university of groningen / faculty of spatial sciences

What is the Impact of School Proximity and Quality on Housing Prices in Allegheny County: Comparing Public and Private Education

by Skovolas Athanasios

S5089549

University of Groningen / Faculty of Spatial Sciences

MSc Real Estate Studies

Master Thesis

Supervisor: Xiaolong Liu

Disclaimer

"Master theses are preliminary materials to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the author and do not indicate concurrence by the supervisor or research staff."

COLOFON

Title	*What is the Impact of School Proximity and Quality on Housing Prices in Allegheny County: Comparing Public and Private Education.*
Author	Athanasios Skovolas
Student Number	S5089549
Supervisor	Xiaolong Liu
Assessor	Ed Nozeman
E-mail	a.skovolas@student.rug.nl
Date	21st May 2024

Abstract

This study investigates the effects of school proximity and school quality on residential properties values in Allegheny County, Pennsylvania. The analysis is accomplished considering both public and private schools as well as level of education. Employing a hedonic pricing model and collecting data for residential transactions and schools for 2022, I found that while both quality of school and proximity can be capitalized into property values, the degree of capitalization depends on the level of education, as well as kind of school. Results suggest that closer proximity to schools at all levels leads to higher housing prices. My analysis revealed that proximity to public primary and private high schools presented the strongest impact on housing prices when properties refrain 700m from schools, by 10.65% and 27.88% respectively. In terms of quality attributes, test scores posed the greatest effect on residential prices in the case of public and private high schools respectively. Finally, my results showed that properties tend to increase the most when locating within 700m from high quality schools by 3.47%.

Keywords: school quality, distance to schools, housing prices, hedonic price model, school types

Table of Contents

Preface

Dear Readers,

This master thesis is titled " What is the *Impact of School Proximity and Quality on Housing Prices in Allegheny County: Comparing Public and Private Education*, represents the end of my academic journey and research in real estate studies.

Proximity to schools, school quality and their relationship with housing prices have emerged as a significant topic today. This thesis aims to investigate the complex relationship between proximity to schools, education quality and the effects on housing prices. Throughout this thesis, I delved into extensive literature, analyzed large datasets, and applied statistical techniques to address this connection.

I am grateful to my supervisor, Xiaolong Liu, for his guidance, expertise, and support throughout this research. His mentorship has played a vital role in shaping the direction and quality of this thesis. Furthermore, I would like to acknowledge the support and encouragement of my family and friends, who have stood by me during the ups and downs of this academic pursuit.

Lastly, I appreciate the academic community and fellow researchers who have paved the way in examining this vital relationship between proximity to schools, school quality and housing prices for several areas. Their pioneering work laid the foundation for this thesis. I believe my thesis can be beneficial for the academic community in examining further this complex relationship.

I hope this thesis provides valuable insights to policy makers and urban planners when examining this relationship. My goal is that the findings presented on this research can contribute to the ongoing discussion and enhance understanding of real estate dynamics.

Athanasios Skovolas

Master’s Student in Real Estate Studies

University of Groningen

1. INTRODUCTION

In the introductory chapter, we discuss the motivation problem of the present thesis. Furthermore, we state the objectives and the research questions are indicated. Finally, the remaining structure of the thesis is explained.

1.1. Motivation

According to Glindro et al. (2011) housing can be perceived as a special type of real estate asset with a dual meaning, as consumption and investment good. The factors that influence the price of property are therefore an issue that has a lot of significance and can be summarized as follows: i) housing characteristics (number of rooms, floor, year built of the property, condition of the property and construction type), ii) neighborhood characteristics (crime rates, safe environment, kind and availability of housing types),and finally iii) other amenities (accessibility to nearby places, public transportation). (Gibbons & Machin, 2003; Bergantino et al., 2022)

The last few years economists and planners have shown a great deal of interest in the use of rent or value data for residential properties to estimate the advantages to households brought about by changes in accessibility to important recreational areas, companies, retail outlets or educational institutions (Tiebout, 1956). The advantages of being close and accessible to these activities can determine property values. Indeed, it is well known that the choice of where to live and the demand for housing are significantly influenced to a great extent by schools. (Tiebout, 1956). Most parents prefer to send their young children to nearby primary schools because they recognize the advantages of living near a school.

As a result, value and accessibility have a complicated connection. Being moderately close to an activity may have a positive impact on value, while being too close to an activity may cause value to decrease (Sah et al., 2015; Huang & Hess, 2018). The two most fundamental benefits of living close to a school are accompanied with safety, reduced travel time and cost to go to that particular school. This is extremely vital for parents who have children going to primary schools (Chin & Foong, 2006; Weiner & Wolkoff, 2001). Another advantage is that parents can monitor and affect their children's activities usually offered by school, especially when proximity to schools is relatively short (Chin & Foong, 2006). However, being close to schools might have a negative impact on property values due to increased crime, traffic, and noise levels (Huang, & Dall' erba, 2020). Subsequently, families will be eager to spend more money for the aforementioned benefits and would prioritize accessibility to schools as a vital reason, when choosing the location of their dwelling (Jud & Watts, 1981).

Moreover, together with proximity, an equally important attribute is to examine how quality of education can be capitalized into housing prices. Several empirical papers have tried to examine the fact that housing prices tend to be higher in areas where schools perform better. This is because schools' reputation will impact individuals' thought of where their house will be located and subsequently its value. This is turn will have a positive effect on demand for housing in that particular area, as parents can take advantage of the benefits when they know their house is surrounded by good quality schools. That fact also justifies parents' willingness to pay more for advanced school performance.(Chin & Foong, 2006). A recent study by Clark & Herrin (2000) conducted by the California Public Education Partnership in the U.S. revealed that residents prioritize quality of schools more compared to issues related to crime reduction or environment. Studies have acknowledged that a neighborhood's quality is important in defining a residential property's demand and price as well as in assessing how education quality affect housing prices. So, when examining educational quality and housing prices, it is essential to understand how this relationship can be capitalized into housing prices, and that depends on the level of education. (Fack & Grenet, 2010; Feng & Lu, 2013). Fack & Grenet, (2010), revealed that parents are more sensitive to student peer quality than to the quantity of schools, existed in an area, and when something like this takes place, then education inequalities can be prevented based on the school rules.

The motivation to study this topic is the following. Through this thesis we are going to investigate how housing prices react to the quality of education as well as proximity to schools. The area of interest will be Allegheny County, in Pennsylvania which presents a variation regarding the number and kind of schools. Furthermore, my analysis will consider both public and private schools located in the area, but also the three different levels of education, that is primary, middle, and high schools. This will be beneficial in order to shed some light on how these two factors affect and to what extent residential prices in the surrounded area, which in turn will determine demand for housing.

This analysis can bridge the gap and contribute in the literature by examining the effects of both proximity to schools and education quality on housing prices by considering both public and private schools simultaneously. Additionally, this analysis will also consider the level of education, that is primary, middle, and high schools in order to have a completed and comprehensive knowledge when it comes to estimate the degree of capitalization into housing prices. This will be the first analysis of its kind for Allegheny County, Pennsylvania, incorporating all these factors concurrently.

1.2. Research Problem Statement

The main objective of this master thesis is to explore the relationship between proximity to schools, education quality and the effects on housing prices, in the area of Allegheny County, Pennsylvania. This area presents a great deal of interest due to its variability of schools, both public and private, as well as availability of housing.

Thus, the aim of this thesis is to contribute by expanding related research on another region in the US, and specifically in Allegheny County, Pennsylvania. By analyzing the effects of proximity to schools and education quality on housing prices when distinguishing between public and private schools and at the same time per level of education, we can close the gap in the literature, since all of these attributes are taken into account. With that respect, it would be interesting to perceive how housing prices react, as well as which type of school and level of education can generate higher housing prices. This analysis will be achieved for the year 2022.

Before proceeding with the research questions that we would like to investigate it is essential to note why the effects might differ between public and private schools. Firstly, private schools are usually located in wealthier areas, and students come from higher socioeconomic backgrounds, and generally achieve higher academic progress. Moreover, private schools are associated with the provision of better education quality compared to public schools, as well as they have a smaller class size, and a range of other amenities which public schools are not equipped with. (Fack & Grenet, 2010; Jayantha & Lam, 2015; Feng & Lu, 2013). Furthermore, many families decide to enroll their children in private schools when they feel that the educational benefits justify the costs, such as tuition fees, extra transportation, and related expenses (Clark & Herrin, 2000). Lastly, public schools are more evenly distributed across the area, where proximity matters and can affect housing prices (Bergantino et al, 2022). Based on the discussion above, the distinction between public and private schools allows us to observe how the degree of capitalization of proximity and education quality can be capitalized into housing prices.

The following research questions have been formulated in order to examine this relationship, namely:

Main Research Question: What is the Impact of School Proximity and Quality on Housing Prices in Allegheny County: Comparing Public and Private Education.

To answer this central research question, the topic is broken down into four sub questions, which will be answered individually within this master thesis.

Research Question 1: What are the key factors influencing residential location choices according to academic literature?

Which can be the crucial determinants that can affect housing prices?

That question seeks to examine and establish the factors associated with residential location decisions. A detailed review of the existing literature on these determinants is required to answer the first question. It is crucial to comprehend the theoretical background of which these factors are, before proceeding with the next research questions for our analysis.

Research Question 2: What are the effects of proximity to schools on housing prices, in Allegheny County?

Is there any proof to support that properties located at a close proximity to schools carry a price premium, compared to the ones located further away?

The following question will be answered through statistical analysis of the empirical data using a hedonic pricing model. The method is further specified in the methodology section below.

Research Question 3: How does the overall education system's quality in Allegheny County affects housing prices?

Is there any evidence to support how housing prices react to different quality metrics when focusing on both types of school and level of education?

The following question presented above will also be answered through statistical analysis of the empirical data using a hedonic pricing model. The method is further specified in the methodology section below.

Research Question 4: How does the immediate proximity to top-performing schools influence the market value of residential properties within Allegheny County?

The following question presented above will also be answered through statistical analysis of the empirical data using a hedonic pricing model, with the help of an interaction variable. The method is further specified in the methodology section below.

As this master thesis seeks to examine the relationship between proximity to schools, education quality and the effects on housing prices, all of the questions described above will be answered considering both public and private schools, as well as all three levels of education for the area of interest, which is Allegheny County, Pennsylvania.

1.3. Conceptual Model

Based on the theory regarding the relationship between education quality, accessibility to schools and the effects on housing prices, a conceptual model is created as illustrated in Figure 1 below. The main objective is to find a quantitative connection between these two factors and housing prices, when separating between types of schools (public & private) and levels of education (primary, middle and high schools)

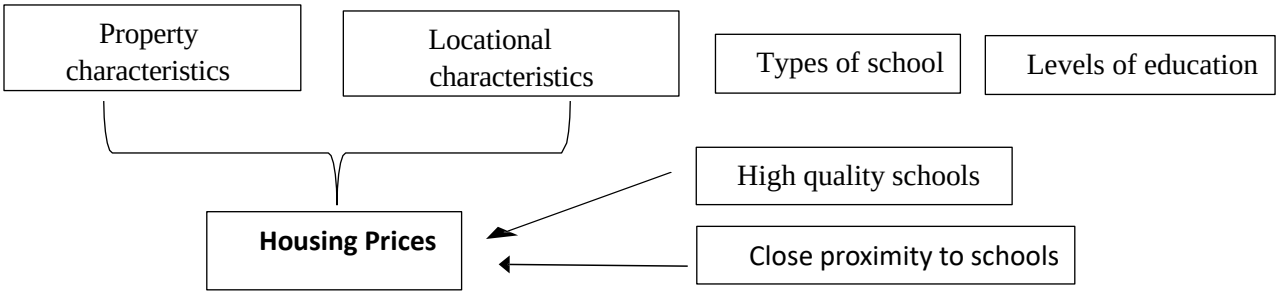


Figure 1: Conceptual model explaining the effect of education quality and accessibility on housing prices.

1.4. Outline

The remainder of this thesis is structured as follows. The Chapter 2 consists of the theoretical framework, together with the literature review and the hypotheses. The 3rd Chapter will examine the data used as well as the methods of the regression analysis. The 4th Chapter will include regression results and findings interpretation of the study. Lastly, Chapter 5 discusses limitations of the study, main conclusion and some suggestions for future research as well as policy implications and recommendations.

2. THEORY, LITERATURE REVIEW, HYPOTHESES

2.1. Theoretical Background

Previous studies have addressed the effects of close proximity to schools and education quality on residential prices. In general, most of the studies suggest that close proximity to schools and higher education quality leads to higher housing prices in that area. This applies also when separating between level of education that is primary, middle and high schools, but also per type of school, public or private. Subsequently, in that regard, it has been proven that families prioritize to a great extent such attributes when choosing the location of their dwelling. Therefore, the effects of proximity to schools as well as education quality on housing prices has been studied systematically, and particularly whether a price premium or a penalty is associated with that and how the real estate market is responding to such changes.

In general, there is a positive link associated with closer proximity to schools at all levels, and a negative link associated with a significantly longer distance to schools. Most of the papers from the literature suggests that properties being close to schools are associated with a price premium. (Owusu-Edusei et al., 2007; Huang & Dall'erba, 2020; Metz, 2015; Bergantino et al., 2022). However, there are cases where that relationship suggests a price penalty, rather than a price premium for residential properties. (Sah et al., 2015). This demonstrates that housing prices drop when locating close to schools, impacting the whole real estate housing area negatively. On the whole, findings suggest that housing prices tend to increase by a certain % when are located close to schools, and as the distance increases then that % drops at an increasing rate.

Moreover, there is also a positive relationship between education quality and housing prices. It is obvious that better education quality of schools is associated with higher housing prices in that area. This outcome can be achieved when looking at different quality metrics for both public and private schools, at all levels of education such as test scores, student progress, report cards, in a way this relationship to be sustained. Most of the papers from the literature suggest that properties that are surrounded by good quality schools are actually sold higher compared to properties located next to low quality school, especially when looking at metrics such test scores and reports cards. (Liu et al., 2022; Lu et al., 2023; Jayantha & Lam, 2015; Fack & Grenet, 2010). However, when looking at different quality metrics, such as number of students, the literature reveals an inverse relationship between education quality and housing prices (Zhou, 2018, Wen et al., 2017). When trying to determine the effects of distance and education quality on housing prices it is essential to

focus on a specific area of interest and at the same time to distinguish between type of schools and level of education, in a way to better capture the overall impact on property values.

However, there are also a number of other additional factors that can determine the value of a property, beyond its proximity to schools and the education quality. Such factors can relate to market conditions, neighborhood amenities and locational factors. For instance, housing characteristics, as well as proximity of properties to railway stations, CBDs, highways, workplaces and shopping malls have also to be taken into account, as they impact property prices (Abidoye & Chan., 2016; Aliyev et al., 2018).

Examining several articles, the effect of proximity to schools and education quality on housing prices has been frequently investigated. However, this has not been thoroughly studied when looking at both public and private schools at all levels of education, that is primary, middle and high schools. Much of the research has focused on specific counties in the United States or on Asian countries like Singapore and China, which differ significantly in school types, availability, and data accessibility. It is essential to discover how state and non-state schools at all levels influence and to what extent housing prices in order to contribute to the academic literature with new knowledge.

