
Östermalm	247	2
Hässelby-Vällingby	211	5–7
Rinkeby-Kista	181	5
Kungsholmen	225	2
Norrmalm	250	4
Södermalm	222	6–10
Skärholmen	198	4–5
Älvsjö	215	2
Enskede-Årsta-Vantör	207	6
Skarpnäck	206	3–4
Farsta	199	3

Skolverket 2016, calculations by author.

loss in variation since a) they have been constant over time, b) more rigidly complied with and c) due to their transparency more readily observed by parents.

5 Empirical Strategy

As proposed in Section 2.1.1, the key challenge of the empirical strategy is that of addressing the endogeneity of the reseach object. At a first stage, the detailed data on locality through block variables, will be of importance. Previous authors have sometimes disagreed as to which control variables to include, giving different—sometimes seemingly arbitrary constructed—control vectors. Following the Cheshire and Sheppard (2004) critique of the popular Black (1999) approach for not being able to control enough for locality, I will be able to estimate a more fine grained fixed block effects model. This gives me a chance to pick up many unobservables such as local amenities like restaurants and parks or dis-amenities such as noise and nearby traffic.

Second, the use of the time variance of grades utilizes the fact that this variation is larger than most other changes at a local level. The reason for utilizing absolute levels rather than changes between years (as done by Jackson et al. (2014) for example), is due to the previous finding that parental residence school preference is to a large extent based on peers. This makes it important to include absolute levels.

Since we are most interested in variation at the district borders, following the assumption made by Black (1999) and many others that residential sorting is smooth whereas school district borders create discrete discontinuities, it is relevant to give the border observations a larger weight. I make an attempt to do so when specifying the second "Distance model" below. Through using an algorithm picking up the best grade within a certain range,[5] and a dummy for whether that (school) grade is within the district, and interacting the two, the probability of being included into the estimation increases as one approaches the border as a trigonometric function of the circle radius. A schematic graph of this connection can be seen in Figure 3. Note that for the distribution to follow Figure 3 for all districts j, schools must be randomly distributed geographically which is not the case. However we do know that schools and grades do express a geographical spread, so through this approach I can find border discontinuities—and I properly separate the effect of school distance from that of district.

The quality of the general house controls, and the extent to which localities affect the estimates, can further be verified through a fixed effects model on house level due to the resales during the period. The assumption here is that the general characteristics of the house has not changed,[6] giving that such qualities as view, compass direction, and similar traits, that we do not have house controls, for will also be absorbed. If these qualities explain a large part of the variation in price, the house fixed effects estimate will deviate from the full sample specification.

5.1 Econometric Specifications

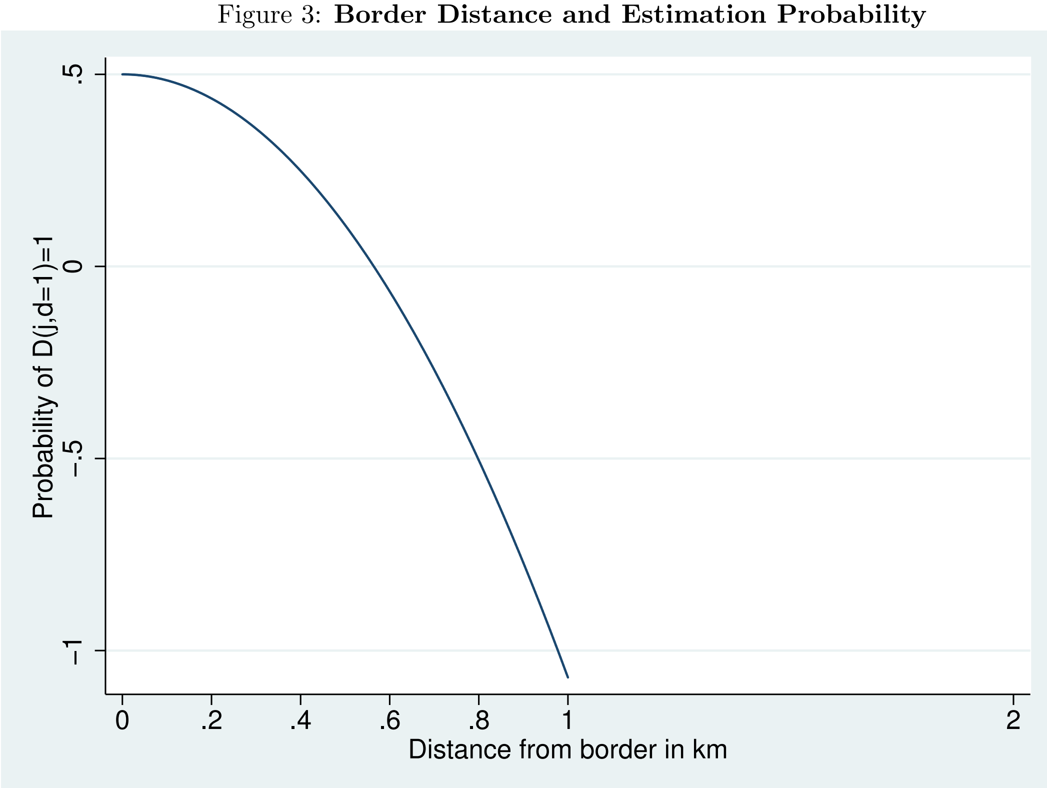
$$\log p_{ijtb} = \alpha + \beta_k x_i + \zeta_l z_{bj} + \rho \times inc_{jt} + \gamma \times school_{j,(t-1)} + \delta_t T + \theta t + u_{it} \tag{1}$$

