THE EFFECTS OF SCHOOL QUALITY AND OTHER UNOBSERVED NEIGHBOURHOOD CHARACTERISTICS ON PROPERTY SALE PRICES

WILLIAM SEREMI THANTSHA

Student Number: 721032

The School of Construction Economics and Management

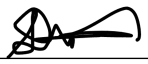
University of the Witwatersrand

A research report submitted to the faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Masters of Science in Engineering.

Johannesburg, October 2021

Declaration

I, William Seremi Thantsha, hereby declare that the research work conducted in this research report is my own unaided work and that all the sources used to support the research study have been accurately acknowledged and reported throughout the document, and that this research report has not been previously submitted, in its entirety, to any University to obtain an academic qualification.

Signature:  Date: 22 October, 2021

Abstract

Purpose – The research aims to investigate the effects of school quality and other unobserved determinants in generating variations in property sale prices. Previous studies have empirically examined the quality of private and public schools without a standard proxy that is accepted in literature. As a result, the research extends knowledge on this dynamic and measures the value of the quality of a former model C school (section 21) and its effect in property price changes in the local markets of the City of Johannesburg (region F).

Research Design/Methodology/Approach – The research adopts the hedonic regression models. It employs a multivariate analysis approach to evaluate and quantify the relationship between the identified predictor variables and property sale prices. The dependent variable was the property sale prices registered in the deeds registry from 2010 to 2020. The independent variables were school quality (i.e. pass rates, sport rankings and quality of facilities) and structural attributes (i.e. erf size, the number of bathrooms and the number of bedrooms) and the year of property sales. The regression models used in the research have been widely applied in previous studies which was useful to ensure validity and reliability of the estimated parameters.

Findings – The research finds that both models (I and II) explain over 26% of the variation in the log of the property sale prices. The estimated price semi-elasticity of the quality of Jeppe Boys High School results to a 2.8% increase in property sale prices in both sampled areas. The dummy variable of the registered year of each property sale was statistically significant, but exhibited a weak relationship with property transaction prices. The logged erf size had the highest and strongest relationship with sale prices. Property structural attributes in both models contributed a capitalisation rate below 10%. The empirical evidence indicates that the VIF's of the explanatory variables were all considerably low (< 5), implying that plausible impacts of multicollinearity were effectively cured and under control. Further proving there were no signs of spatial autocorrelation, spatial autoregression, and heteroscedasticity being problematic in biasing the estimated regression coefficients.

Research Limitations/Implications – The researcher also attempted to estimate the effect of traffic noise in property transactions. As a result of time-constraints and limited resources reliable data could not be obtained which, however, did not affect the overall findings or conclusions. Further research is ad-hoc to contribute insight on the role of such dynamics in property markets.

Practical implications - The results indicate that quality former model C schools play a role in explaining the variation of property sale prices in the short-run. The research provides insight that policymakers, investors and real estate developers can adopt to enable them to isolate univariates that have high-sensitivity issues of spatial correlation and spatial autoregression based on the data and the appropriate functional forms used in hedonic modeling.

Originality/Value – The main contribution is uncovering the relationship between the school quality of a former model C school and property transaction prices in the local markets, using Kensington and Observatory as sampled areas. The research also moves further to test for the presence of multicollinearity, heteroscedasticity, omitted variable bias, particularly, spatial autoregression and spatial autocorrelation problems in hedonic modeling which are prevalent in some previous studies.

Key words: property sale prices; school quality; hedonic pricing models; location; environmental determinants; property-related structural attributes.

Dedication

The research report is a tribute to my dearly departed grandmother, grandfather and brother, Modikwa Albertina Kekana (1945 - 2003), Madimetja Fred Kekana (1938 - 2005), and Fred Thabiso Masilela (1985 - 2008).

Your presence in our lives is sincerely missed.

Acknowledgment

I extent my hand of gratitude to my beloved family who have been understanding and provided their anointing love and endless support during the difficult times I have experienced this year and during the course of my studying. The successful completion of this Masters thesis would have not been possible without your assistance. To Kedibone Maggie Thantsha, Maisha Alpheus Thantsha, Chuene Selly Thantsha, Moshupologo Sally Thantsha and Leonard Mantsho, you have been the pillar of my endurance that has kept me moving forward. To Aly Karam, thank you for the unconditional care throughout my academic years. Through your constant believes, emotional support and prayers that have acted as a capstone throughout this journey, I have managed to get this far and I am thankful, God bless you all in abundance.