2.2. Theoretical framework of property value determinants

As has already been stated above, the value of a property is influenced by a set of elements which have been separated into attributes. (Abidoye & Chan, 2016; Aliyev et al., 2019; Charles, 2012; Kim et al., 2015). Such attributes according to Chin & Chau, (2003) and Wen et al. (2005) refer to locational, structural, and neighbourhood variables, all of which have a different impact on property values. By taking the first derivative of prices in regard to one of these attributes, it can be estimated that if the derivative is positive, there is a positive effect on housing prices and vice versa (Wu et al. (2013); and Bayoh et al. (2006). At the previous section of this analysis, it has been stated what kind of aspects each of these attributes include, so there is no need to reiterate those aspects again. With that in mind, a framework can be constructed as appears below, describing the relationship between those attributes and the respective effects on housing prices. This part of the analysis answers the first question of this master thesis, that is which are the factors affecting residential location choices according to the literature.

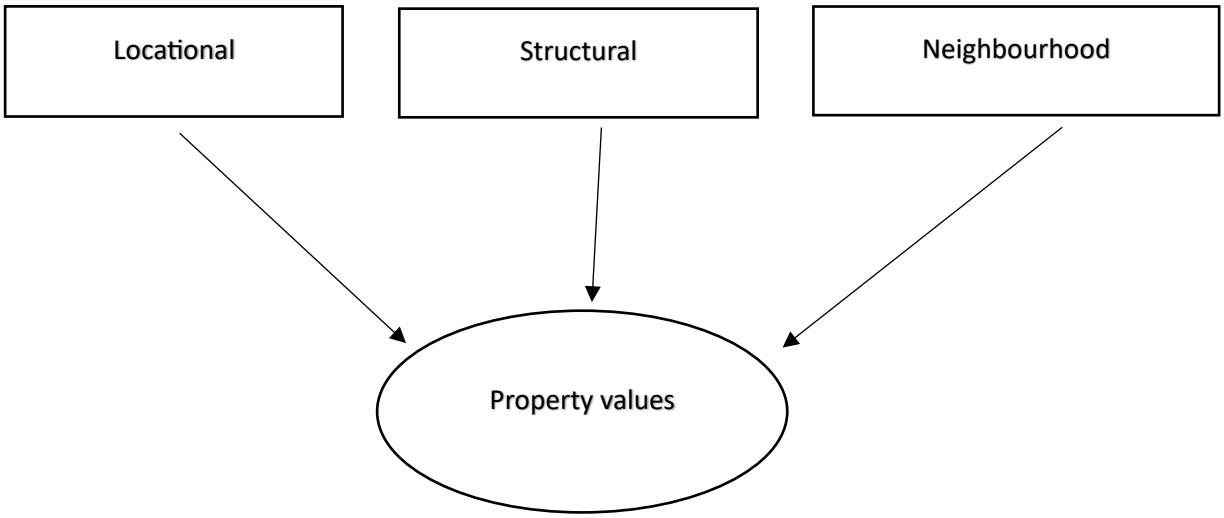


Figure 1: Framework property values determinants scheme.

2.3. Literature Review

School Proximity and Housing Prices

Proximity to schools and the effects on housing prices is a topic which has been widely discussed and examined by researchers, as it influences to a great extent property values and housing valuations. This section deals with the theoretical background. At this part, emphasis is given on the literature in a way to have a more completed overview of this relationship, as well as the methodology used and the main findings. Here the studies can be distinguished as follows i) studies that estimate the impact of school proximity on housing market prices through a pure hedonic model; ii) other studies employ different techniques such as the spatial approach, or difference in difference approach (DID) rather than a hedonic price model. As for the relationship between proximity to schools and housing prices, a pure hedonic price model is mainly used to determine the respective effects. In all of the articles (Huang & Hess, 2018; Sah et al. 2015; Metz, 2015; Bergantino et al. 2022) the authors use the sales price of the property as dependent variable, expressed in a logarithmic form and proximity to schools, as the main independent variable alongside, other variables such as housing and neighborhood characteristics, to determine the effects on housing prices.

Des Rosiers et al. (2001) in their analysis, employed a hedonic model to estimate the effect of both the size and distance of primary schools on residential property values in Quebec, Canada. The findings suggest that primary schools which are close to properties within a range of 300m-500m affect property values the most. Another study by Chin & Foong (2006) which coincides with the previous study, revealed that accessibility to prestigious primary schools is more crucial than access to junior high schools. Metz (2015) also observed that proximity to schools at all levels has a positive impact on housing prices. Another study conducted by Sah et al. (2015), investigated the public and private school proximity effect on residential prices. When distinguishing between public and private schools they concluded that home prices rise by about 0.755 % for each 1,000 ft of distance from the closest public elementary school, as opposed to private schools where property prices rise by 1.4%.

As explained before, other studies investigate/examine the relationship between proximity to schools and housing prices using different methods, like quantile regressions and DID models. For instance, Huang & Hess (2018) investigated the relationship between distance of residential properties to schools and the effects on housing prices in Oshkosh, Wisconsin, U.S considering all levels of education. Adapting a quantile regression, results revealed that distance to all levels of

education has a positive impact on housing prices, but also distance to primary and middle schools is more crucial than that of high schools. A similar study by Huang & Dall' erba (2020) found that the relationship between housing prices and distance to middle schools appears to be nonlinear, meaning that property prices increase when they are close to schools, but decrease above 3.6 km. Bergantino et al (2022), investigated the relationship between accessibility to kindergartens and housing prices in Italy, in the eleven major Italian Municipalities' neighborhoods, and they concluded that proximity to a kindergarten is capitalized into property values and prove that close proximity to a kindergarten has a significant and positive effect on housing prices.

School Quality and Housing Prices

According to Sah et al. (2015) school quality has a positive relationship with housing prices. For instance, Clark and Herrin (2000) investigated how housing prices react to quality of schools in California and they found that at specific districts when quality of schools is good, prices tend to be sold at a premium compared to other districts. A number of studies has been conducted in a way to determine the relationship between housing prices and education quality, using different mechanisms of quality (students' expenses, student to teacher ratio, average class size, report cards, and test scores). Initial studies were focusing on students' expenditure as an effective way to determine the impact of education quality on housing prices (Oats, 1969; Edel & Sclar, 1974; Gustely, 1976; Sonstelie & Portney, 1980). These studies revealed that the higher the amount of expenditure per student, the better the education quality can get.

According to Wen et al. (2018) test scores and schools ranking are an effective and proper way to determine the effects of education quality on housing prices. Evidence in general reveals that high education quality is accompanied with higher housing prices . Although this is true and can be applied to many countries (Wen et al. 2014; Figlio & Lucas, 2004; Kane, 2006; Lu et al. 2023) there is not much evidence regarding what is the effect of education quality when making distinctions by type of schools or school levels on housing prices (primary, middle, high school). Some studies examine the relationship between school quality and housing prices for a specific level of education, some others differentiate between public and private schools, while others make comparisons between multiple levels of education. Wen et al. (2014), observed a positive link between school quality and housing prices. When separating between schools they found that primary and junior high schools have a substantial impact on the surrounding school district. The total value of the property in this school district rises by 2.020% or 5.443%, respectively, for every degree that elementary and junior high school education quality improves.

Figlio and Lucas (2004) observed that there is a strong relationship between school grades and housing prices and residential location. This is because good grades lead to better education quality. For instance, a grade "A" can lead up to 19.5% increase in housing sales prices compared to a grade "B". Another study by Seo and Simmons (2009) revealed that school district ratings and overall performance of a school, which appear to be the most appropriate measures for education quality, have a strong impact on residential prices and so have test scores. Brasington and Haurin (2006), found that expenditure per student contribute positively on housing prices as well. This can lead up to a 7% increase in the housing prices. These results coincide with the findings of Black (1999) regarding test scores, where the increase on housing prices can be up to 2.3%, as well as Kane's

(2006) with an increase of 0.5%, Clapp et al, (2008) with an increase on housing prices of 1.4% and Gibbons & Machin (2003) with an increase of 0.67% and 11% in the study by Chiodo et al. (2005), and Lu et al. (2022) for the area of Sydney, Australia. Feng & Lu (2013) investigated the relationship between education quality and housing prices after the reformation of China's education system, with the addition of Experimental Model Senior High Schols. By acquiring panel data for over 50 residential areas in Shanghai, China, and using a hedonic model specification, the authors proved that the presence of an additional Experimental School caused an increase in the housing prices by almost 17%. Liu et al. (2022), investigated the relationship between education quality and housing prices in the case of primary schools in Shanghai, China. The authors proved that high quality education offered by primary schools can lead up to 16% increase in housing prices.

Hypotheses

Based on the literature review examined above, two hypotheses can be stated and specified to answer the composing research questions:

H1: Proximity to schools at all levels, whether public or private, tends to boost housing prices.

H2: Excellent quality of education at all levels, in both types of schools, is positively linked to higher housing prices.

3. DATA & METHODOLOGY

This chapter represents the dataset. Firstly, I describe from where the datasets were obtained by explaining and analyzing the necessary variables. Then a hedonic pricing model is selected to investigate and clarify the relationship between proximity to schools, education quality and the effects on housing prices. This choice was based on the nature of the dependent variable, allowing me to examine the link between different independent variables and the effects on prices.

3.1. Context

Pennsylvania is one of the fifty states of the United States, located to the northeast side. It is the fifth-most populous state in the United States, with over 13 million residents and 33rd largest by area, just after Mississippi. One of the biggest metropolitan cities is Philadelphia, Pittsburgh and Delaware Valley. Some neighboring states are Delaware to its southeast, Maryland to its south, West Virginia to its southwest, Ohio to its west, Lake Erie, and New York to its north.

The State counts 67 Counties, namely, Philadelphia, Allegheny, Montgomery, Bucks, Delaware, Lancaster and Chester. From them Allegheny County, which is used also as a case study for this master thesis has 1.2mil inhabitants and it is considered the most populated County in Pennsylvania together with Philadelphia with 1.5mil inhabitants. Moreover, Allegheny County accounts for 130 municipalities, with some being Bethel Park, Penn Hills, Monroeville and Pittsburgh. Figures 3 and 4 depict a map showcasing all Pennsylvania's Counties and all of the municipalities in Allegheny County respectively.

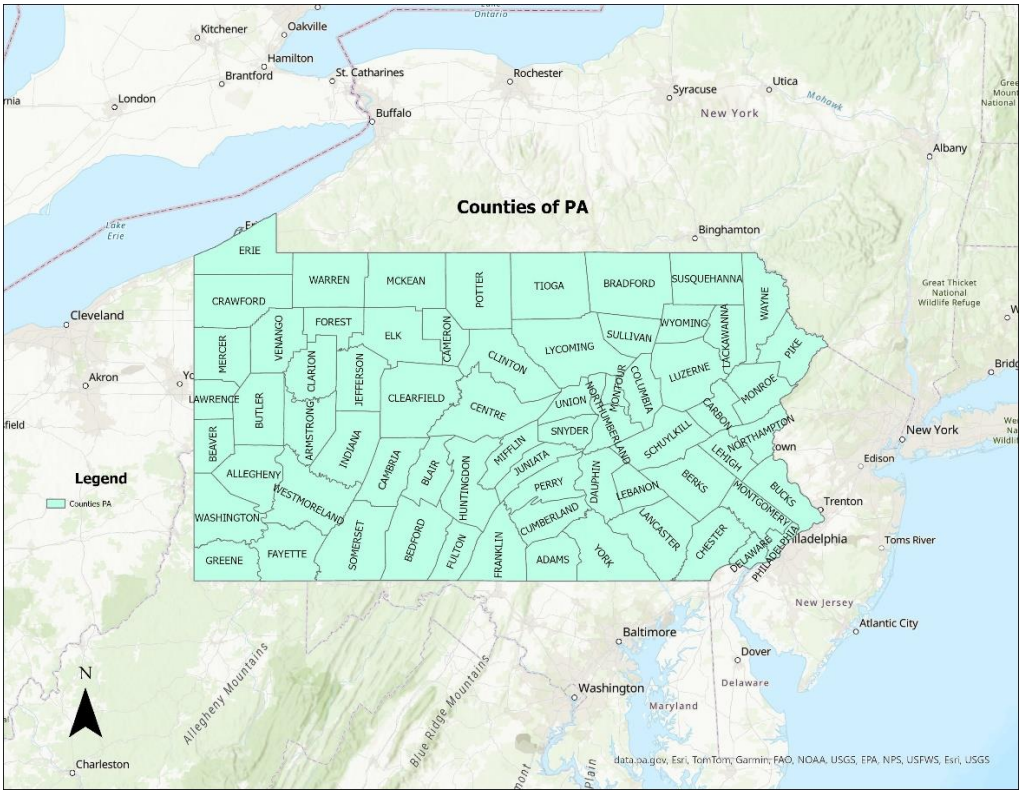


Figure 3: Map showcasing all Counties of Pennsylvania

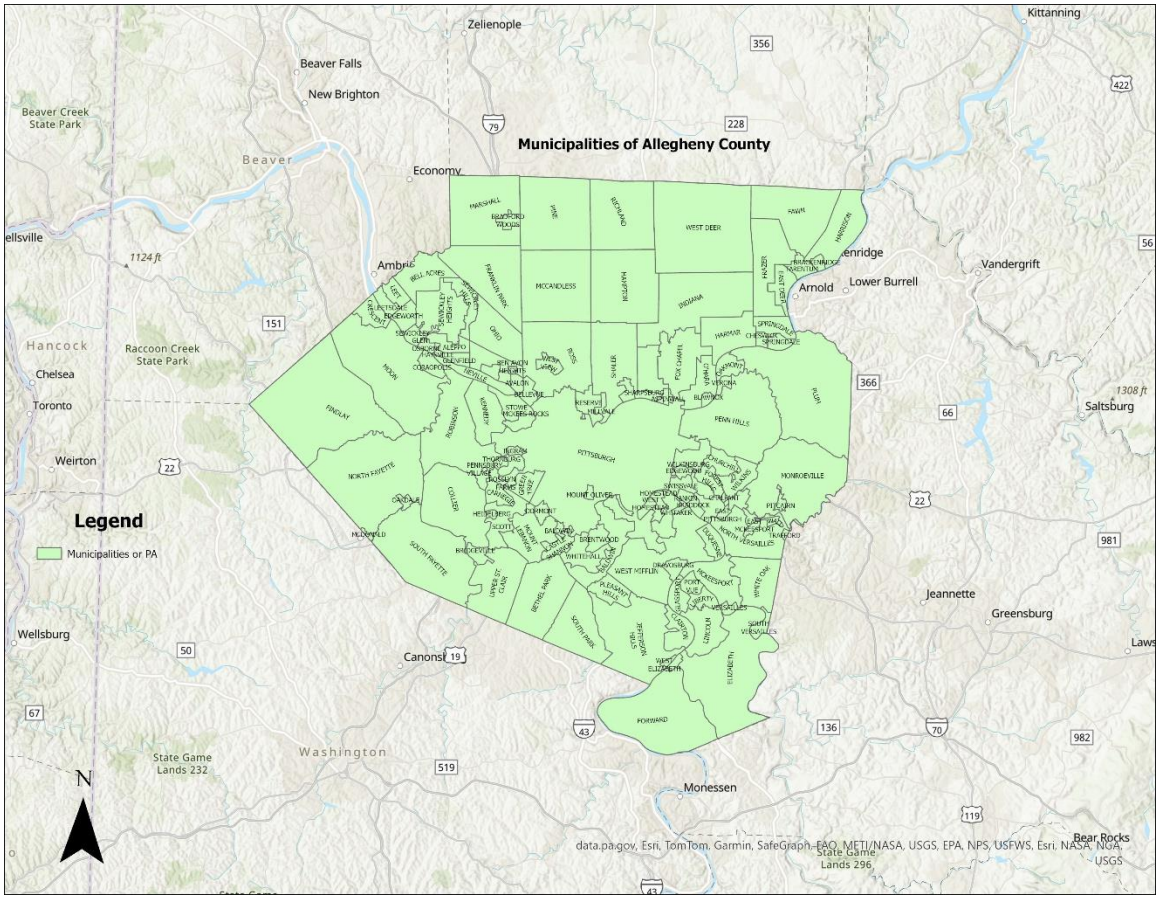


Figure 4: Map showcasing Allegheny's Municipalities

3.2. Dataset Background

Housing Transactions

This research investigates the relationship between proximity and quality of schools and what are the effects on residential home prices specifically for Allegheny County, Pennsylvania. As a result, this quantitative research considers various sources from which the necessary datasets were gathered in order to analyze and examine that aforementioned relationship, when considering both public and private schools at all levels of education. The housing transactions data were collected from this source (WPRDC, n.d.). This dataset contains all the necessary information regarding residential real estate transactions that have taken place in Allegheny County over a period of years. More specifically, this dataset contains housing transactions data starting from 2010 till 2022. The analysis for this research will be implemented for 2022.