The month demeaned price, p of detached house i in Stockholm municipality city district j at time t in block b, is determined by a set of vectors. First a vector x of k house specific variables.

[5] The algorithm is coded to find all schools within a specified radius, and then return the grade of the highest graded school among them.

[6] On the one hand we can expect such things as renovation—but on the other we will also see depreciation.

Figure 3: **Border Distance and Estimation Probability**



Assumes random geographical distribution of schools. Rendered by author using Stata.

Second, a vector z of l spatial locality controls for neighborhood characteristics and localities. Third, an income control for each region. Fourth follows the lagged school quality specifications.[7] Last, yearly time controls and a general time trend. Last an error terms where $E(u) = 0$.

5.2 Choice of Variables

Building on previous research regarding the use of *value added*-variables, and following from that within-municipality variance of inputs is negligible during the period, this study will utilize variation in output variables. Final grades, that are readily observable, is the preferred choice of independent variable.

For a first stage, not presented here, I test both an average district variable as discussed above as well as a yearly max, which is constructed simply by taking the highest grade among the schools within the city district for that particular year. I conclude that they are very correlated (corr = .8113) and that they produce similar estimates. The economic interpretation of average grade is more clear, as well as providing a more plausible variation through reducing the effect of outliers from single year single school variation affecting the district estimate.

I use the grade lagged from the previous year (price at t and grades at $t-1$) since the simultaneous year grades are naturally biased by the characteristics of resident inflow and not observed at the time of purchase so can not be seen as exogenous. I thus assume that grades are observed the year before purchase. Since the average grades for each district do vary, but correlate positively over time, I have chosen to exclude further lags to reduce the complexity of the model.

To test my second hypothesis I have a choice between splitting the sample into yearly cross-sections or construct an interaction between the general time trend and the average for each district. Since the first method would give a lot of estimates and not utilize the entire sample, I will begin with testing a time trend that can then be further analyzed with specific intervals or years if relevant.

5.3 Average School District Effect

Similar to most literature we are first interested in the school district effect:

$$\log p_{ijtb} = \gamma_1 average_{jt} + \Pi_{jtb} + v_{it} \tag{2}$$

Where: $\Pi_{it} + v_{it}$ is the control vectors and error terms as defined in equation 1; $average$ is the average final grades for the municipal schools within city district j at time t. To test the second hypothesis of how this estimate changes over time I also include an interaction term between the time trend and the yearly average:[8]

$$\log p_{ijtb} = \gamma_1'' average_{jt} + \Phi t \times average + \Pi_{jtb} + v_{it}'' \tag{3}$$

Utilizing the large amount of repeated sales (Table 4) I can also estimate a fixed effect specification on house level. The sample is thus restricted to all observations where resales > 1, which is 18,134 observations as seen in Table 4.

$$\log p_{ijt} = \gamma_1' \overline{average}_{jt} + \Pi_{jt}' + v_{it}' \tag{4}$$

[7]Note that for ease of notation, I normalize t in the following specifications so that $t_{school} = (t+1)_{price}$.

[8]Note that the t-term that is interacted with $average$ is included in the Π-vector

Where Π' is a reduced control vector, interpreted in the same way as in equation 1, but lacking controls for city district and house qualities since they are absorbed by the specification:

$$\Pi' = \rho' \times inc_{jt} + \delta_t' T + \theta' t \tag{5}$$

5.4 Distance or Endowment

Central for our specification is finding the difference between distance and endowment. One source of endogeneity could be captured with the distance itself to quality schools—and we could also expect to pick up some of the both endowment effect (when testing only distance) and the convenience effect (reflecting time and transportation costs).

To achieve this I map each house sale to its city district. I then include the best school, as measured with final grades (*grademax*), within radiuses .5–2.0 km with .5 km increments. Coordinates allows me to discern whether the best school for each increment is *within* the school district or not—to address the question of whether proximity has an added endowment value for entitlement proximity.