I would also like to extend my appreciation to my supervisor Professor Samuel Azasu, I humbly thank you for your guidance, encouragements, fruitful and insightful discussions throughout this research experience.

Table of Contents

List of Figures

List of Tables

List of Graphs

List of Maps

List of Acronyms

Acronym	Description
CAMA	*computer-assisted mass appraisal*
CBA	*Cost-Benefit Analysis*
CBD	*Central Business District*
CEM	*Construction, Economics and Management*
CNY	*Chinese Yuan Renminbi*
CoJ	*City of Johannesburg*
CPEP	*California Public Education Partnership*
CSA	*Cricket South Africa*
DID	*Difference-in-difference (DID) method*
DoE	*Department of Education*
DRT	*Department of Road and Transport*
EMM	*Johannesburg Metropolitan Municipality*
EPD	*Performance and Development Programme*
GIS	*Geotechnical Information System*
GDP	*Gross Domestic Product*
GWR	*Geographical Weighted Regression model*
HPM	*House Pricing Model*
IDP	*Integrated Development Plan*
JHSB	*Jeppe High School for Boys*
JMSA	*Johannesburg Metropolitan Statistical Area*
LPG's	*Local Public Goods*
MRA	*Multi-Variate Regression Analysis*
NDP	*National Development Plan*
OECD	*Organisation for Economic Co-operation and Development*
OLS	*Ordinary Least Square model*
Quintile 4 Schools	*Former model C schools in Quintile 4*
RSA	*Republic of South Africa*
SAFMA	*South African Facilities Management Association*
SANS	*South African National Standards*
SASA	*South African Schools Act, 1996*
Section 20	*Section 20 of the South African Schools Act, 1996*

Acronym	Description
Section 21	*Section 21 of the South African Schools Act, 1996*
SEU	*Social Exclusion Unit*
SGB	*School Governing Body*
SP	*Sale Price*
SPSS	*Statistical Package for the Social Sciences*
SQ	*School Quality*
TE model	*Treatment Effect Model*
UK	*United Kingdom*
USA	*United States of America*
VECM	*Vector Error Correction Model*
VIF	*Variance-of-Influence*
WHO	*World Health Organisation*
WTP	*Willingness to Pay*

Chapter One: Introduction and Background

1.1. Introduction

This chapter outlines the key elements that provides a background for the research motivation, aims and objectives. It also sheds light on the methodological techniques employed in the research and the possible contributions towards policy-design. To end with, the chapter outlines the report structure.

1.2. Research Background

The role of real estate constitutes paramount value to the development of many countries, household's wealth portfolio, and the overall social growth. For decades, property sale prices have exhibited highly localised differences that cannot be explained solely by variations in the physical attributes compositing a property (Gibbons, 2011). As a result, continuous attempts to understand the idiosyncrasies of property markets, particularly the underlying dynamics of macroeconomic and microeconomic attributes that contribute to the implicit price of a property, are mostly at the national and regional scale as opposed to the local levels. Wilhelmsson and Ceccato (2018) note that scholars have long applied multiple regression models to make inferences to a series of transactions on the demand, supply and value attributed to environmental and locational determinants. In reality, untangling this bundle of composite characteristics is a mammoth task as a result of selecting the appropriate functional form for the data used. This considers accurately estimating the effects of unique characteristics with regards to structural attributes, locational and environmental quality in generating property sale price variations.