Subsequently, the dataset includes the following property characteristics: property city, address & house number, main use of the property. Furthermore, it includes information regarding sales date and sales price(in $) of the property, zip codes of properties, year built, type of construction which refers as exterior finish (brick, frame, concrete, stone, etc.), condition of the property (excellent, average, fair, good). Finally, there is information about the total number of rooms, bedrooms, full and half baths, lot area (in sq. ft), living area (in sq. ft) number of stories, age of the property, number of fireplaces, basements and garages.

Moreover, the dataset also includes essential records regarding ZIP codes. That simply demonstrates that there is geographical information regarding the latitude and longitude of the properties. This appears to be extremely beneficial for two main reasons. Firstly, the location of these properties can easily be depicted on a map, using the software ArcGIS Pro, which we did on this research. Secondly, with such geospatial data points, the distance between residential properties and schools can easily be estimated in a rapid and effective way.

Education Quality

The second dataset was collected from National Center for Education Statistics (NCES, n.d.) and (Private Schools, n.d.). It contains useful information regarding public and private schools for the whole US, but also the whole dataset is divided into counties as well. For this research, as the first dataset that we have regarding housing sales transactions refers to Allegheny County, Pennsylvania, the dataset for schools has to consider the same area.

The original dataset for Allegheny County consisted of 300 public schools and 170 private schools. After going through the dataset, some schools were permanently closed, so they were not taken into account for the analysis. Consequently, the final dataset consists of 266 public and 101 private schools. In order to capture the effects of education quality on housing prices, this dataset also refers to the school year of 2022.

This dataset contains all the necessary information from primary to high school level. More specifically, it includes the school's address and name, but also the city in which the school is located. Apart from that there is also information about the number of student enrollments. Finally, in this dataset information regarding ZIP codes is contained, together with the coordinates of where the school is located. Once again, as has been stated above this facilitates the calculation of the distance between residential properties and schools.

To measure the quality of schools two sources are mainly used to collect data regarding the performance of schools, either public or private. (GreatSchools, n.d.) and (Private School Review, n.d.). GreatSchools is an independent non-profit organization which aims to provide parents with information such as overall rating of the school, test scores, school demographics, teachers and staff, learning outcomes, but also student with disabilities focusing on discipline and attendance. By adding the necessary quality metrics on the existing datasets, for which information could be taken, an expanded version of them has been created. These datasets will be considered when it comes to estimating the effects of quality of education on housing prices.

Based on the information that we obtained from these two platforms the existing datasets for public and private schools, were further expanded, based on the data availability. A detailed overview of which variables were used for this analysis can be found in the Descriptive Statistics section of this research paper.

Cleaning Process

To determine the effects of proximity to schools on housing prices, we had to clean the data before processing them. Initially, values with 0 or no value were removed from the dataset. Then the dataset was further narrowed down to properties built between 1900-2021 in order to remove outliers and to narrow down the difference between min and max. Next, emphasis was given to properties featuring 1-10 total rooms, 1-8 bedrooms, 1-5 full baths, and 0-4 half baths. Finally, by removing the lower and the upper 2.5% percentile from the sales price, we decreased the sample even more. After all these considerations, sales price of the properties for this research range from $80.000 to $717.000. As this research investigates the effects of distance on housing prices, and since this variable will act as our main independent variable, we had to remove outliers and narrow down the difference between min and max by also removing the upper 2.5% of the observations. This was implemented for all levels of education in the case of both public and private schools. Consequently, this led to a further reduction regarding the number of observations, leading to a final dataset of 8.827 observations.

On the contrary, to determine the effects of education quality on housing prices the same process was followed as before. As a result, this led to a final dataset consisting of 10.275 observations, which applies both to public and private schools, after narrowing down the sample.

3.3. Descriptive Statistics

The tables below depict the descriptive statistics for all of the variables that will be included in our analysis in order to examine and investigate the relationship between proximity to schools and education quality and respective effects on housing prices. As has already been stated, this research is focusing on both public and private schools at all levels of education. With that respect, separate tables have been created for all of the variables.

More specifically, the descriptive statistics show the number of observations, the mean, the standard deviation, the minimum, and the maximum for each variable that they will be considered in the analysis. Table 1 illustrates the variables for all housing characteristics.

Table 2 provides information about the key independent variable which is distance to the nearest school, when separating between type of school and level of education. Particularly, Table 2 considers distance to the nearest school as a continuous variable,

The final Tables 4 and 5 provide information regarding quality metrics for both public and private schools at all levels of education. This is essential in order to realize how housing prices react to those quality metrics. As it can be seen from the tables 8 quality metrics will be considered in case of public schools and 6 in private schools. More information about what each variable mean and what it represents can be found in the notation table, which can be found in the Appendix part of this research.

Table 1: Descriptive Statistics for Housing Characteristics

Variable	Obs	Mean	Std. Dev.	Min	Max
Saleprice	8.827	247051.71	125454.26	80.000	717.000
Totalrooms	8.827	6.304	1.352	1	10
Bedrooms	8.827	2.96	.772	1	8
Fullbaths	8.827	1.409	.563	1	5
Halfbaths	8.827	.493	.558	0	4
Fireplaces	8.827	.475	.662	0	5
Exteriorfinish	8.827	2.329	1.538	1	7
Stories	8.827	1.575	.496	1	3.5
Garage	8.827	.823	.801	0	4
Basement	8.827	4.705	.991	1	5
Age of property	8.827	70.519	25.314	1	122
Sales Month	8.827	5.634	2.691	1	11
Yearblt	8.827	1951.481	25.314	1900	2021

Table 2: Descriptive statistics for the main independent variable, as continuous

Variable	Obs	Mean	Std. Dev.	Min	Max
Distance to nearest pub prim	8.827	1.148	.648	.022	3.344
Distance to nearest pub mid	8.827	1.893	1.068	.048	5.215
Distance to nearest pub high	8.827	1.931	.963	.064	4.598
Distance to nearest priv prim	8.827	2.004	1.266	.053	6.811
Distance to nearest priv mid	8.827	3.567	2.465	.032	11.579
Distance to nearest priv high	8.827	3.633	1.751	.031	7.965

Table 4: Descriptive statistics for the effects of public education quality on housing prices

Variable	Obs	Mean	Std. Dev.	Min	Max
Public Primary					
No. of students pub prim	10.275	416.105	191.349	120	1028
Test score pub prim	10.275	6.91	2.237	2	10
Student progress pub prim	10.275	6.079	1.35	2	10
Equity Score pub prim	10.275	5.685	1.474	2	10
Suspension rate pub prim	10.275	.046	.063	0	.45
Student to teacher pub prim	10.275	13.248	2.292	9	19
Teacher salary pub prim	10.275	69792.471	13405.291	33054	112500
Pupil to counselor pub prim	10.275	211.19	116.538	70	280
Public Middle					
No. of students pub mid	10.275	473.409	236.204	150	700
Test score pub mid	10.275	7.333	2.113	3	10
Student progress pub mid	10.275	6.149	1.366	3	9
Equity Score pub mid	10.275	5.588	1.206	3	9
Suspension rate pub mid	10.275	.079	.079	.01	.31
Student to teacher pub mid	10.275	12.413	2.199	9	17
Teacher salary pub mid	10.275	72776.196	13457.26	71000	112600
Pupil to counselor pub mid	10.275	315.497	106.686	107	540
Public High					
No. of students pub high	10.275	1003.169	575.558	400	2024
Test score pub high	10.275	6.544	1.846	3	10
Student progress pub high	10.275	7.038	2.215	4	10
Equity Score pub high	10.275	5.7	1.633	3	9
Suspension rate pub high	10.275	.076	.082	.01	.46
Student to teacher pub high	10.275	13.169	2.118	8	18
Teacher salary pub high	10.275	74745.279	13464.549	34700	113800
Pupil to counselor pub high	10.275	283.21	71.778	120	640

Table 5: Descriptive statistics for the effects of private education quality on housing prices

Variable	Obs	Mean	Std. Dev.	Min	Max
Private Primary					
No. of students priv prim	10.275	127.835	75.336	20	190
Summary Rating priv prim	10.275	7.489	.905	6	9
Test scores priv prim	10.275	7.414	.97	6	9
Student progress priv prim	10.275	7.339	.783	6	9
Equity Score priv prim	10.275	7.271	.859	6	9
Student to teacher priv prim	10.275	13.203	3.292	3	21
Private Middle					
No. of students priv mid	10.275	234.252	171.42	35	315
Summary Rating priv mid	10.275	7.425	.749	6	9
Test scores priv mid	10.275	6.945	.872	6	9
Student progress priv mid	10.275	7.307	.597	6	9
Equity Score priv mid	10.275	7.485	.718	6	9
Student to teacher priv mid	10.275	12.124	4.387	5	22
Private High					
No. of students priv high	10.275	257.765	192.662	60	433
Summary Rating priv high	10.275	7.502	.893	6	9
Test scores priv high	10.275	7.521	1.055	6	9
Student progress priv high	10.275	7.238	.8	6	9
Equity Score priv high	10.275	7.261	.948	6	9
Student to teacher priv high	10.275	10.27	2.945	4	17

3.4. Methodology

This research aims to investigate the relationship between proximity to schools, education quality and housing prices, while considering both public and private schools at all levels of education, that is primary, middle and high school. Therefore, our model depicts the impact of the key independent variable, distance to the nearest public and private school at all levels of education, and education quality metrics on the dependent variable, the transaction price that the property was sold. To avoid biased and inaccurate estimates of my result, the natural logarithm of transaction prices was considered. Subsequently, this leads to a normal distribution and it also helps with the linearity between the dependent and independent variables that will be included in the model.

To determine the effects of proximity to the nearest school on property prices, a hedonic pricing model was employed. Hedonic models are often used to investigate the relationship between housing prices and external effects. In order to examine the research question 2 that was raised above, named "*What are the effects of proximity to schools on housing prices, in Allegheny County",* the following empirical model was implemented. To answer this question, distance to the nearest school, public or private, at all levels is going to be considered as a continuous variable, in order to see the effects on housing prices.

Before proceeding with the empirical analysis, it is essential to note why the effects between public and private might differ. Below three reasons why this happens will be mentioned. Initially, many families decide to enroll their children in private schools, because they get some additional benefits that public schools do not offer, that is pick up and drop off of their children, with school buses from their homes to schools and vice versa. So, in that sense distance is not so much of an issue for parents sending their children to private schools. (Clark & Herrin, 2000). Whereas, in public schools as parents must go their children to school distance for them matters at a greater extent. Moreover, families that prioritize private education may have different housing preferences or financial capabilities, which affects the location of their home and how much they are willing to pay for a property, independently of school proximity. On the contrary, families from different socioeconomic status, that are unable to support private education would be interested in locating close to public schools (Feng & Lu, 2013). Lastly, public schools are more evenly distributed across the area, where proximity matters and can affect housing prices (Bergantino et al, 2022). Finally, public schools operate within specific catchment areas, (school districts), so families that want to send their children to public schools must live close to such school zones, prioritizing proximity to schools. On the contrary, private schools do not have catchment areas giving families more flexibility of where to live, so they do not focus that much on proximity to schools (Chin & Foong,

2006). These are some reasons why the effects of distance might differ in the context of public and private schools.

Model for the effects of distance on housing prices as a continuous variable

$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ to\ Nearest\ Publ\ School_{i,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t}$ ***(1)***

$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ to\ Nearest\ Priv\ School_{i,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t}$ ***(2)***

In the above equations $\ln(P_{i,t})$ represent the natural logarithm of transaction price, α represents the constant, β is the coefficient for the independent variables to be estimated, and $\varepsilon_{i,t}$ is the error term, which captures the unpredicted factors not included in the model that can affect the dependent variable Y. The notation of i represents the transactions being made: i = 1, 2, 3, 4…., N, while the notation of t formulates the time referring to the months that the properties were sold: t = 1, 2, 3, 4…., N. In both of the models we have included location effect and time effects. As for the first ones, the variable zip code denotes where the property is located and is labelled as b3*zip code. As for the time effects, sales months denotes the month which the property was transacted and in the model is labelled as b4*sales month. Both of these variables will be considered as categorical variables in our analysis. As for the remaining variables, Y is a vector 11 different housing characteristics, referring to total rooms of the property, number of bedrooms, full and half baths, fireplaces, garage, basements, type of exterior finish, year built of the property. These variables will be represented as categorical variables in the analysis. Whereas age of the property and number of stories will be considered as continuous variables in the analysis. In the model it is labelled as b2*Y*Housing Characteristics. Finally, for the key independent variables X is a vector considering the three levels of education, that is primary, middle and high school. Model 1 captures the effects of distance to the nearest public school at all levels on housing prices, while Model 2 captures the effects of distance on housing prices in the case of private schools at all levels. This is labelled as ***b1*X*Distance to Nearest Publ School and b1*X*Distance to Nearest Priv School*** for both models respectively. As has already been stated above, here distance to the nearest public and private schools at all levels will be considered as a continuous variable.[1]

Now that we have established the correlation between the proximity to schools and housing prices, a further analysis is required to explore the ways in which quality of education affects property

[1] To investigate the nonlinearity between housing prices and proximity to schools, the regression specifications and corresponding results, which treat distance as a categorical variable, are provided in the Appendix section of this research.

values. This is because the reputation of a school, together with their educational outcomes and resources can have an impact on housing prices. In order to examine the research question 3 that was raised above, named "*How does the overall education system's quality in Allegheny County affects housing prices?",* the following empirical models were implemented. For this analysis, we constructed distinct models for public and private schools, as different quality metrics for each were taken into account. Moreover, we also split up the models per level of education to assess the individual effects of these metrics on property values. This analysis is crucial in order to comprehend the relationship between qualitative educational attributes/outcomes and housing prices, thus providing an overview of the determinants of property values. To directly address research question 3 regarding the impact of educational quality on housing prices in Allegheny County, we developed the following empirical models.