I argue that a reasonable distance specification is 1.0 km, running detailed specifications to test robustness for this distance and estimate distances in both directions in Table 8. When approaching the borders (d=.5) we lose statistical power since the number of observations $D > 0$ is reduced, as well as variation from the schools placed more than .5km from the border. 1 km corresponds roughly to the average diameter of a school district, but I test for the distances around school district size to show plausibility of the estimates. The reason for not using the de facto school district coding is discussed in Section 4.

$$\log p_{ijtb} = \sigma_{1d} grademax_{dts} + \sigma_{2d} D_{djts} + \sigma_{3d} D_{dj} \times grademax_{dts} + \Pi_{jtb} + w_{it} \tag{6}$$

Where $\Pi_{jt} + w_{it}$ is the control vector and error term as defined in equation 1, max_{dts} is the highest graded school s within distance d at time t, D is a dummy for whether the max_d-school s is also within the same school district as house i as well as a municipal school with district allocation. Last an interaction term that will take the same value as max_{dts} if the school s is within the same district district as house i. This specification will then be repeated with 500 meter increments for d= .5–2 km.

I include an interaction effect similar to that in the average specification above, adding $t \times grademax_{dts}$ to the right hand side of equation 6:

$$\log p_{ijt} = \sigma''_{1d} grademax_{dts} + \sigma''_{2d} D_{djts} + \sigma''_{3d} D_{dj} \times grademax_{dts} + \Pi_{it} + \Phi^* [t \times grademax_{dts}] + w''_{it} \tag{7}$$

Last I estimate the Distance specification using house fixed effects following the Average specification fixed effects description above:

$$\log p_{ijt} = \sigma'_{1d} \overline{grademax}_{dts} + \sigma'_{2d} D_{dj} + \sigma'_{3d} D_{dj} \times \overline{grademax}_{dts} + \Pi'_{jt} + w'_{it} \tag{8}$$

5.5 Hypothesis Specifications

1. **Positive house price premium.**

District Average model from equation 2:

H_0: $lyearaverage > 0$
H_1: $lyearaverage \leq 0$

Distance model from equation 6:

H_0: $(\sigma_{2d} + \sigma_{3d}) > 0 \mid grademax =$ in sample
H_1: $(\sigma_{2d} + \sigma_{3d}) \leq 0 \mid grademax =$ in sample

2. **Reduced premium over time from free school penetration.**

District Average model from equation 3:

H_0: $\Phi < 0$
H_1: $\Phi \geq 0$

Distance model from equation 7:

H_0: $\Phi^* < 0$
H_1: $\Phi^* \geq 0$

6 Results

6.1 School District Estimates

I begin to run the estimates without the full vector of controls to get a picture of the level of bias that can be expected. When using only year dummies the grade estimates probably capture neighborhood and district effects for each city district and block, including locality effects.

The price effect of one increased point in average school district grade lies at .05 percent per grade as seen in the Table 7, column 1. Put into sample perspective this corresponds to a 21 percent price premium ($37 \times \beta_1 = .208$) when moving from a mean average grade city district (grade=211) to the highest average grade district (grade=248). The maximum price premium when moving from the bottom district to the highest average district during a particular year ($\Delta = 64$) is 36 percent. In SEK for an average house in the sample during the period, this corresponds to respectively TSEK 521 and TSEK 901 ($63,000 and $91,000 at present exchange rate).

Note also that all the estimates go in the expected directions along with the school district premium, giving significant explanatory value to house quality and size measures. Please refer to the appendix and Table 11 for a full output Table, including other house variables such as construction decade.

When adding fixed effects for city districts, in column 2 of Table 7, the effect of average grade is actually increased somewhat by .085 extra marginal effect per average grade change. Adding controls for city district income reduces the estimate somewhat as is expected (column 3) but the effect is still of similar magnitude.

In column 4, I add an interaction variable between the general time trend and the yearly average. It becomes significant with a positive sign, at a lower magnitude than the previous lagged average estimates—rendered insignificant in this specification.

Last, the same specification but using the full set of locality controls in column 5 of Table 7. The estimate for average grade is reduced somewhat but still remains at the same order of magnitude and corresponds to an in-sample effect of 15.3 percent (Δ = TSEK 383 = \$46,000) and 26.6 percent (Δ =TSEK 666=\$808,000) at average city district to max, and bottom to max, respectively. The estimates for most of the house quality variables are also reduced somewhat. I last utilize the subsample repeated sales variation for the house fixed effects specification in column 6. The estimates are similar to the full sample regressions, but note that I am not able to use district fixed effects at the same time.