The property market is highly heterogeneous and this cannot be overemphasized. Dodds (2010), Owusu-Ansah (2010; 2018) Wen et al. (2019), argued that some of the studies using Multi-Variate Regression Analysis (MRA) (formerly known as, *hedonic price modeling*) to examine a plethora of property indicators have drawn major criticisms. The research adds further that studies failed to incorporate the full-range of relevant neighbourhood characteristics and other important unobserved property attributes in some regression models, particularly, during property valuations. The problem linked with large datasets is the difficulty to identify possible outliers in the hedonic model (Wilhelmsson, 2019). Chiodo et al. (2010) earlier noted that, coefficient estimates of parameters have become plagued by endogeneity issues and standard omitted variable bias that are usually caused by reversed causality or selection bias. In addition, other identified problems that can create bias estimated coefficients is the presence of parameter heterogeneity and spatial dependence. Greene (2008) also

identified and described some of these issues as main sources of hedonic problems that included multicollinearity and heteroscedasticity, which are further discussed in the literature review chapter. Wild (2016) later added that such problems are exacerbated by the exorbitant use of numerous variables, average coefficient measures, and a mixture of regression methods to test the same phenomenon as seen in few previous studies. Hence, providing valid information on property sale prices to policymakers, real estate developers, households and banks responsible for most valuations is deemed as necessary in the research. For future studies, adopting property databases from reliable sources will contribute fruitful insight in further exploring and understanding property market dynamics and issues raised in the research.

As a backdrop for the research, the lack of a standardized proxy to accurately estimate the added-value of school quality has resulted in scholars, including Ka-wai (2010), Wild (2016), and Hummel (2017) adopting a myriad of varying school attributes to represent the real value of school quality on burgeoning prices in those areas being investigated. To date, high uncertainties are still associated with some of the multivariate estimations with scholarly debates interrogating the reliability of the findings due to the cumbersome nature of the dataset or school inputs and outputs used in modeling. This is further compounded by the selection of inappropriate functional forms that generally result in problems such as, spatial autoregression and spatial correlation (in the error terms), which the research also addresses. The research constructed a proxy based on the available school quality parameters to define and help explain how the variable is employed (see section 1.3.2). School quality as a variable has been dismantled exponentially, and it is notable that not all input and output measures embodying the quality of a school are valued equally. This is shown by scholars either adopting; expenditure per pupil, household income per pupil, salary per teacher, proficiency rates, Mathematics, English and Science average test-scores, and reading-scorecards. The use of these explanatory dummy variables as proxies to represent the real educational quality of the overall school is somewhat acutely obtuse, inconsistent and bias (also see, Tiebout, 1956; Oates, 1969; Kane, et al., 2005; Fack, et al., 2009; Nguyen-Hoang, 2011; Khachiyan, 2013; Yadavalli and Florax, 2014; and Wild, 2016). Most of these studies found coefficient estimates based on average values to represent the real price of each of the parameters in housing prices.

Despite the literature criticisms, overcoming the relenting heirlooms of bias statistical findings needs prioritisation to avoid overlooking the intrinsic value-adding nature of quality schools on property sale prices. Comprehending the overall influence of good neighbourhood schools on property price indices has not been immensely necessitated in the South African Urban Economics literature. To untangle this, the research alternatively focused on a former model

C (Section 21) school situated in region F of the City of Johannesburg (CoJ), using the hedonic pricing model. This aimed to improve information reliability instead of only focusing on the effects of public or private educational institutions, which are vastly investigated in previous studies using large datasets or repetitive variables with different sample sizes. This attempt revealed that substantial discrepancies are observable in property price changes when a different school type is examined. The focus of the research is grounded on four highly relevant policy issues - school quality, location, structural attributes and traffic noise. A comprehensive analysis of the empirical findings highlights both 'the known' and 'the unknown' on these underlying policy issues. The research aimed to unravel this phenomenon to purposively aid in the evaluation and policy-design processes. Although serious empirical hurdles persist, argumentatively the research on property sale prices (based on vigilant research designs) has generated relevant and credible non-bias valuations of the value of place-specific amenities and physical attributes.