Model for the effects of education quality on housing prices

$Ln(P_{i,s,t}) = a + \mathbf{b1*Quality\ Metrics\ Public\ Primary_{i,s,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(9)***

$Ln(P_{i,s,t}) = a + \mathbf{b1*\ Quality\ Metrics\ Public\ Middle_{i,s,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(10)***

$Ln(P_{i,s,t}) = a + \mathbf{b1*\ Quality\ Metrics\ Public\ High_{i,s,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(11)***

$Ln(P_{i,s,t}) = a + \mathbf{b1*\ Quality\ Metrics\ Private\ Primary_{i,s,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(12)***

$Ln(P_{i,s,t}) = a + \mathbf{b1*\ Quality\ Metrics\ Private\ Middle_{i,s,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(13)***

$Ln(P_{i,s,t}) = a + \mathbf{b1*\ Quality\ Metrics\ Private\ High_{i,s,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(14)***

We have already explained the meaning of the variables above. As for the key independent variable Quality Metrics, this acts as a vector variable, considering 8 metrics in the case of public schools at all levels, and 6 metrics in the case of private schools. These variables are all the same across all levels. More specifically, for public schools, we consider variables such as number of students, test scores, student progress, equity score, teacher salary, student to teacher ratio and student per counselor as effective metrics for estimating the effects on housing prices. On the contrary, in the case of private schools, quality metrics like number of students, summary rating, test scores, student

progress, equity score and student to teacher ratio were taken into account. In these models, all of the quality metrics that will be used in the model will represent continuous variables.

Finally, to answer the last question of this research, namely "*How does the immediate proximity to top-performing schools influence the market value of residential properties within Allegheny County?"* the following empirical model will be implemented. At this point since we are interested in examining the effects of properties being close to schools on property values, an interaction variable will be used between high quality schools and properties located within 700m from such schools. To determine the effects of the interaction variable on housing prices, separate models will be created for the different levels of education and type of school.

Model for the effects of interaction variable on housing prices

$Ln(P_{i,s,t}) = a + b1*X*Distance\ Group_{i,t} + b2*High\ Quality\ Publ\ Prim_{i,s,t} + \mathbf{b3*X*Distance\ Group*High\ Quality\ Publ\ Prim_{i,s,t}} + b4*Y*Housing\ Characteristics_{i,t} + b5*Zip\ code_{i,t} + b6*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(15)***

$Ln(P_{i,s,t}) = a + b1*X*Distance\ Group1_{i,t} + b2*High\ Quality\ Publ\ Mid_{i,s,t}\ \mathbf{b3*X*Distance\ Group1*High\ Quality\ Publ\ Mid_{i,s,t}} + b4*Y*Housing\ Characteristics_{i,t} + b5*Zip\ code_{i,t} + b6*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(16)***

$Ln(P_{i,s,t}) = a + b1*X*Distance\ Group2_{i,t} + b2*High\ Quality\ Publ\ High_{i,s,t} + \mathbf{b3*X*Distance\ Group2*High\ Quality\ Publ\ High_{i,s,t}} + b4*Y*Housing\ Characteristics_{i,t} + b5*Zip\ code_{i,t} + b6*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(17)***

$Ln(P_{i,s,t}) = a + b1*X*Distance\ Group3_{i,t} + b2*High\ Quality\ Priv\ Prim_{i,s,t} + \mathbf{b3*X*Distance\ Group3*High\ Quality\ Priv\ Prim_{i,s,t}} + b4*Y*Housing\ Characteristics_{i,t} + b5*Zip\ code_{i,t} + b6*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(18)***

$Ln(P_{i,s,t}) = a + b1*X*Distance\ Group4_{i,t} + b2*High\ Quality\ Priv\ Mid_{i,s,t} \mathbf{+ b3*X*Distance\ Group4*High\ Quality\ Priv\ Mid_{i,s,t}} + b4*Y*Housing\ Characteristics_{i,t} + b5*Zip\ code_{i,t} + b6*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(19)***

$Ln(P_{i,s,t}) = a + b1*X*Distance\ Group5_{i,t} + b2*High\ Quality\ Priv\ High_{i,s,t} + \mathbf{b3*X*Distance\ Group5*High\ Quality\ Priv\ High_{i,s,t}} + b4*Y*Housing\ Characteristics_{i,t} + b5*Zip\ code_{i,t} + b6*Sales\ Month_{i,t} + \varepsilon_{i,s,t}$ ***(20)***

As we have already explained most of the variables' meaning above, there is a variable that we have not referred to. High quality is a dummy variable, focusing on all levels of education per type

of school, and takes value of 1 if tests scores of students is between 8-10 and 0 otherwise. We decided to consider test scores as a measurement of high quality, since this variable was included in both models in the case of public and private schools, when determining the effects of education quality metrics on housing prices. In the models it is labelled as b2*High Quality. Finally, as for the main variable of interest which is labelled as ***b3*Distance Group*High Quality,*** we want to see if properties located within 700m from high quality schools, command higher prices, compared to the ones located further away from such schools.

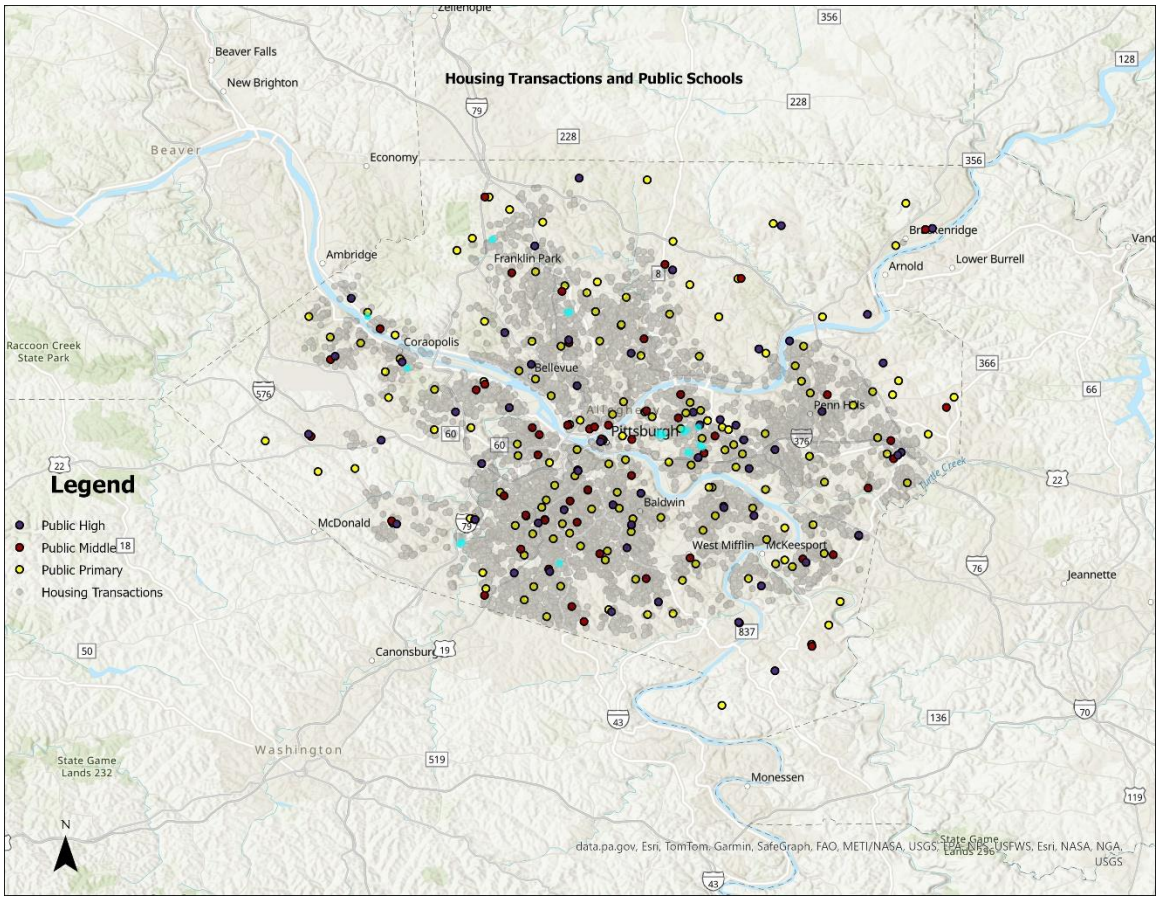


Figure 5: Housing Transactions and Public Schools per level of education

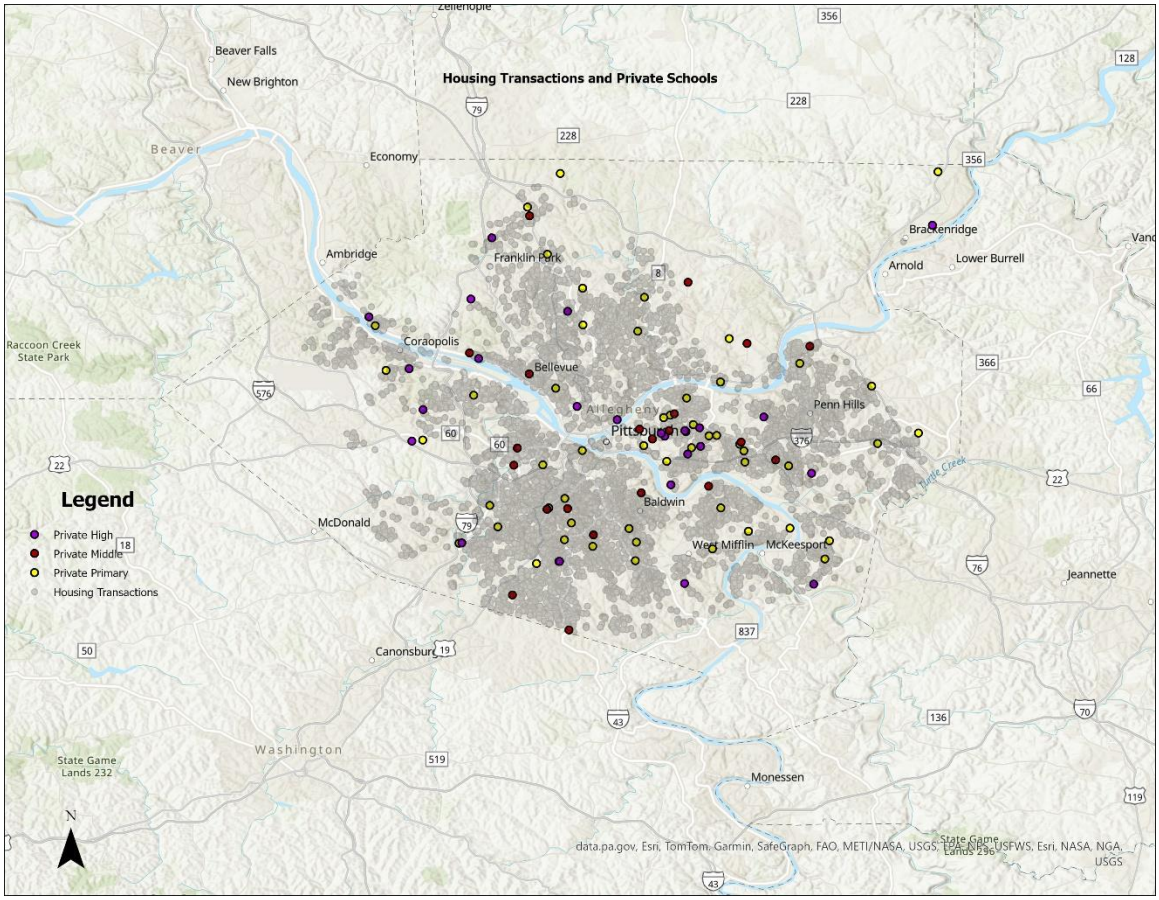


Figure 6: Housing Transactions and Private Schools per level of education

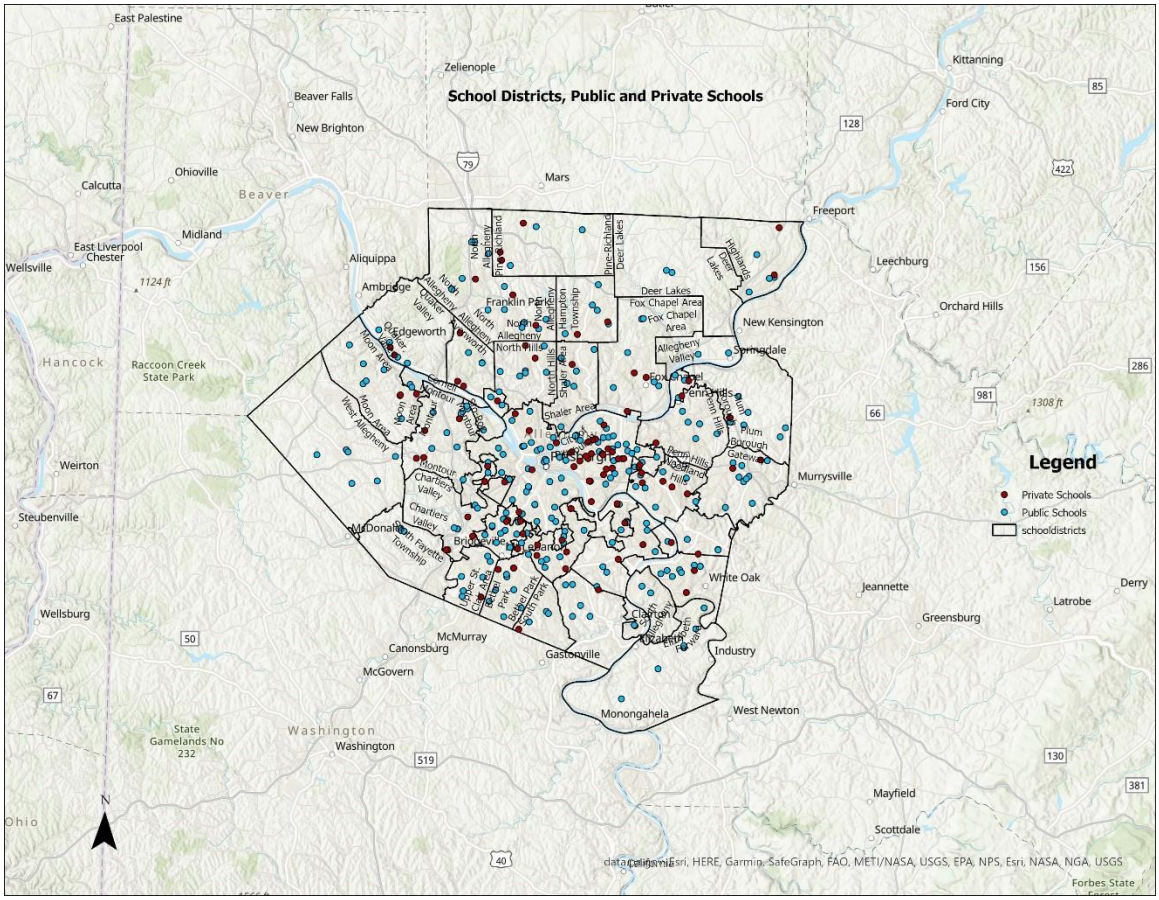


Figure 7: School Districts & All schools.

4. RESULTS & DISCUSSION

4.1. Results

In this chapter, the regression results of the hedonic pricing models described above are presented. This research aims to investigate the relationship between proximity to schools and education quality and the effects on housing prices. Therefore, we are going explore how housing prices react when separating between type of school, that is public and private as well as level of education, that is primary, middle and high school.

In the tables below, the regression results are presented, using the natural logarithm of housing transaction prices as the dependent variable. Moreover, to deal with heteroscedasticity, where the variance of the errors is not constant, robust standard errors were applied to the models during regression runs. The regression results show the effects of the independent variables included in the model on the dependent variable, which is the natural logarithm of housing sales price.