All specifications are all regressed using Stata's robust estimation option. I also test with clustering on an interaction between year and school district to get the proper amount of clustering units, and this has no effect on estimates and only increases standard errors. The reason for not using this clustering as base line is that I want to use the same clustering for the full block dummy specification. For the latter, I can not combine year and district clustering due to requiring too many degrees of freedom.

For the District average specification, I can therefore not reject *hypothesis 1* and there is a significant effect of grades upon house prices both economically and statistically. I do, however, reject *hypothesis 2* since the time trend of the lagged yearly average estimate is not negative—but instead displayed the opposite sign.

6.2 The Price of Endowment

A first test was run (not shown) for significance of the *max* variable without the interactions, giving an estimate for value of having the best school within each increment independent of it being within the same district. These estimation displayed highly insignificant results for all specifications and distances.

The interpretation of $\sigma_{21} + \sigma_{31}$ should be that of value added (subtracted) per school s being *within* the city district ($j_i = j_s$), as depending on the previous year t grade of the best school s within 1 km. From column 2 follows that the school itself in the interacted specification does not seem to have an effect as different from zero. However, if the same school is also within the same district as the house, the effect is negative (following the negative intercept $-.33$ for $D = 1$) at lower school grades. This effect turns positive from the interaction effect for higher grades (>229), with a marginal effect for each grade increase of the best school of .146 percent.

This gives that for a house within 1 km of a best school in the radius, whose lagged final grades exceed 229 points, the marginal effect of school grades upon the house price is .146 percent—at an in-sample maximum (grade=283) of $54 \times \sigma_{31} = 7.8$ percent. This corresponds to a house price premium of SEK197,500 (\$23,900) at the average sample house price.

Estimates around the most plausible identification distance at 1km behaves as expected: the estimates for shorter and longer distances are similar to the 1km specification estimates. It can be noted that the positive effect of the interaction variables occur at roughly similar grades: $(\sigma_{2,d=0.5} + \sigma_{3,d=0.5} \times max) > 0 \mid max \geq 243$ and ditto $(\sigma_{2,d=1.5} + \sigma_{3,d=1.5} \times max) > 0 \mid max \geq 220$. Note also that the interaction effect becomes insignificant at d=2, which is a distance exceeding the span of most school districts.

Last I also test a similar time trend as in the Average specification (not shown) but it is not significant at any statistical level and does not affect the other estimates.

I conclude noting I can not reject *hypothesis 1* in this specification either: the added value of the best school being in city district for in-sample values of the interaction variable for *max* is positive. However, for all significance levels *hypotesis 2* can be rejected in this specification as not different from 0.

Table 7: **District Average Specification**, Estimation Results

$y = \log(price_{itj})$	(1) Base	(2) Districts	(3) Inc	(4) Time	(5) Full	(6) F.E.
standard sum	0.0128	0.0112	0.0114	0.0113	0.00742	
	$(10.29)^{***}$	$(9.04)^{***}$	$(9.23)^{***}$	$(9.18)^{***}$	$(5.90)^{***}$	
areal, m^2	0.000765	0.000713	0.000712	0.000714	0.000302	
	$(35.01)^{***}$	$(31.16)^{***}$	$(31.17)^{***}$	$(31.25)^{***}$	$(8.15)^{***}$	
avg. grade	0.00563	0.00648	0.00500	0.00205	0.00416	0.00669
	$(16.69)^{***}$	$(6.78)^{***}$	$(5.02)^{***}$	(1.48)	$(4.10)^{***}$	$(3.39)^{***}$
avg. income			0.00000364	0.00000137	0.00000499	0.00000545
			$(5.98)^{***}$	(1.46)	$(7.73)^{***}$	$(4.97)^{***}$
avg. grade × t				0.000412		
				$(3.31)^{***}$		
N	33060	33060	33060	33060	33060	18134
R^2	0.551	0.560	0.560	0.560	0.623	0.462

t statistics in parentheses

Both avg. grade and income on city district level, with grades lagged one year. Specification (1) controls for time dummies and trend (see table 11 for detailed control estimates) and all house controls; (2) adds city district fixed effects; (3) adds city district avg. yearly income; (4) a time interaction variable between time trend and avg. grade; (5) is using the full set of block locality controls (N=2,729) and (6) uses fixed effects on house level through a restricted resale sample.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: **Distance Specification**, Estimation Results