1.2.1. The Attributable Implicit-Value of School Quality on Property Sale Prices

Gibbons (2011) highlights that the implicit value of properties provide shadow prices of non-market goods, which has been immensely recognised in environmental economics for their contribution to cost-benefit analyses. Quality educational facilities within urban vicinities (or around employment areas) are highly valued by households and potential homebuyers. Inconspicuously, their housing purchasing behaviours reveal a somewhat lack of comprehensive understanding of the actual monetary value or effect quality schools have on their housing selling prices (Wen, et al., 2019). Machin (2011) asserts that numerous locational factors can act as determinants of housing expenditures due to their state of quality that can be measured in value and capitalised into housing selling prices. This includes, but is not limited to, other quality public goods such as recreational areas and facilities (e.g. parks, swimming pools, and tennis courts), community services (e.g. libraries, clinics), and landmarks (e.g. transport nodes, museums, historic buildings and statues) within close-proximity to residential properties. Among these determinants, school quality in the local context has not been applied effectively as there is little attention given to this subject area and some of the literature results are arguably inconclusive.

For developed markets, extensive work investigating the relationship between house prices and other property indicators such as school quality is widely seen in countries such as China, United Kingdom (UK), and United States of America (USA) (see, for example, Kane, et al., 2005; Bayer, et al., 2007; Clapp, et al., 2008; and Annexure A). In China, for example, research conducted on the housing market in Shanghai includes scholars such as Feng et al. (2010)

indicating that an additional quality senior educational facility improves the value of a property by 6.9%. A study in Beijing by Hu and Wang (2016) demonstrates that the per square meter value of a residential property within close-proximity to a quality primary school catchment area has a higher price than that of properties located outside that catchment area, pricing to approximately 2266 CNY (Chinese Yuan Renminbi) (6.8% of the average selling price). Another empirical study shows that homebuyers in Hong Kong are willing to pay just about 27%-39% more to be situated within a key school district (Jayantha, et al. 2015). Considering the above, Hu and Wang (2016) argue that most households surveyed in Beijing choose to relocate to different locations based on a myriad of neighbourhood factors and not just the presence of quality schools in urban nodes.

The most relatable framework associated with the above empirical observations builds on Tiebout's (1956) research paper which presents what became to be labelled as the 'Tiebout model'. It set the groundwork for years of empirical studies on housing values and the parent's preferences for Local Public Goods (LPGs) (Wulsin, 2009). According to the model, people are more prone to 'vote with their feet' when making residential and locational decisions in specific neighbourhoods with characteristics such as quality schools (Yadavalli and Florax, 2014). However, neighbourhoods with the most valued services such as good quality schools, ceteris paribus, become more expensive to inhabit. Historical literature has widely adopted the hedonic price model as an ideal method of measurement in quantifying the value-added by good school qualities to the price of housing (Tiebout's, 1956; Oates, 1969; Black, 1999; Fen, et al., 2010; Nguyen-Hoang, 2011; Khachiyan's, 2013; Yadavalli, et al., 2014; Wild, 2016; Jayantha, et al., 2015; and Hu and Wang, 2016). In this model, a buyer's demand of a property is considered an aggregation of their demand for every single attribute of their housing. A few of these studies indicate that using a direct comparison of two homes in a similar location sharing equal local amenities - if proximity to a quality school is a valued attribute then the value of the property in the better school catchment area will reflect that value (Wild, 2016). Therefore, the determinants that cause variations in property values may be explained by the shifts in relationship between demand and supply or variations in the property attributes traded in the local areas under investigation.

1.3. The Significance of School Quality Purposing the Study

The seminal studies conducted by Oates (1969), Rosen (1974) and Black (1999) explain that consumers maximise their utility through their decision to reside in a neighbourhood that provides the greatest surplus in terms of benefits over cost. As shown by Khachiyan (2013), for many people, the quality of a good school is an imperative consideration when buying a

house. However, the numerous attempts in quantifying the link between property sale prices and good school characteristics have been obscured and a mammoth task to uncover. The failure to control for certain variables and their variations can result in biased estimations, which are often used to represent the extent school quality capitalisations influence property sale prices. Further criticisms stress that existing literature also fails to prove that the capitalisation effects of the school variables do not apply equally for both low-end and high-end properties, even when they are situated within the same urban context. It can be argued that observable variables tend to have a higher correlation with a myriad of neighbourhood characteristics that homebuyers prudently consider, in contrast to unobservable school variables.