Hence, the coefficients describe the relationship between property prices proximity to schools and education quality, after controlling for other effects of property characteristics, time and location effects. To interpret the resulting regression coefficients, we used the following formula:

$$\%\Delta \text{ in sales price} = (e^{\text{coefficient}} - 1) * 100$$

Table 6: Effects of Proximity to Public and Private Schools on Property Values

	Model (1)	***Model (2)***
	Public Schools	***Private Schools***
Variables	*Coefficients S.E.*	*Coefficients S.E.*
Distance to nearest pub prim	-0.04896*** (0.00652)	
Distance to nearest pub mid	-0.0235*** (0.00393)	
Distance to nearest pub high	0.01412*** (0.00433)	
Distance to nearest priv prim		-0.03129*** (0.00331)
Distance to nearest priv mid		-0.02051*** (0.00171)
Distance to nearest priv high		-0.0204*** (0.00226)
Control Variables		
Housing Characteristics	Yes	Yes
Time Effects		
Sales Month	Yes	Yes
Location Effects		
Zip code	Yes	Yes
Constant	12.36*** (0.42351)	12.51*** (0.41732)
N	8.827	8.827
R^2	0.5876	0.5878
Adj. R^2	0.5754	0.5756

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, considering distance to the nearest public or private school at all levels as a continuous variable. The dependent variable is the natural logarithm of transaction prices. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 5% significance level and * at the 10% significance level.*

At first glance, we can conclude from the Table 6 that the key independent variable for examination demonstrates statistical significance and seem to have an impact on the dependent variable. Based on the findings, distance to the nearest public or private school at all levels of education is statistically significant, causing the respective effect on housing prices.

Model (1) in Table 3 examines public schools at the primary, middle, and high school levels. The key independent variable's effect on housing prices is statistically significant across all levels. Distance to the nearest primary and middle schools has negative coefficients of -0.04896 and -0.0235, significant at the 1% level. This means that, ceteris paribus, as the distance increases, property prices decrease by 4.78% and 2.32%, respectively. Conversely, distance to high schools has a positive coefficient of 0.01412, also significant at 1%, indicating a 1.42% increase in property prices. Finally, based on the results an R^2 of 58.76% was found. This suggests that approximately 58.76% of the variation in the dependent variable is explained by the independent variables in the regression model.

Model (2) in Table 6 examines private schools at the primary, middle, and high school levels. The key independent variable's effect on housing prices is statistically significant across all levels. Distance to all levels of education has negative coefficients of -0.03129, -0.02051, and -0.0204, significant at the 1% level. This means that, ceteris paribus, as the distance increases, property prices decrease by 3.02%, 2.03%, and 2.02% for primary, middle, and high schools, respectively. The decrease in home prices over distance is most severe for primary schools, followed by middle and then high schools. Finally, based on the results an R^2 of 58.78% was found. This suggests that approximately 58.78% of the variation in the dependent variable is explained by the independent variables in the regression model.

When comparing the individual effects of distance between public and private schools and level of education we observe the following. We found a negative coefficient for primary and middle school for both public and private schools. However, when comparing the effects of distance in the case of high schools, we established a positive coefficient for public high schools and a negative coefficient for private high schools. The fact that high schools present an opposite coefficient, compared to primary and middle schools might be attributed to the following reasons. On the one hand, residents might prioritize proximity to primary and middle school more compared to high schools, as they consider early education more beneficial for their children. Apart from that, high school students can easily travel greater distances by themselves as they are more independent as opposed to private high school students, who can use the school buses provided by their schools.

Table 9: Effects of Public School Quality Metrics on Housing Prices by Education Level

Variables	*Coefficients S.E.*	*Coefficients S.E.*	*Coefficients S.E.*
	Model 9 Primary	*Model 10 Middle*	*Model 11 High*
Quality Characteristics			
Number of students	-0.00006*	-0.00003	-0.00005***
	(0.00002)	(0.00002)	($8.38e^{-07}$)
Test scores	0.0089***	0.0323***	0.0405***
	(0.0024)	(0.0029)	(0.0032)
Student Progress	0.0287***	0.0132***	0.0093***
	(0.0030)	(0.0033)	(0.0022)
Equity Score	0.0064***	0.0192***	0.0359***
	(0.0039)	(0.0045)	(0.0038)
Suspension Rate	-0.4867***	-0.5814***	-0.1949***
	(0.0642)	(0.0642)	(0.0587)
Teacher Salary	$2.28*10^{-6}$***	$1.90*10^{-6}$***	$1.04*10^{-6}$***
	($2.79*10^{-7}$)	($2.83*10^{-7}$)	($3.13*10^{-7}$)
Student to teacher	-0.0067***	-0.0003	-0.0068**
	(0.0017)	(0.0018)	(0.0024)
Pupils to counselor	-0.00007	-0.00002	-0.0003***
	(0.00004)	(0.00004)	(0.00005)
Constant	11.88***	11.88***	12.05***
	(0.4135)	(0.4065)	(0.4121)
Control Variables			
Property Characteristics	Yes	Yes	Yes
Time effects			
Sales Month	Yes	Yes	Yes
Location effects			
Zip code	Yes	Yes	Yes
N	10.275	10.275	10.275
R^2	0.6051	0.6053	0.6051
Adj. R^2	0.5938	0.5939	0.5937

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, the effects of Public School Quality Metrics on Housing Prices by Education Level. The dependent variable is the natural logarithm of transaction prices. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 10% significance level and * at the 5% significance level.*

The next tables that will be created and analyzed in this research will focus on how education quality metrics have an impact on housing prices. The first table will focus on the effects of such metrics in the case of public schools at all levels, while the following table will focus on private schools. As has already been stated above, the quality metrics will be the same for all levels of education across the two types of schools. Consequently, in the case of public schools, metrics like number of students, test scores, student progress, equity score, teacher salary, student to teacher ratio, and student per counselor will be presented. On the contrary, in the case of private schools, metrics such as summary rating, test scores, student progress, equity score and student to teacher ratio will be stated.

At first glance, we can conclude from Table 9 that the key independent variables for examination demonstrate statistical significance and seem to have an impact on the dependent variable. Here, our analysis is focusing on the effects of public school education metrics on housing prices for the different levels of education. Having said that, now we are going to discuss the effects of these variables for all levels of education on housing prices.

In the case of **public primary schools**, Model 9 test scores, student progress and equity score have a positive coefficient of 0.0089, 0.0287 and 0,0064, which are statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), a unit increase in those variables is linked to a 0.89%, 2.91% and 0.64% increase in sales price respectively. A one unit increase in suspension rate is associated with a decrease in sales price by approximately 38.57%

In the case of **public middle schools**, Model 10 , test scores, student progress and equity score have a positive coefficient of 0.0323, 0.0132 and 0.0192, which are statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), a unit increase in those variables is linked to a 3.28%, 1.33% and 2.94% increase in sales price respectively. A one unit increase in suspension rate is associated with a decrease in sales price by approximately 44.09%.

In the case of **public high schools**, Model 11 test scores, student progress and equity score have a positive coefficient of 0.0405, 0.0093 and 0.0359, which are statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), a one unit increase in these variables is associated with 4.13%, 0.94% and 3.66% increase in sales price respectively. A one unit increase in suspension rate is associated with a decrease in sales price by approximately 17.71%

After having analyzed and interpreted the coefficients for the different quality metrics it is essential to also talk about why the coefficients differ across the three levels of education, that is primary, middle and high school. First and foremost, parents may place different levels of significance of education quality across the different levels of education. This justifies the fact why some parents may prioritize primary school over middle and high school, as it constitutes the formative years of a child's education in terms of writing, speaking and developing behavioral skills (Wen et al, 2017). Apart from that, each level of education has different courses and curriculum, which means that at primary level, education is considered easier and mandatory, whereas as a student proceeds to the next levels it becomes more difficult and demanding, and sometimes not obligatory, in the case of high schools. This reflects on different quality metrics like tests scores, number of students and student progress and can explain why different coefficients might arise (Owusu Edusei et al, 2007). Moreover, performance indicators play a different role across all levels of education, which influence how education quality is perceived. For instance, test scores, student progress, suspension rates and other metrics may vary in significance across primary, middle, and high schools. (Owusu Edusei et al., 2007; Wen et al., 2014). Furthermore, the racial and the ethnic groups within a school, at all levels of education, might affect to a great extent education quality, as they might be students from different countries, having issues adapt at the school environment. Finally, students with disabilities and learning difficulties might affect education quality at all levels of education, as they need extra support from teachers and perhaps different forms of exams. (Bayer et al, 2017; Kahn, 2024)

Finally, based on the results a similar R2 of 60.52% for all three levels of education was found. This suggests that approximately 60.52% of the variation in the dependent variable is explained by the independent variables in the regression model.

Table 10: Effects of Private School Quality Metrics on Housing Prices by Education Level

Variables	*Coefficients S.E.*	*Coefficients S.E.*	*Coefficients S.E.*
	Model 12 *Primary*	*Model 13* *Middle*	*Model 14* *High*
Quality Characteristics			
Number of students	-0.0003***	-0.0002***	-0.0001***
	(0.00005)	(0.00002)	(0.00002)
Summary scores	0.0591***	0.0568***	0.0376***
	(0.0085)	(0.0090)	(0.0103)
Test Score	0.0360***	0.0285***	0.0512***
	(0.0056)	(0.0064)	(0.0068)
Student Progress	0.0319***	0.0282***	0.0068***
	(0.0074)	(0.0087)	(0.0073)
Equity Score	0.0046***	0.0037	0.0333**
	(0.0073)	(0.0058)	(0.0088)
Student to teacher	-0.0041	-0.0050***	-0.0162***
	(0.0013)	(0.0009)	(0.0015)
Constant	12.07***	12.24***	12.17***
	(0.4168)	(0.4182)	(0.4178)
Control Variables			
Property Characteristics	Yes	Yes	Yes
Time effects			
Sales Month	Yes	Yes	Yes
Location effects			
Zip code	Yes	Yes	Yes
N	10.275	10.275	10.275
R^2	0.6054	0.6052	0.6051
Adj. R^2	0.5942	0.5939	0.5738

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, the effects of Private School Quality Metrics on Housing Prices by Education Level. The dependent variable is the natural logarithm of transaction prices. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 10% significance level and * at the 5% significance level.*

At first glance, we can conclude from Table 10 that the key independent variables for examination demonstrate statistical significance and seem to have an impact on the dependent variable. Here, our analysis is focusing on the effects of public school education metrics on housing prices for the different levels of education. Having said that, now we are going to discuss the effects of these variables for all levels of education on housing prices.

In the case of **private primary schools**, Model 12 test scores, student progress and equity score have a positive coefficient of 0.0360, 0.0319 and 0,0046, which are statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), a unit increase in those variables is linked to a 3.67%, 3.24% and 0.46% increase in sales price respectively.

In the case of **private middle schools**, Model 13 test scores and student progress have a positive coefficient of 0.0285 and 0.0282 which are statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), a unit increase in those variables is linked to a 2.89% and 2.86% increase in sales price respectively. Finally, student to teacher ratio contributes negatively on housing prices.

In the case of **private high schools**, Model 12 test scores, student progress and equity score have a positive coefficient of 0.0512, 0.0068 and 0.0333, which are statistically significant at 1% and 10% level of significance respectively. This means that, holding all other independent variables constant (ceteris paribus), a unit increase in those variables is linked to a 5.25%, 0.68% and 3.39% increase in sales price correspondingly. Finally, student to teacher ratio contributes negatively on housing prices.

Finally, based on the results a similar R2 of 60.53% for all three levels of education was found. This suggests that approximately 60.53% of the variation in the dependent variable is explained by the independent variables in the regression model.

A final remark when analyzing quality metrics and the effects on housing prices, is that when housing prices increase, education quality improves too and vice-verse, sustaining the overall relationship. These estimations are similar to the theoretical background that found a positive relationship between test scores and housing prices. Nguyen-Hoang & Yinger (2011) observed that a one unit increase in student test scores in the case of high schools tends to raise housing prices by 4%. Figlio & Lucas (2004) observed also the same positive relationship. Clapp, Nanda and Ross (2008) concluded that an improvement of the overall rating of a school is associated with a 1.3% and 1.4% increase in property values, in the case of public primary and middle school. Rosiers,

Lagana & Thériault (2001), found a negative relationship between number of students and housing prices. Seo & Simons, (2009) found a positive link between student performance and the effects on housing prices. Focusing on elementary level, they found that property values tend to rise by 3.39%. Bransigton & Haurin (2006), observed the same pattern with an increase in sales price of properties by 7.1%. Bayer et. al, (2017) found that when disadvantaged students are getting the help they need from teachers, this is associated with a 2% increase in sales price of the properties. Kahn (2024) also concluded that support from teachers contributes positively on housing prices.

At this stage and after having concluded our conversation regarding how different education metrics affect and to what extent housing prices, next in the analysis is to perceive if properties located close to high quality can actually influence property values. More specifically we want to examine if properties located close to high quality schools, within a radius of 700m, are actually sold more compared to properties located further away. In that respect, the final tables focus on the interaction variables between distance and high-quality schools. Moreover, to determine if a school is considered of a high quality, we created a dummy variable for all types and levels of education based solely on test scores and specifically if the grade is larger than 8 and lower than 10. We decided to proceed like that for two main reasons. Firstly, because according to Figlio & Lucas, (2004) test scores is considered an appropriate and effective measure, which allows us to determine education quality at all levels. Apart from that we wanted to select a variable that is common on both regression results. Subsequently, the interaction term was created as explained above.

Table 11: Effects of Interaction Variable on Housing Prices in the case of Public Schools

Variables	*Coefficients S.E.*	*Coefficients S.E.*	*Coefficients S.E.*
	Model 15 *Primary*	*Model 16* *Middle*	*Model 17* *High*
Interaction Variable			
DistanceGroup*HighQualityPubPrim	0.0272***		
	(0.0251)		
DistanceGroup1*HighQualityPubMid		0.0070*	
		(0.0250)	
DistanceGroup2*HighQualityPubHigh			0.0341**
			(0.0280)
Constant	12.16***	12.10***	12.16***
	(0.4176)	(0.4114)	(0.4152)
Control Variables			
Property Characteristics	Yes	Yes	Yes
Time effects			
Sales Month	Yes	Yes	Yes
Location effects			
Zip code	Yes	Yes	Yes
N	9.642	9.642	9.642
R^2	0.6044	0.6046	0.6047
Adj. R^2	0.5923	0.5925	0.5925

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, the effects of interaction variable in the case of public schools by Education Level. The dependent variable is the natural logarithm of transaction prices. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 10% significance level and * at the 5% significance level.*

Table 12: Effects of Interaction Variable on Housing Prices in the case of Private Schools

Variables	*Coefficients S.E.*	*Coefficients S.E.*	*Coefficients S.E.*
	Model 18 *Primary*	*Model 19* *Middle*	*Model 20* *High*
Interaction Variable			
DistanceGroup3*HighQualityPrivPrim	0.0137**		
	(0.0215)		
DistanceGroup4*HighQualityPrivMid		0.0100*	
		(0.0280)	
DistanceGroup5*HighQualityPrivHigh			0.0237**
			(0.0358)
Constant	12.21***	12.22***	12.23***
	(0.4189)	(0.4195)	(0.4196)
Control Variables			
Property Characteristics	Yes	Yes	Yes
Time effects			
Sales Month	Yes	Yes	Yes
Location effects			
Zip code	Yes	Yes	Yes
N	9.525	9.525	9.525
R^2	0.6054	0.6053	0.6059
Adj. R^2	0.5931	0.5930	0.5936

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, the effects of interaction variable in the case of public schools by Education Level. The dependent variable is the natural logarithm of transaction prices. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 10% significance level and * at the 5% significance level.*

At first glance, we can see that the key independent variables on Tables 11,12 are statistically significant. That means properties actually sold more, when locating within 700m from high quality schools. When focusing on Table 11, which considers public schools, we also have positive and statistically significant coefficients for the main variable of interest. When looking at the three levels of education we observe positive coefficients of 0.0272, 0.0070, and 0.0341. This means that holding, all other independent variable constant, properties located close to high quality schools at all levels, are actually sold at a premium of 2.76%, 0.7%, and 3.47%, compared to the ones that do not in the case of public primary, middle and high schools respectively.