$y = \log(price_{itj})$	(1) 0.5 km	(2) 1.0 km	(3) 1.5 km	(4) 2.0 km	(5) Full, 1.0 km	(6) F.E., 1.0 km
max0.5km	0.0000889					
(grade, σ_1)	(0.42)					
district0.5	-0.452					
(dummy, σ_2)	(-3.83)***					
district0.5 $\times$ grade	0.00186					
(interaction, σ_3)	(3.32)***					
max1km		-0.0000433			0.000235	0.000285
		(-0.49)			(1.28)	(0.83)
district1		-0.334			-0.209	-0.860
		(-4.35)***			(-1.62)	(-3.57)***
district1 $\times$ grade		0.00146			0.000788	0.00366
		(4.13)***			(1.27)	(3.22)**
max1.5km			-0.0000565			
			(-0.48)			
district1.5			-0.225			
			(-2.56)*			
district1.5 $\times$ grade			0.00102			
			(2.58)**			
max2km				0.000297		
				(1.20)		
district2				-0.0372		
				(-0.33)		
district2 $\times$ grade				0.000152		
				(0.30)		
N	33060	33060	33060	33060	33060	18134
R^2	0.560	0.560	0.560	0.560	0.623	0.462

t statistics in parentheses
Where: (1) through (4) estimates for different distance specifications; (5) repeats the 1km-estimation using the full set of block locality controls and (6) uses fixed effects on house level through a restricted resale sample.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7 Discussion

The estimates of the economic effect of average city district grade upon house prices are very high, which is in line with some previous studies such as Haurin and Brasington (1996). However the estimates seem unreasonably high and probably reflect the endogeneous nature of the regressors. As this is performed using a very good set of control variables, it puts some doubt on previous literature's estimates where they have not managed to ensure proper identification: the indication is that bias stem from residential sorting almost exclusively rather than house, block or district characteristics.

One source of overestimation could be to the fact that we only take detached housing into account. It is common to leave the central city and higher square meter priced multi-family housing when increasing family size. This would focus the school quality premium towards the single family houses in the suburb that are represented in the sample, and is something to bear in mind when discussing external validity of this study. Detached housing has been a large part of the sample in many similar studies however, since they are popular in large parts of both the UK and US, so comparison is still relevant.

The controls for house quality are seemingly of high quality, which can be confirmed by the fixed effects model for both specifications. The fixed effects estimates do not change notably from the full sample regressions. This points towards other sources of bias, since the fixed effects estimates are also similarly high as the basic specification. It also seems the income controls captures some sorting which is expected. The increased estimate effects from including the city districts is somewhat surprising though, and it is not clear what drives this.

School quality also seem to have some correlation with locality, which can be seen in the house quality estimates reduced somewhat in both specifications—along with the downward corrected estimates for school effect. The reduced estimate of quality could be a sign that a highly valued block (due to say neighborhood amenities, location or aesthetics) is positively correlated with house size and house standard sum value. That this would in turn correlate with sorting of residents not captured by mean income, as indicated by a slightly reduced school estimate for the full specification, is not unreasonable.

That the block controls renders the distance interaction effects insignificant in the second specification could be due to lack of statistical power: after all we now have more than 2,000 controls, including interaction effects. They do show the same signs and magnitudes however which is somewhat reassuring.

The positive effect from time interaction between average grades and time in the first specification is somewhat unexpected, but given that it turns insignificant in the improved second specification could hint at some general endogeneity of the first model. Therefore a zero effect seems more plausible, which is a interesting contrast to the Fack and Grenet (2010) study. One explanation would be that the private free schools are not seen as a perfect substitute to the high quality municipal schools. This is not an unreasonable hypothesis, since a substantial proportion of the Swedish free schools are niche schools of particular religious or pedagogic specialty. It could also reflect the fact that peer effects might take some time to build up, before the right students and the reputation is in place. The municipality schools' peer effect might therefore not be replicated by the more recently opened private schools, and they are seen as imperfect substitutes.

Note that the distance in itself to quality schools is not significant. This could be seen as somewhat strange, given that transport and time costs increase linearly for the families. It could still have an effect of correcting the interaction estimates, and it is possible that other specifications of distance might have a significant effect. For finding this, one could include such variables as travel time using the fastest mode of transport (as available by online map algorithms) or non-linear distance specifications.

So how much causality can we infer from these results? There is a clear, constant, value added of school districts for high quality schools that we only find at distances to the border corresponding with school district sizes. Any high quality school outside this radius is not relevant for the house price. It seems reasonable to argue that this captures a true school district effect. Endogeneity is clearly reduced when compared to the basic specification, and the model is clearly robust to all added levels of controls and fixed effects—along with all non-relevant estimates behaving properly and showing the expected signs.

I would last point towards the interesting negative effect from lower grade schools being in district. Some studies have observed that bad schools will reduce house prices—but note that it is only the effect of bad schools being in the school district that has a negative effect. Given that we still have a probable endogeneity issue from residential sorting, this gives rise to what Kane et al. (2006) refers to as a second order effect on prices. Our model is not able to distinguish this effect from a first order price effect following explicit demand. Rather, we also find that within each city district, centered around schools, residential sorting would include into our estimates the characteristics and the demand function of the residents. This could in turn explain the negative intercept: around the schools with lower grades we find residents that differ from those around higher grade schools.

This allows us to draw the conclusion that we have significantly improve upon a basic hedonic model when using a geographic specification. This augmented specification significantly improves identification, but some endogeneity in the form of residential sorting and second grade price effects, is likely to remain.

7.1 Conclusions

My estimated results are similar in magnitude to the studies from UK, US and France during the period, when the bias of a common hedonic model is reduced with an improved specification including a a continuous distance variable. The effect of added value from having the best nearby school within the district is positive for high grades (>219) and has an effect of .15 percentage points per final school grade the previous year—for a maximum in-sample premium of 7.8 percent which corresponds to SEK197,500 ($23,900) at the average house value.

I argue that we can draw some causal conclusions from this and that identification is significantly improved from a more basic hedonic model. However, an unknown level of bias is likely to remain. While the link between school quality and house prices might indeed be one of the strongest established in the field of economics, we do not quite know the magnitude—only that it probably lies somewhere between 1.5–35 percent.

Three important findings can be drawn from this study. For one, it further points towards school premiums being fairly constant across institutional contexts, since the results hold up well for Sweden despite the large differences to the US or UK. Second, the study confirms that the approach taken by most previous studies is likely deeply endogenous and generously upwardly biased. Despite the detail level of controls I can show through my correction of estimates when improving identification, that all remotely more naive specifications probably overestimate the house price premium to a large degree. Third, it indicates that residential sorting is the main source of bias. I am able to show that the quality of controls is very high, and even when controlling for individual blocks using a very fine grained specification, a large part of the bias remain. Socioeconomic controls reduce the estimates somewhat but do not seem to capture a large part of the sorting.

From a policy perspective, this paper indicates that school districts for top-level schools in Stockholm municipality have added a noticeable price premium on detached housing during 2001–2013. The premium persist despite the rapid increase in private free school penetration during the period. It seems that municipal schools—maybe by strength of age and heritage—have an attraction and status that the, generally recent, private free schools have not acquired. The consequence is that private free schools during this period, does not seem to be especially good substitutes for quality municipality schools during this period. Last, given that these results go in line with previous literature that shows peer quality is the most important component of school choice rather than value added, it is not clear what incentives choice and competition provides for the schools to improve. One recommendation would therefore be that despite school premiums, the rapid introduction of competitive free schools, is a hasty conclusion.

7.2 Further Studies

First, it would be natural to expand this study to include apartments, since we now lack transactions for the most central districts. This data is of lesser quality, since Sweden lacks a central register making the data available that gathered by private actors. Second, the field is open for gathering the yearly school district ledgers and constructing the algorithm mapping each address to its respective school, which would improve the border variation. It would also indicate whether this information set is well enough known among parents. A third relevant addition would be to include the present period, since the school district policy has been dissolved completely from the year 2016.

8 References

References

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics*, 114(2):577–599.

Bogarth, W. T. and Cromwell, B. A. (2000). How much is a neighborhood school worth? *Journal of Urban Economics*, 47:280–305.

Brasington, D. and Hauring, D. R. (2006). Educational outcomes and house values: A test of the value added approach. *Journal of Regional Science*, 46(2):245–268.

Brasington, D. M. (1999). Which measures of school quality does the housing market value? *Journal of Real Estate Research*, 18(3):395–413.

Ceccato, V. and Wilhelmsson, M. (2011). The impact of crime on apartment prices: Evidence from Stockholm, Sweden. *Geografiska Annaler: Series B*, 93(1):81–103.

Cheshire, P. and Sheppard, S. (2004). Capitalising the value of free schools: The impact of supply characteristics and uncertainty. *Economic Journal*, 114:F397–F424.

Chiodo, A. J., Hernandez-Murillo, R., and Owyang, M. T. (2010). Nonlinear effects of school quality on house prices. *Federal Reserve Bank of St. Louis Review*, 92(3):185–204.

Fack, G. and Grenet, J. (2010). When do better schools raise housing prices? evidence from paris public and private schools. *Journal of Public Economics*, 94:59–77.

Fiva, J. H. and Kirkebøen, L. J. (2008). Does the housing market react to new information on school quality? *Statistics Norway Discussion Papers no. 541*, Statistics Norway, Oslo.

Friedman, M. (1955). The role of government in education in economics and the public interest in (ed.) Solo, R.A. *Economics and the Public Interest*, Rutgers University Press, New Brunswick NJ.