Based on the results a similar R2 of 60.45% for all three levels of education was found. This suggests that approximately 60.45% of the variation in the dependent variable is explained by the independent variables in the regression model.

On the contrary, when focusing on Table 12, which considers private schools, we also have positive and statistically significant coefficients for the main variable of interest. When looking at the three levels of education we observe positive coefficients of 0.0137, 0.0100, and 0.0237. This means that holding, all other independent variable constant, properties located close to high quality schools at all levels, are actually sold more by 1.38%, 1%, and 2.40%, compared to the ones that do not in the case of private primary, middle and high schools respectively.

Based on the results we can see that the effects on housing prices is most severe in the case of high schools, followed by primary and then middle schools. This applies to both public and private schools. A potential reason for that might be that high schools is the last level of education, before students enter Universities, so parents are willing to pay a premium in order for their house to be close to such schools, especially if this associated with the provision of high education quality.

Based on the results a similar R2 of 60.55% for all three levels of education was found. This suggests that approximately 60.55% of the variation in the dependent variable is explained by the independent variables in the regression model.

These estimations are similar to the theoretical background that found a positive relationship between properties being close to high quality schools and housing prices. Qiu et al., (2019) found that properties located close to high quality primary schools in Beijing are associated with a 2.1% increase in sales price. Also, a study by the Department of Education in the UK, (2017) found that properties located close to high quality primary and middle schools is associated with an 8% and 6.8% increase in the sales price of properties.

5. CONCLUSIONS & LIMITATIONS

5.1. Conclusion

The main objective of this master thesis was to investigate the relationship between proximity to schools, education quality and the effects on housing prices, in the area of Allegheny County Pennsylvania. Our research can contribute to the existing literature in two ways. Initially, we explore what are the effects of this relationship are on housing prices, in a different State and County of the USA which has not been explored before. Secondly, we close the gap in the literature since this study is the first to consider both types of school, at all levels of education. The location of a school is significant for residents that are willing to pay a premium for their property in order to be at close proximity from such schools. Moreover, school quality is one other attribute that residents emphasize when choosing a property. Consequently, it is essential to understand that relationships as schools can determine to a great extent property values. The main question of this research was: "*What is the Impact of School Proximity and Quality on Housing Prices in Allegheny County: Comparing Public and Private Education.*

To answer this main question of our research, we created separate regression models, separating between types of school and level of education. Firstly, we investigated the effects of proximity to schools on housing prices, treating this variable as continuous and then as categorical. In the first case, all the coefficients presented negative coefficients and statistically significant indicating that when distance to schools increases, property values drop. Based on the regression results it is evident that property prices affect the most when a primary school is located far away from properties, as housing prices drop by 4.78% in the case of public schools and 3.08% in the case of private.

When also investigating the second relationship, we proved that good education quality is associated with high housing prices. More specifically, we our regression results suggested positive coefficients for specific variables like test scores and student progress indicating an increase in property values, and a negative coefficient for variables like suspension rates and number of students, suggesting a decrease. Based on our results we found that test scores tend to increase housing prices the most by 4.13% and 5,25% in the case of public and private high schools respectively. On the other hand, high suspension rates are associated with a decrease in housing prices by 44.09% in the case of public middle schools.

5.2. Limitations & Future Research

During this research we identified a number of limitations that should be taken into account for next research. Such limitations provide a balanced perspective on the research and highlight areas where further investigation may be conducted to better examine the respective relationship between proximity to schools, education quality and housing prices.

Initially, one major limitation of this thesis is the absence of a longitudinal analysis which would allow to capture the effects on housing prices for two or three years. This could be beneficial in order to look for any patterns observed over the years since housing markets often present fluctuations. This could be an essential part of this research, as here we focused only on 2022.

Secondly, by also examining and comparing different Counties, we could also observe how different areas behave and which are more volatile to prices. Based on the results we could acquire knowledge about the residential location decision of people, and simultaneously which types of school and at which level are valued the most, contributing positively on housing prices. This is because different Counties present variability in housing prices, type and number of schools.

Moreover, the inclusion of additional variables which were not taken into account for the regression models could potentially influence the final result. For instance, by including neighborhood and economic characteristics, such distance to the closest freeway, park, library, or data regarding people's salaries this could maybe alter the relationship. By also incorporating, variables like % of students who repeat a class, % of male or female teachers, % of teachers under a certain age for the second relationship this could lead us to different effects.

Finally, a significant area that remains unexplored is the impact of these attributes on rental markets. While most studies focus on property sales, investigating these effects from the perspective of renters could reveal how the results differ in terms of magnitude and significance. This can be attributed to the fact that the dynamics of sales and rental markets can vary considerably.

This research also suggests some policy recommendations that can be considered. One vital recommendation is improving public transportation systems to ensure that high-quality schools are accessible from diverse residential areas. This could help balance the housing market by reducing the premium on homes near such schools. Diverse access to high-quality education is ensured by encouraging mixed income housing close to schools. Moreover, investing in the quality of all schools can reduce disparities between districts and stabilize housing prices across regions. Apart from that, by implementing school choice programs, families have more flexibility of choosing where to live, without sacrificing the education quality of their children.

REFERENCES

Abidoye, R.B. and Chan, A.P.C. (2016). Critical determinants of residential property value: professionals' perspective. *Journal of Facilities Management*, 14(3), pp.283–300. doi:https://doi.org/10.1108/jfm-02-2016-0003.

Aliyev, K., Amiraslanova, M., Bakirova, N. and Eynizada, N. (2019). Determinants of housing prices in baku: Empirical analyses. *International Journal of Housing Markets and Analysis*, 12(2), pp.281–297. doi:https://doi.org/10.1108/ijhma-08-2018-0062.

Bayer, P., Casey, M., Ferreira, F. and McMillan, R. (2017). Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, [online] 102, pp.91–105. doi:https://doi.org/10.1016/j.jue.2017.07.004.

Bayoh, I., Irwin, E.G. and Haab, T. (2006). Determinants of residential location choice: How important are local public goods in attracting homeowners to central city locations?*. *Journal of Regional Science*, 46(1), pp.97–120. doi:https://doi.org/10.1111/j.0022-4146.2006.00434.x.

Bergantino, A.S., Biscione, A., de Felice, A., Porcelli, F. and Zagaria, R. (2022). Kindergarten Proximity and the Housing Market Price in Italy. *Economies*, 10(9), p.222. doi:https://doi.org/10.3390/economies10090222.

Black, S.E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2), pp.577–599. doi:https://doi.org/10.1162/003355399556070.

Brasington, D. and Haurin, D.R. (2006). Educational outcomes and house values: A test of the value added approach*. *Journal of Regional Science*, 46(2), pp.245–268. doi:https://doi.org/10.1111/j.0022-4146.2006.00440.x.

Brooks, C. and Sotiris Tsolacos (2010). *Real Estate Modelling and Forecasting*. Cambridge, Uk ; New York: Cambridge University Press.

Charles, S.L. (2012). Understanding the determinants of single-family residential redevelopment in the inner-ring suburbs of chicago. *Urban Studies*, 50(8), pp.1505–1522. doi:https://doi.org/10.1177/0042098012465908.

Chin, H.C. and Foong, K.W. (2006). Influence of School Accessibility on Housing Values. *Journal of Urban Planning and Development*, 132(3), pp.120–129. doi:https://doi.org/10.1061/(asce)0733-9488(2006)132:3(120).

Chiodo, A.J., Rubén Hernández-Murillo and Owyang, M.T. (2005). Nonlinear Hedonics and the Search for School District Quality. *RePEc: Research Papers in Economics*.

Clapp, J.M., Nanda, A. and Ross, S.L. (2008). Which school attributes matter? The influence of school district performance and demographic composition on property values. *Journal of Urban Economics*, 63(2), pp.451–466. doi:https://doi.org/10.1016/j.jue.2007.03.004.

Clark, D.E. and Herrin, W.E. (2000). The Impact of Public School Attributes on Home Sale Prices in California. *Growth and Change*, 31(3), pp.385–407. doi:https://doi.org/10.1111/0017-4815.00134.

data.wprdc.org. (n.d.). *Property Data with Geographic Identifiers - CKAN*. [online] Available at: https://data.wprdc.org/dataset/property-data-with-geographic-identifiers [Accessed 23 Oct. 2023].

Department of Education, (2017). Available at: https://assets.publishing.service.gov.uk/media/5a82a832ed915d74e3402e69/House_prices_and_schools.pdf [Accessed 20 Feb. 2024].

Edel, M. and Sclar, E. (1974). Taxes, spending, and property values: Supply adjustment in a tiebout-oates model. *Journal of Political Economy*, 82(5), pp.941–954. doi:https://doi.org/10.1086/260248.

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(1-2), pp.59–77. doi:https://doi.org/10.1016/j.jpubeco.2009.10.009.

Feng, H. and Lu, M. (2013). School quality and housing prices: Empirical evidence from a natural experiment in Shanghai, China. *Journal of Housing Economics*, 22(4), pp.291–307. doi:https://doi.org/10.1016/j.jhe.2013.10.003.

Figlio, D.N. and Lucas, M.E. (2004). What's in a Grade? School Report Cards and the Housing Market. *American Economic Review*, 94(3), pp.591–604. doi:https://doi.org/10.1257/0002828041464489.

Gibbons, S. and Machin, S. (2003). Valuing English primary schools. *Journal of Urban Economics*, 53(2), pp.197–219. doi:https://doi.org/10.1016/s0094-1190(02)00516-8.

Glindro, E.T., Subhanij, T., Szeto, J. and Zhu, H. (2011). Determinants of House Prices in Nine Asia-Pacific Economies. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.1333646.

GreatSchools.org. (n.d.). *Find A Great School*. [online] Available at: https://www.greatschools.org/ Accessed (25 Aug. 2023).

Gustely, R.D. (1976). Local taxes, expenditures and urban housing: A reassessment of the evidence. *Southern Economic Journal*, 42(4), p.659. doi:https://doi.org/10.2307/1056258.

Huang, P. and Hess, T. (2018). Impact of distance to school on housing price: Evidence from a quantile regression. *The Empirical Economics Letters*, 17(2), pp.149–156.

Huang, Y. and Dall'erba, S. (2020). Does Proximity to School Still Matter Once Access to Your Preferred School Zone Has Already Been Secured? *The Journal of Real Estate Finance and Economics*, 62(1). doi:https://doi.org/10.1007/s11146-020-09761-w.

Jayantha, W.M. and Lam, S.O. (2015). Capitalization of secondary school education into property values: A case study in hong kong. *Habitat International*, 50, pp.12–22. doi:https://doi.org/10.1016/j.habitatint.2015.07.011.

Jud, G.D. and Watts, J.M. (1981). Schools and Housing Values. *Land Economics*, 57(3), p.459. doi:https://doi.org/10.2307/3146025.

Kahn, M.E. (2024). Racial and ethnic differences in the financial returns to home purchases. *Real Estate Economics*. doi:https://doi.org/10.1111/1540-6229.12475.

Kane, T.J. (2006). School quality, neighborhoods, and housing prices. *American Law and Economics Review*, 8(2), pp.183–212. doi:https://doi.org/10.1093/aler/ahl007.

Kim, H., Park, S.W., Lee, S. and Xue, X. (2015). Determinants of house prices in seoul: A quantile regression approach. *Pacific Rim Property Research Journal*, 21(2), pp.91–113. doi:https://doi.org/10.1080/14445921.2015.1058031.

Liu, Z., Ye, J., Ren, G. and Feng, S. (2022). The Effect of School Quality on House Prices: Evidence from Shanghai, China. *Land*, 11(11), p.1894. doi:https://doi.org/10.3390/land11111894.

Lu, Y., Shi, V. and Pettit, C.J. (2023). The impacts of public schools on housing prices of residential properties: A case study of Greater Sydney, Australia. *ISPRS International Journal of Geo-Information*, [online] 12(7), p.298. doi:https://doi.org/10.3390/ijgi12070298.

Metz, N.E. (2015). Effect of distance to schooling on home prices. *Review of Regional Studies*, 45, pp.151–171. doi:https://doi.org/10.52324/001c.8060.

NCES (2017). *National Center for Education Statistics (NCES) Home Page, part of the U.S. Department of Education*. [online] Ed.gov. Available at: https://nces.ed.gov (Accessed 20th Aug. 2023).

Nguyen-Hoang, P. and Yinger, J. (2011). The capitalization of school quality into house values: A review. *Journal of Housing Economics*, 20(1), pp.30–48. doi:https://doi.org/10.1016/j.jhe.2011.02.001.

Oates, W.E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy*, 77(6), pp.957–971. doi:https://doi.org/10.1086/259584.

Owusu-Edusei, K., Espey, M. and Lin, H. (2007). Does Close Count? School Proximity, School Quality, and Residential Property Values. *Journal of Agricultural and Applied Economics*, 39(1), pp.211–221. doi:https://doi.org/10.1017/s1074070800022859.

Private School Review. (2019). *Private School Review - Profiles of USA Private Schools*. [online] Available at: https://www.privateschoolreview.com/ Accessed (25 Aug. 2023).

public.opendatasoft.com. (n.d.). *US Private Schools*. [online] Available at: https://public.opendatasoft.com/explore/dataset/us-private-schools/table/ [Accessed 24 Oct. 2023]. Accessed (25 Aug. 2023).

Qiu, L., Guo, D., Zhao, X. and Zhang, W. (2019). The value of school in urban China: a spatial quantile regression with housing transactions in Beijing. *Applied Economics*, pp.1–12. doi:https://doi.org/10.1080/00036846.2019.1683146.

Rosiers, F.D., Lagana, A. and Theriault, M. (2001). Size and proximity effects of primary schools on surrounding house values. *Journal of Property Research*, 18(2), pp.149–168. doi:https://doi.org/10.1080/09599910110039905.

Sah, V., Conroy, S.J. and Narwold, A. (2015). Estimating School Proximity Effects on Housing Prices: the Importance of Robust Spatial Controls in Hedonic Estimations. *The Journal of Real Estate Finance and Economics*, 53(1), pp.50–76. doi:https://doi.org/10.1007/s11146-015-9520-5.

Seo, Y. and Simons, R. (2009). The effect of school quality on residential sales price. *Journal of Real Estate Research*, 31(3), pp.307–328. doi:https://doi.org/10.1080/10835547.2009.12091255.

Sonstelie, J.C. and Portney, P.R. (1980). Gross rents and market values: Testing the implications of tiebout's hypothesis. *Journal of Urban Economics*, 7(1), pp.102–118. doi:https://doi.org/10.1016/0094-1190(80)90029-7.

Tiebout, C.M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5), pp.416–424.

Weimer, D.L. and Wolkoff, M.J. (2001). School Performance and Housing Values: Using Non-Contiguous District and Incorporation Boundaries to Identify School Effects. National Tax Journal, 54(2), pp.231–254. doi:https://doi.org/10.17310/ntj.2001.2.03.

Wen, H., Sheng-hua, J. and Xiao-yu, G. (2005). Hedonic price analysis of urban housing: An empirical research on hangzhou, china. *Journal of Zhejiang University-SCIENCE A*, 6(8), pp.907–914. doi:https://doi.org/10.1631/jzus.2005.a0907.