Gibbons, S. and Machin, S. (2003). Valuing English primary schools. *Journal of Urban Economics*, 53(2):197–219.

Harjunen, O., Kortelainen, M., and Saarimaa, T. (2014). Best education money can buy? Capitalization of school quality in Finland. *Vatt Working Papers no. 58*, Government Institute for Economic Research, Helsinki.

Haurin, D. R. and Brasington, D. (1996). School quality and real house prices: Inter- and intrametropolitan effects. *Journal of Housing Economics*, 5:351–368.

Herath, S. and Maier, G. (2010). The hedonic price method in real estate and housing market research. A review of the literature. *SRE Discussion Papers no. 3*, WU Vienna University of Economics and Business, Vienna.

Imberman, S. A., Naretta, M., and O'Rourke, M. (2015). Capitalization of charter schools into residential property values. *NBER Working Paper No. 20990*, National Bureau of Economic Research, Cambridge MA.

Jackson, C. K., Johnson, R., and Persico, C. (2014). The effect of school finance reforms on the distribution of spending, academic achievement, and adult outcomes. *NBER Working Paper No 20118*, National Bureau of Economic Research, Cambridge MA.

Kane, T. J., Reigg, S. K., and Staiger, D. O. (2006). School quality, neighborhoods and housing prices. *American Law and Economics Review*, 8(2):183–212.

Li, W. (2012). Does school quality affect real estate prices? The effect of top-tier elementary schools on property prices in Shanghai. *Honors Projects*, Paper 44, Economics Department, Macalester College, Saint Paul MN.

Maslianskaia-Pautrel, M. and Baumont, C. (2016). The nature and impacts of environmental spillovers on housing prices: A spatial hedonic analysis. *Working Paper 2016.04*, FAERE, French Association of Environmental and Resource Economists.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55.

Rosenthal, L. (2003). The value of secondary school quality. *Oxford Bulletin of Economics and Statistics*, 65(3):329–355.

Rothstein, J. 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):1333–1353.

Sahlgren, G. H. and Jordahl, H. (2016). Information – ett verktyg för bättre skolsystem. SNS Förlag, Stockholm.

Söderström, M. and Uusitalo, R. (2010). School choice and segregation: Evidence from an admission reform. *Scandinavian Journal of Economics*, 112(1):55–76.

USK (2012). Elevpendling Stockholms Stad 2010–2011, Stockholms Stads utrednings- och statistikkontor. Technical report.

A Terms Used

Term	Explanation or Swedish Equivalent
City district	Stockholm municipality is divided into 14 *stadsdelsnämnd* as seen in Table 10, that have been responsible for primary school assignment.
Free school	Electable schools that are non municipal entities whose growth, parallel to the municipal school system, can be seen in Figure 1.
School district	For Stockholm municipality, each primary school is assigned a *skolpliktsområde* within which they are responsible for all students.

B Tables and Figures

Table 9: **Prices by Year and Stadsdel**, TSEK

Stadsdel	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Bromma	806	1992	3456	3328	3982	4155	5460	6236	4263	5916	6746	7235	7380
Hägersten-Liljeholmen	979	1737	2774	2774	3114	3325	4167	4285	3287	4677	5550	5664	5934
Spånga-Tensta	1252	1816	2399	2438	2694	2873	3583	3429	3334	3633	4112	4119	3750
Östermalm	-	-	-	-	-	-	-	-	-	-	-	-	-
Hässelby-Vällingby	649	1567	2024	2022	2254	2267	2894	2958	2381	2764	3369	3260	3185
Rinkeby-Kista	371	947	1520	1370	1205	1422	2084	2335	1526	2029	2396	2452	2583
Kungsholmen	2439	5588	5018	4105	5801	6509	8554	7900	8240	7573	10413	8982	-
Norrmalm	-	-	-	-	-	-	-	-	-	-	-	-	-
Södermalm	-	-	3300	6200	-	-	9500	6238	-	6000	7000	-	-
Skärholmen	467	891	1580	1351	1119	1611	2033	2414	1465	1960	2840	2652	2716
Älvsjö	848	1442	2151	2003	2184	2141	3139	3900	2745	3352	4456	4590	4711
Enskede-Årsta-Vantör	648	1618	2338	2506	2746	3230	4063	4493	3343	4429	5087	5056	4385
Skarpnäck	486	1058	1981	2096	2483	2343	3416	3961	2763	3095	4751	5156	5295
Farsta	416	811	1677	1523	1742	2060	2757	3305	2155	2862	3760	3932	3296

Lantmäteriet 2014, calculations by author.

Table 10: **Average Income and Population**, by Year and Stadsdel

Stadsdel	Year:	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Rinkeby-Kista	Pop:	31488	32179	32557	32838	32677	32497	33031	32841	32583	32796	33456	34214	34786	35221
	Inc.:	145178	153871	156942	158266	158185	158838	160693	167763	173792	173173	175283	181581	187808	192534
Spånga-Tensta	Pop:	23263	23431	23630	23656	23627	23786	24365	24750	25154	25678	26018	26434	26745	27188
	Inc.:	181275	191161	196670	201151	205485	209485	213374	222069	231284	231050	235700	242594	248767	253784
Hässelby-Vällingby	Pop:	43467	43808	43944	43856	43691	43973	44831	45653	46516	47082	48121	49333	50380	51374
	Inc.:	214907	224622	230003	232659	236702	240389	244521	251675	260417	261178	263210	268047	274272	279841
Bromma	Pop:	45476	45464	45453	45349	45433	45941	46283	47100	47858	48536	49325	50616	52621	53960
	Inc.:	261787	278796	283046	286151	293580	305634	321354	335702	349701	347516	354583	365404	373858	384239
Kungsholmen	Pop:	46315	46631	46679	46954	47454	47644	48761	49145	49475	50682	52259	53851	55860	57057
	Inc.:	243500	259867	263155	267150	276116	288698	301333	321202	337354	339916	348845	361600	374263	383843
Norrmalm	Pop:	52381	52377	52418	52771	52848	53228	53185	53142	54005	55009	55733	56292	56897	57316
	Inc.:	253229	270342	273619	279011	285588	297508	311819	334837	351653	354222	362985	378854	390343	400525
Östermalm	Pop:	52707	52695	52299	52199	51938	52130	52165	52331	53199	54324	55173	55919	56487	58008
	Inc.:	265320	282432	284623	284755	292727	306510	319305	343149	356371	356906	360616	375561	387219	397525
Södermalm	Pop:	86394	87125	88411	89693	90522	91693	92944	95430	98068	99601	100935	102232	103268	104427
	Inc.:	220660	235059	241367	246855	252916	262842	274600	286160	300348	303964	311319	323127	333771	341667
Enskede-Årsta-Vantör	Pop:	63944	64069	63943	63868	63848	64082	64854	66184	67695	69225	70349	71555	72786	73784
	Inc.:	189678	200742	205979	210759	214166	219448	226376	235994	246454	248871	253237	260963	269669	277783
Skarpnäck	Pop:	30582	30802	30729	30801	30801	30748	30907	31250	31749	32834	33516	33898	34450	34779
	Inc.:	188219	198253	203940	208494	212826	218739	225456	234544	245585	249024	254614	262746	271305	279541
Farsta	Pop:	15285	15259	15382	15360	15457	15834	16134	16435	17109	17721	18166	18533	19094	19679
	Inc.:	228022	239183	246989	253004	261676	270353	280393	295312	310308	311368	319428	329377	337104	344857
Älvsjö	Pop:	15285	15259	15382	15360	15457	15834	16134	16435	17109	17721	18166	18533	19094	19679
	Inc.:	228022	239183	246989	253004	261676	270353	280393	295312	310308	311368	319428	329377	337104	344857
Hägersten-Liljeholmen	Pop:	48947	49036	49106	49227	49639	49957	51204	52721	54039	57270	59274	61110	62805	64213
	Inc.:	200998	211666	217628	222280	228168	235553	247553	258026	270406	274981	281163	291791	302311	311609
Skärholmen	Pop:	22584	22980	22911	22669	22406	22347	23089	23527	23803	24151	24660	25290	25741	26181
	Inc.:	164613	172183	176687	180589	181450	182288	184300	188157	192466	191516	191778	196371	199955	204258

Statistics Sweden 2016, calculations by author.

Table 11: **District Average Specification**, Estimation Results Full Controls

	(1) avgfull
standard_sum	0.00742
	(5.90)***
areal_m_2	0.000302
	(8.15)***
nr_of_estates	-0.0470
	(-0.33)
living_space	0.00182
	(7.96)***
non_living_space	0.000172
	(1.41)
lyearaverage	0.00416
	(4.10)***
pre1900	-0.155
	(-0.63)
siecle	0.126
	(1.48)
tens	0.0628
	(0.76)
twenties	0.112
	(1.56)
thirties	0.0707
	(0.99)
forties	-0.146
	(-2.05)*
fifties	-0.107
	(-1.44)
sixties	-0.00627
	(-0.09)
seventies	0.00737
	(0.11)
eighties	0.0794
	(1.09)
new	0.0692
	(0.87)
t	0.131
	(20.02)***
N	33060
R^2	0.623

t statistics in parentheses

Variables $pre1900$–new are dummies for construction decade.

Year dummies not included here but for an overview of price development please refer to table 9.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$