Wen, H., Xiao, Y., Hui, E.C.M. and Zhang, L. (2018). Education quality, accessibility, and housing price: Does spatial heterogeneity exist in education capitalization? *Habitat International*, 78, pp.68–82. doi:https://doi.org/10.1016/j.habitatint.2018.05.012.

Wen, H., Xiao, Y. and Zhang, L. (2017). School district, education quality, and housing price: Evidence from a natural experiment in Hangzhou, China. *Cities*, 66, pp.72–80. doi:https://doi.org/10.1016/j.cities.2017.03.008.

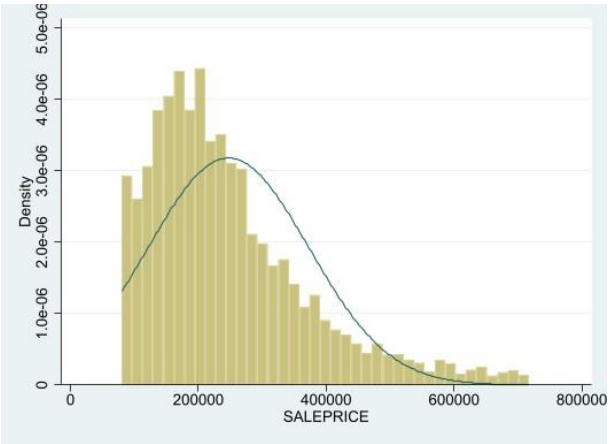
Wen, H., Zhang, Y. and Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in Hangzhou, China. *Habitat International*, 42, pp.155–163. doi:https://doi.org/10.1016/j.habitatint.2013.12.004.

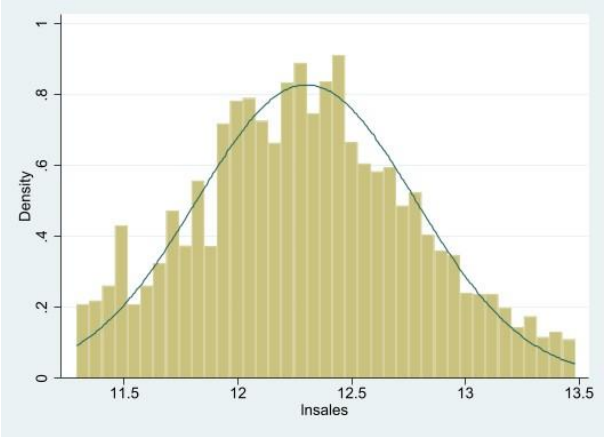
Wu, W., Zhang, W. and Dong, G. (2013). Determinant of residential location choice in a transitional housing market: Evidence based on micro survey from beijing. *Habitat International*, 39, pp.16–24. doi:https://doi.org/10.1016/j.habitatint.2012.10.008.

Zhou, T. (2018). Capitalization of Public School Quality into Housing Prices: An Empirical Analysis Based on School District Housing for Public Primary Schools in Shanghai. Chinese Studies, 07(04), pp.286–327. doi:https://doi.org/10.4236/chnstd.2018.74025

Appendix

Diagnostics Tests





Variance Inflation Factor

	VIF	1/VIF
Total rooms	3.018	.331
Bedrooms	2.631	.380
Ageofproperty	2.045	.489
Bsmtgarage	1.797	.556
Fullbaths	1.555	.643
Basement	1.448	.691
Stories	1.373	.729
Halfbaths	1.332	.751
Propertyzip	1.257	.796
Fireplaces	1.194	.837
Month	1.002	.998
Distance to nearest pub prim	1.348	.742
Distance to nearest pub mid	1.288	.777
Distance to nearest pub high	1.269	.788
Distance to nearest priv prim	1.212	.825
Distance to nearest priv mid	1.120	.893
Distance to nearest priv high	1.056	.947
Exteriorfinish	1.043	.959
Mean VIF	**1.499**	**.**

Diagnostic Tests

After having run the regression models in order to determine the effects of proximity to schools and education quality on housing prices, it is essential to run the appropriate and necessary diagnostic tests to make sure that all of the assumptions of OLS are hold. Once all of these have been determined it is necessary to conduct diagnostic tests to check the model's quality and accuracy. This section covers exactly this topic. According to Brooks & Tsolacos (2010), there are five assumptions of the OLS regression models that should be met, namely the error term has a population mean of zero (1), all independent variables are not correlated with the error term (2), observations of the error term are uncorrelated with each other (3), the error term has a constant variance (no heteroscedasticity) (4), the error term is normally distributed (5).

After conducting several tests including Breusch Pagan, Shapiro-Wilkes, Linktest the following problems were pinpointed. Breusch Pagan test is used to check for heteroscedasticity in a regression model. Heteroscedasticity occurs when the variance of the errors from a regression model is not constant across all levels of the independent variables. The Linktest is used to check for model specification. Finally, Shapiro-Wilkes test is used to check for the normality of the residuals.

The following assumptions are violated, namely heteroscedasticity and the fact that residuals are not normally distributed. To solve the issue of heteroscedasticity, the natural logarithm of the dependent variable of sales price was used. In the Appendix section of this research, after the transformation of this variable in a logarithmic form, a normal distribution is followed. Moreover, robust standard errors were also used at the end of each regression model, firstly for solving issues with heteroscedasticity and secondly for making estimates more accurate, effective and reliable. Finally, the non-normality of the residuals and the appropriate functional form was also tested, by applying the tests described above.

Another important test that always has to be conducted in order to assure that there are no biased estimates of our results is to check for multicollinearity. This shows the degree to which the independent variables included in the model are correlated with one another, which we do not want based on the assumptions mentioned in the first section of this chapter. To check for this issue, we used the correlation matrix, which shows the correlation between the dependent and independent variables included in the model. Finally, we also check for collinearity between variables using VIF which can also be found in the Appendix section of this research. When examining these two techniques together, we can say that there is no multicollinearity detected, since there is no correlation between the independent variables included in the model.

Model for the effects of distance on housing prices as a categorical variable

As has been mentioned before the previous model consider distance to the nearest school as a continuous variable. Here, we also want to discover the nonlinear effects of the different levels and types of school on housing prices. This will represent how different sets of distances affect and to what extend housing prices, based on the distance that properties refrain from schools. The categories that have been created for the different types of school, public and private, at all levels of education, primary, middle and high school will be presented below. For all levels of education, we have created different categories increasing by 700m. Since we use categorical variables, these take values from 0 to 1, and that means that properties take a value of 0, if they fall outside the defined range and 1 if they lie within that range. The distinction between public and private schools at all levels is presented below. The reason why we chose these specific categories in the case of public and private schools was based on two aspects. Firstly, we followed the same categorization to determine the effects on housing prices as Sah et al, 2015; Metz, 2015; Owusu-Edusei et al, 2007. Secondly, by looking at the max of the key independent variable, that is distance to the nearest school, either public or private, we determined how many different ranges of distances we will have. Consequently, we have three different sets of distance in the case of public primary schools, and four sets of distance for the rest of the types of school and level of education, without including the reference category. As has already been stated for our analysis, distance splines increase by 700m. According to Sah et al, using smaller distance ranges (700m) allows for a more thorough examination of how proximity to schools affects and to what extent housing prices. Finally, we decided to create ranges increasing by 700m for all types of school and level of education, in order to use that information to answer the last question of this research, which will also be explained later in our analysis.

A detailed overview of the reference category and the corresponding distance ranges can be found in the Notation Table at the end of this research paper.

Once we have explained and analyzed, why we will treat distance to school as a categorical variable the following models can be constructed:

$$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ group}_{i,t} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t} \quad \mathbf{(3)}$$

$$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ group1}_{i,t} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t} \quad \mathbf{(4)}$$

$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ group2_{i,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t}$ **(5)**

$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ group3_{i,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t}$ **(6)**

$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ group4_{i,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t}$ **(7)**

$Ln(P_{i,t}) = a + \mathbf{b1*X*Distance\ group5_{i,t}} + b2*Y*Housing\ Characteristics_{i,t} + b3*Zip\ code_{i,t} + b4*Sales\ Month_{i,t} + \varepsilon_{i,t}$ **(8)**

We have already explained the meaning of the variables above, so here we will focus only on the key independent variable considering distance as a categorical. Finally, the variables of interest distance group, distance group1, distance group2 are categorical variables referring to distance to the nearest public primary, middle and high school, whereas distance group3, distance group4, and distance group5 referring to distance to the nearest private primary, middle and high. X in all cases is a vector considering the different sets of distance which have been created to capture the effects on housing prices. In the model it is labelled as ***b1*X*Distance group.***

Table 3: Descriptive statistics for the main independent variable, as categorical

Variable	Obs	Mean	Std. Dev.	% of obs. (1)	Min	Max
distance group						
Dist. (2.122km+) reference	8.827	.093	.291	9.30%	0	1
Dist.(0.022-0.722km)	8.827	.304	.46	30.40%	0	1
Dist. (0.723-1.422km)	8.827	.391	.488	39.10%	0	1
Dist. (1.423-2.121km)	8.827	.212	.408	21.20%	0	1
distance group1						
Dist. (2.848km+) reference	8.827	.191	.393	19.10%	0	1
Dist. (0.048-0.748km)	8.827	.141	.348	14.10%	0	1
Dist. (0.749-1.446km)	8.827	.254	.435	25.40%	0	1
Dist. (1.447-2.146km)	8.827	.247	.431	24.70%	0	1
Dist. (2.147-2.847km)	8.827	.168	.374	16.80%	0	1
distance group2						
Dist. (2.866km+) reference	8.827	.186	.389	18.60%	0	1
Dist. (0.064-0.764km)	8.827	.115	.319	11.50%	0	1
Dist. (0.765-1.465km)	8.827	.244	.43	24.40%	0	1
Dist. (1.466-2.165km)	8.827	.263	.44	26.30%	0	1
Dist. (2.166-2.865km)	8.827	.192	.394	19.20%	0	1
distance group3						
Dist. (2.857km+) reference	8.827	.24	.427	24.00%	0	1
Dist. (0.053-0.753km)	8.827	.15	.357	15.00%	0	1
Dist. (0.754-1.454km)	8.827	.253	.435	25.30%	0	1
Dist. (1.455-2.155km)	8.827	.219	.414	21.90%	0	1
Dist. (2.156-2.856km)	8.827	.138	.345	13.80%	0	1
distance group4						
Dist. (2.837km+) reference	8.827	.524	.499	52.40%	0	1
Dist. (0.032-0.732km)	8.827	.077	.266	7.70%	0	1
Dist. (0.733-1.433km)	8.827	.128	.335	12.80%	0	1
Dist. (1.434-2.134km)	8.827	.139	.346	13.90%	0	1
Dist. (2.135-2.836km)	8.827	.132	.338	13.20%	0	1
distance group5						
Dist. (2.835km+) reference	8.827	.655	.475	65.50%	0	1
Dist. (0.031-0.731km)	8.827	.044	.206	4.40%	0	1
Dist. (0.732-1.432km)	8.827	.08	.272	8.00%	0	1
Dist. (1.433-2.133km)	8.827	.092	.29	9.20%	0	1
Dist. (2.134-2.834km)	8.827	.128	.334	12.80%	0	1

Table 7: Effects of Public-School Distance Categories on Housing Prices by Education Level

Variables	*Coefficients S.E.*	*Coefficients S.E.*	*Coefficients S.E.*
	Model 3 Primary	*Model 4 Middle*	*Model 5 High*
Distance Group (base 2.122km+)			
Dist. (0.022-0.722km)	0.1012*** (0.0148)		
Dist. (0.723-1.422)	0.0680*** (0.0143)		
Dist. (1.423-2.121)	0.0290** (0.0154)		
Distance Group1 (base 2.848km+)			
Dist. (0.048-0.748km)		0.0931*** (0.0139)	
Dist. (0.749-1.446km)		0.0585*** (0.0119)	
Dist. (1.447-2.146km)		0.0532*** (0.0022)	
Dist. (2.147-2.847km)		0.0066 (0.0130)	
Distance Group2 (base 2.866km+)			
Dist. (0.064-0.764km)			-0.0588*** (0.0152)
Dist. (0.765-1.465km)			-0.0110 (0.0124)
Dist. (1.466-2.165km)			-0.0160 (0.0120)
Dist. (2.166-2.865km)			-0.0441*** (0.0128)
Constant	12.17*** (0.4208)	12.17*** (0.4210)	12.21*** (0.4220)
Control Variables			
Property Characteristics	Yes	Yes	Yes

Time effects			
Sales Month	Yes	Yes	Yes
Location effects			
Zip code	Yes	Yes	Yes
N	8.827	8.827	8.827
R^2	0.5903	0.5901	0.5904
Adj. R^2	0.5780	0.5777	0.5779

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, the effects of Public School Distance Categories on Housing Prices by Education Level. The dependent variable is the natural logarithm of transaction prices. Distance group is a categorical variable for the distance to the nearest public primary school, distance group1 is a categorical variable for the distance to the nearest public middle school and distance group2 is a categorical variable for the distance to the nearest public high school. For distance group the reference category is set to be 0, and this denotes properties located at a distance of 2.122km and above from public primary schools. For distance group1 the reference category is set to be 0, and this denotes properties located at a distance of 2.848km and above from public middle schools. For distance group2 the reference category is set to be 0, and this denotes properties located at a distance of 2.866km and above from public high schools. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 10% significance level and * at the 5% significance level.*

This section here treats distance to nearest school, either public or private at all levels of education, as a categorical variable. This will provide useful insights regarding how properties located at different ranges from schools affect and to what extent housing prices. We are doing this in order to look for nonlinear effects regarding the relationship between proximity to schools and housing prices. The first table will examine this relationship in the case of public schools at the three levels of education, and then the next table will refer to private schools.

We can conclude from the Table 7 that the key independent variable for examination demonstrates statistical significance and seem to have an impact on the dependent variable. Here, our analysis is focusing on the effects of distance on housing prices, but we consider distance to schools as a categorical variable. As has been stated above, distance group, distance group1 and distance group2 are categorical variables referring to the distance to the nearest public primary, middle, and high school respectively. The reference categories in the case of public primary schools, refers to properties located at a distance of 2.122km and above from public primary schools, 2.848km and above in the case of middle and 2.866km and above in the case of high schools.

In the case of public primary schools, Model 3 properties located 0.022-0.722km away from schools have a positive coefficient of 0.1012 which is statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus),

properties within this range are sold at a premium of 10.65% compared to the reference category. At the next ranges, properties are also sold at a premium but at a smaller rate. (7.04% & 2.94%)

In the case of public middle schools, Model 4 properties located 0.048-0.748km away from schools have a positive coefficient of 0.0931 which is statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), properties within this range are sold at a premium of 9.76% compared to the reference category. At the next ranges, properties are also sold at a premium but at a smaller rate. (6.02% & 5.46%)

In the case of public high schools, Model 5 properties located 0.064-0.764km away from schools have a negative coefficient of -0.0588 which is statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), properties within this range are sold at a discount of 5.71% compared to the reference category. Also, properties at the last range are sold at a discount of 4.31%.

Finally, based on the results a similar R2 of 59.03% for all three levels of education was found. This suggests that approximately 59.03% of the variation in the dependent variable is explained by the independent variables in the regression model.

Table 8: Effects of Private School Distance Categories on Housing Prices by Education Level

Variables	*Coefficients S.E.*	*Coefficients S.E.*	*Coefficients S.E.*
	Model 6 *Primary*	*Model 7* *Middle*	*Model 8* *High*
Distance Group3 (base 2.857km+)			
Dist. (0.053-0.753km)	0.1668***		
	(0.0132)		
Dist. (0.754-1.454km)	0.1301***		
	(0.0113)		
Dist. (1.455-2.155km)	0.0719***		
	(0.0116)		
Dist. (2.156-2.856km)	0.0284*		
	(0.0132)		
Distance Group4 (base 2.837km+)			
Dist. (0.032-0.732km)		0.1666***	
		(0.0158)	
Dist. (0.733-1.433km)		0.1367***	
		(0.0126)	
Dist. (1.434-2.134km)		0.0619***	
		(0.0120)	
Dist. (2.135-2.836km)		0.0214	
		(0.0121)	
Distance Group5 (base 2.835km+)			
Dist. (0.031-0.731km)			0.2459***
			(0.0198)
Dist. (0.732-1.432km)			0.0723***
			(0.0147)
Dist. (1.433-2.133km)			0.0634***
			(0.0137)
Dist. (2.134-2.834km)			0.0208
			(0.0120)
Constant	12.19***	12.21***	12.18***
	(0.4207)	(0.4214)	(0.4218)

Control Variables			
Property Characteristics	Yes	Yes	Yes
Time effects			
Sales Month	Yes	Yes	Yes
Location effects			
Zip code	Yes	Yes	Yes
N	8.827	8.827	8.827
R^2	0.5874	0.5882	0.5881
Adj. R^2	0.5752	0.5760	0.5759

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

*Note: The table shows the regression results, the effects of Private School Distance Categories on Housing Prices by Education Level. The dependent variable is the natural logarithm of transaction prices. Distance group3 is a categorical variable for the distance to the nearest private primary school, distance group4 is a categorical variable for the distance to the nearest private middle school and distance group5 is a categorical variable for the distance to the nearest private high school. For distance group3 the reference category is set to be 0, and this denotes properties located at a distance of 2.857km and above from private primary schools. For distance group4 the reference category is set to be 0, and this denotes properties located at a distance of 2.837km and above from private middle schools. For distance group5 the reference category is set to be 0, and this denotes properties located at a distance of 2.835km and above from private high schools. Parentheses include standard errors, *** indicates that the variable is statistically significant at the 1% significance level,** at the 10% significance level and * at the 5% significance level.*

We can conclude from the Table 8 that the key independent variable for examination demonstrates statistical significance and seem to have an impact on the dependent variable. Here, our analysis is focusing on the effects of distance on housing prices, but we consider distance to schools as a categorical variable. As has been stated above, distance group3, distance group4 and distance group5 are categorical variables referring to the distance to the nearest private primary, middle, and high school respectively. The reference categories in the case of public primary schools, refers to properties located at a distance of 2.857km and above from private primary schools, 2.837km and above in the case of middle and 2.835km and above in the case of high schools.

In the case of private primary schools, Model 6 properties located 0.053-0.753km away from schools have a positive coefficient of 0.1668 which is statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), properties within this range are sold at a premium of 18.15% compared to the reference category.

Properties at the next ranges, are sold at a premium, but this amount is decreasing compared to the initial (13.89%, 7.45%, 5%, 2.88%)

In the case of private middle schools, Model 7 properties located 0.032-0.732km away from schools have a positive coefficient of 0.1666 which is statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), properties within this range are sold at a premium of 18.13% compared to the reference category. Properties at the next ranges are sold at a premium, but this amount is decreasing compared to the initial (14.65%, 6.39%).

In the case of private high schools, Model 8 properties located 0.031-0.731km away from schools have a positive coefficient of 0.2459 which is statistically significant at 1% level of significance. This means that, holding all other independent variables constant (ceteris paribus), properties within this range are sold at a premium of 27.88% compared to the reference category. Properties at the next ranges, are sold at a premium, but this amount is decreasing compared to the initial (7.50%, 6.55%)

Finally, based on the results a similar R2 of 58.79% for all three levels of education was found. This suggests that approximately 58.79% of the variation in the dependent variable is explained by the independent variables in the regression model.

After analyzing Tables 7 and 8, we can conclude that these estimations are similar to the theoretical background that found a positive relationship between proximity to schools and housing prices. Owusu-Edusei, et al. (2007), found that for houses located within 800m from elementary and middle schools there is a positive effect on housing prices by 7.4% and 10% respectively. Huang and Hess, (2018) found that for each additional mile a house is away from the closest elementary school, middle school and high school, the value of a house is expected to decline. Metz (2015), found that properties located at a distance less than 500ft from an elementary school, are sold higher by 2.5%, whilst properties located less than 1000ft from middle and high school a 3.2% and 2% increase can be observed.

To sum up, we could say that when analyzing the effects of distance as a categorical variable, considering all three levels of education and types of school different pricing patterns emerge. This can be attributed to a plethora of reasons such as parent's priorities, location desirability, safety issues, and demographic factors Each of the above factors contribute to the overall attractiveness of a neighborhood, affecting housing prices and demand for housing.

Notation Glossary

Empirical Model	
$Ln(P_{i,t})$	Natural logarithm of transaction prices
I	Transaction i = 1, 2, …, N
t	Time t = 1, 2, …, N refer to months that the property was transacted
Cons	Constant to be estimated
βi	Parameters to be estimated
Totalrooms	Categorical variable, describing the total rooms that the property has
Bedrooms	Categorical variable, describing the number of bedrooms that the property has
Fullbaths	Categorical variable, describing the number of full baths the property has
Halfbaths	Categorical variable, describing the number of half baths the property has
Fireplaces	Categorical variable, describing the number of fireplaces the property has
Exteriorfinish	Categorical variable, describing the exterior material that the property is built
Stories	Continuous variable, describing the number of stories the property has
Garage	Categorical variable, describing the number of garages the property has
Basement	Categorical variable, describing the number of basements the property has
Age of property	Continuous variable, describing the age of the property
Sales Month	Categorical variable, describing the transaction month that the property was sold
Yearblt	Categorical variable, describing the year that the property was built
Zip code	Categorical variable, describing the zip code of where the property sold is located
Distance to nearest pub prim	Continuous variable, describing the distance of properties to the nearest public primary school
Distance to nearest pub mid	Continuous variable, describing the distance of properties to the nearest public middle school
Distance to nearest pub high	Continuous variable, describing the distance of properties to the nearest public high school
Distance to nearest priv prim	Continuous variable, describing the distance of properties to the nearest private primary school
Distance to nearest priv mid	Continuous variable, describing the distance of properties to the nearest private middle school
Distance to nearest priv high	Continuous variable, describing the distance of properties to the nearest private high school
distance group	Categorical variable referring to the variable of interest Distance to nearest pub prim
0	Reference category, that ranges from (2.122 and above), and takes values 0 if the property falls outside that range, and 1 if it lies within this range, denoting the distance

	that properties refrain from public primary schools at that range.
1	Ranges from (0.022-0.722) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public primary schools at that range.
2	Ranges from (0.723-0.1.422) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public primary schools at that range.
3	Ranges from (1.422-2.121) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public primary schools at that range.
distance group1	Categorical variable referring to the variable of interest Distance to nearest pub mid
0	Reference category, that ranges from (2.848 and above), and takes values 0 if the property falls outside that range, and 1 if it lies within this range, denoting the distance that properties refrain from public middle schools at that range.
1	Ranges from (0.048-0.748) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public middle schools at that range.
2	Ranges from (0.749-1.446) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public middle schools at that range.
3	Ranges from (1.447-2.146) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public middle schools at that range.
4	Ranges from (2.147-2.847) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public middle schools at that range.
distance group2	Categorical variable referring to the variable of interest Distance to nearest pub high
0	Reference category, that ranges from (2.866 and above), and takes values 0 if the property falls outside that range, and 1 if it lies within this range, denoting the distance that properties refrain from public high schools at that range.
1	Ranges from (0.064-0.764) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public high schools at that range.
2	Ranges from (0.765-1.465) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public high schools at that range.
3	Ranges from (1.466-2.165) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public high schools at that range.
4	Ranges from (2.166-2.865) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from public high schools at that range.
distance group3	Categorical variable referring to the variable of interest Distance to nearest priv prim
0	Reference category, that ranges from (2.857 and above), and takes values 0 if the property falls outside that range, and 1 if it lies within this range, denoting the distance that properties refrain from private primary schools at that range.
1	Ranges from (0.053-0.753) and takes values 0 if the property falls outside that range

	and 1 if it lies within this range, denoting the distance that properties refrain from private primary schools at that range.
2	Ranges from (0.754-1.454) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private primary schools at that range.
3	Ranges from (1.455-2.155) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private primary schools at that range.
4	Ranges from (2.156-2.856) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private primary schools at that range.
distance group4	Categorical variable referring to the variable of interest Distance to nearest priv mid
0	Reference category, that ranges from (2.837 and above), and takes values 0 if the property falls outside that range, and 1 if it lies within this range, denoting the distance that properties refrain from private middle schools at that range.
1	Ranges from (0.032-0.732) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private middle schools at that range.
2	Ranges from (0.733-1.433) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private middle schools at that range.
3	Ranges from (1.434-2.134) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private middle schools at that range.
4	Ranges from (2.135-2.836) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private middle schools at that range.
distance group5	Categorical variable referring to the variable of interest Distance to nearest priv high
0	Reference category, that ranges from (2.835 and above), and takes values 0 if the property falls outside that range, and 1 if it lies within this range, denoting the distance that properties refrain from private high schools at that range.
1	Ranges from (0.031-0.731) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private high schools at that range.
2	Ranges from (0.732-1.432) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private high schools at that range.
3	Ranges from (1.433-2.133) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private high schools at that range.
4	Ranges from (2.134-2.834) and takes values 0 if the property falls outside that range and 1 if it lies within this range, denoting the distance that properties refrain from private high schools at that range.
District	Categorical Variable denoting the school district
No. of students pub prim	Continuous variable denoting the total number of students public primary schools have
Test score pub prim	Continuous variable denoting how public primary school students perform at tests

Student progress pub prim	Continuous variable denoting how public primary school students perform/progress
Equity Score pub prim	Continuous variable denoting the extent to which public primary schools help the disadvantaged students
Suspension rate pub prim	Continuous variable denoting the suspension rates in the case of public primary schools
Student to teacher pub prim	Continuous variable denoting student to teacher ratio for public primary schools
Teacher salary pub prim	Continuous variable denoting teachers' salary for public primary schools
Pupil to counselor pub prim	Continuous variable denoting student to counselor ratio for public primary schools
No. of students pub mid	Continuous variable denoting the total number of students public middle schools have
Test score pub mid	Continuous variable denoting how public middle school students perform at tests
Student progress pub mid	Continuous variable denoting how public middle school students perform/progress
Equity Score pub mid	Continuous variable denoting the extent to which public middle schools help the disadvantaged students
Suspension rate pub mid	Continuous variable denoting the suspension rates in the case of public middle schools
Student to teacher pub mid	Continuous variable denoting student to teacher ratio for public middle schools
Teacher salary pub mid	Continuous variable denoting teachers' salary for public middle schools
Pupil to counselor pub mid	Continuous variable denoting student to counselor ratio for public middle schools
No. of students pub high	Continuous variable denoting the total number of students public high schools have
Test score pub high	Continuous variable denoting how public high school students perform at tests
Student progress pub high	Continuous variable denoting how public high school students perform/progress
Equity Score pub high	Continuous variable denoting the extent to which public high schools help the disadvantaged students
Suspension rate pub high	Continuous variable denoting the suspension rates in the case of public high schools
Student to teacher pub high	Continuous variable denoting student to teacher ratio for public high schools
Teacher salary pub high	Continuous variable denoting teachers' salary for public high schools
Pupil to counselor pub high	Continuous variable denoting student to counselor ratio for public high schools
No. of students priv prim	Continuous variable denoting the total number of students private primary schools have
Summary Rating priv prim	Continuous variable denoting the overall score of private primary schools
Test scores priv prim	Continuous variable denoting how private primary school students perform at tests
Student progress priv prim	Continuous variable denoting how private primary school students perform/progress
Equity Score priv prim	Continuous variable denoting the extent to which private primary schools help the disadvantaged students
Student to teacher priv prim	Continuous variable denoting student to teacher ratio for private primary schools
No. of students priv mid	Continuous variable denoting the total number of students private middle schools have

Summary Rating priv mid	Continuous variable denoting the overall score of private middle schools
Test scores priv mid	Continuous variable denoting how private middle school students perform at tests
Student progress priv mid	Continuous variable denoting how private middle school students perform/progress
Equity Score priv mid	Continuous variable denoting the extent to which private middle schools help the disadvantaged students
Student to teacher priv mid	Continuous variable denoting student to teacher ratio for private middle schools
No. of students priv high	Continuous variable denoting the total number of students private high schools have
Summary Rating priv high	Continuous variable denoting the overall score of private high schools
Test scores priv high	Continuous variable denoting how private high school students perform at tests
Student progress priv high	Continuous variable denoting how private high school students perform/progress
Equity Score priv high	Continuous variable denoting the extent to which private high schools help the disadvantaged students
Student to teacher priv high	Continuous variable denoting student to teacher ratio for private high schools
High quality pub prim	Dummy variable that takes values of 1 if test scores pub prim>=8, and 0 otherwise
High quality pub mid	Dummy variable that takes values of 1 if test scores pub mid>=8, and 0 otherwise
High quality pub high	Dummy variable that takes values of 1 if test scores pub high>=8, and 0 otherwise
High quality priv prim	Dummy variable that takes values of 1 if test scores priv prim>=8, and 0 otherwise
High quality priv mid	Dummy variable that takes values of 1 if test scores priv mid>=8, and 0 otherwise
High quality priv high	Dummy variable that takes values of 1 if test scores priv high>=8, and 0 otherwise

Variable	Obs	% of Obs.
distance group	**8.827**	
Dist. (2.122km+) reference	825	9.30%
Dist.(0.022-0.722km)	2.685	30.40%
Dist. (0.723-1.422km)	3.450	39.10%
Dist. (1.423-2.121km)	1.867	21.20%
distance group1	**8.827**	
Dist. (2.848km+) reference	1.684	19.10%
Dist. (0.048-0.748km)	1.241	14.10%
Dist. (0.749-1.446km)	2.239	25.40%
Dist. (1.447-2.146km)	2.176	24.70%
Dist. (2.147-2.847km)	1.487	16.80%
distance group2	**8.827**	
Dist. (2.866km+) reference	1.639	18.60%
Dist. (0.064-0.764km)	1.016	11.50%
Dist. (0.765-1.465km)	2.157	24.40%
Dist. (1.466-2.165km)	2.320	26.30%
Dist. (2.166-2.865km)	1.695	19.20%
distance group3	**8.827**	
Dist. (2.857km+) reference	2.121	24.00%
Dist. (0.053-0.753km)	1.324	15.00%
Dist. (0.754-1.454km)	2.231	25.30%
Dist. (1.455-2.155km)	1.936	21.90%
Dist. (2.156-2.856km)	1.215	13.80%
distance group4	**8.827**	
Dist. (2.837km+) reference	4.626	52.40%
Dist. (0.032-0.732km)	676	7.70%
Dist. (0.733-1.433km)	1.134	12.80%
Dist. (1.434-2.134km)	1.228	13.90%
Dist. (2.135-2.836km)	1.163	13.20%
distance group5	**8.827**	
Dist. (2.835km+) reference	5.784	65.50%
Dist. (0.031-0.731km)	390	4.40%
Dist. (0.732-1.432km)	709	8.00%
Dist. (1.433-2.133km)	815	9.20%
Dist. (2.134-2.834km)	1.129	12.80%