It is still debatable as to which specific variables count as appropriate measures of school quality and how they contribute to property price variations. However, several studies virtually indicate a consensus that there is a positive correlation between certain school quality attributes and property prices (Kane, et al. 2005). As a backdrop, further research is ad-hoc to determine the specific measures that should epitomize school-quality in the property transaction market. The research provides valid econometric inferences using yearly school pass rates, conditions of educational facilities and sporting profiles (based on performance rates) as key school quality determinants households with children prioritize when making housing-decisions. As an effort to appropriate a proxy for this report, the focus on section 21 schools provides an idealistic basis as this school type has not been investigated by scholars. The research conceptualised a proxy best-fit for the subject under investigation that was guided by the scholarly theoretical standpoints that provide significant meaning to how section 21 schools function.

1.3.1. Section 21 Schools and their Allocated Functions - Former Model C Schools

The South African Schools Act (SASA) of 1996 made provisions for School Governing Bodies (SGBs) in public schools to manage their recurrent expenditures to ensure they become more responsible and progressive. This led to the inauguration and constitutionally legislating section 21 as an Act adopted from the inherent functions of section 20 schools. Adopting this function requires additional capacity and responsibility from SGBs. Nyambi and Heystek (2007: 39) note the following:

a) Managerial capacity to account and handle public funds;

b) The aptitude to ensure sound financial decision-making; and

c) The ability to meet continual contract onuses to suppliers of services.

Among others, these prerequisites are governed by SASA, section (16) (1) which states as follows:

> *"The governance of every public school is vested in its governing body and the principal with the academic staff is responsible for the professional management of the school" (RSA, 1996: 14).*

Subject to the SASA Act, allocation of section 21 functions provides schools with such an opportunity of gaining a somewhat level of capital autonomy. The awarding of the section 21 status or functions to schools in South Africa was considered an imperative stage in the process of democratising and improving schools. Post-1994, the first schools that applied and obtained the section 21 functions and status were predominantly the former model C schools (see also RSA 1996: section 21(6)). This occurred soon after the promulgation of the SASA Act as most former model C schools were deemed capacitated in school management and governance before 1994 (Heystek and Nyambi, 2007: 39). Pre-1994 a large portion of these schools were well-equipped with infrastructure and basic resources that were financed from tuition fees paid by parents who in general had better salaries than parents in other parts of South Africa. The enactment of this rule was an opportunity for the government to save their state funds from financing school expenses.

Model C schools under the section 21(1) rule, as guided by the SASA section (16) (1), were also awarded powers to appoint more educators than the standard allocation which is also financed by the school fees (Heystek and Nyambi, 2007: 39). The power and status related to section 21 functions correspond with the human skills and financial resources available in these model C schools (Harris, et al., 2003: 22). However, these functions are not realised in rural schools and schools in lower-income communities with insufficient funds to exercise these powers and status. Karlsson (2002: pp. 327-336) underscores the financial burden associated with this form of power or advantage that is limited to schools or parents with capital means to pay for certain benefits. Schools without adequate funding to maintain or upgrade the educational facilities and grounds are not awarded these benefits. Therefore, capital independence is a prerequisite to ensure a maintained provision of quality education, resources and facilities.

The extracted Section 21 functions (Nyambi and Heystek, 2007: 229) for the purpose of the research include the following:

a) School building upgrades, maintenance of the school grounds and the overall landed property, which includes but is not limited to, school boarding houses;

b) Defining the school's extramural curriculum and the option of subject choices in terms of provincial curriculum policy;
c) Procuring educational materials and equipment's; and
d) Paying for municipal services, such as electricity, toilets and potable water to the school.

Relevant to this research paper, a study in the Western Cape government department stresses that it is mandatory for Section 21 schools to ensure the day-to-day maintenance of educational facilities and resources (Schreuder, 2016). In addition, regular maintenance should be prioritised daily to prevent school buildings and facilities from deteriorating or undergoing obsolescence as an outcome of negligence. Therefore, the research signifies submerging into understanding the importance of maintenance, improvements, and conditions of educational facilities to determine the effects of school quality. The complexities of how quintile 4 and 5 former model C schools choose to finance quality upgrades on the school infrastructure to achieve aesthetic competency are critical to this paper. This will extend knowledge in the public finance literature that has a long tradition of attempting to determine the monetary value added by school quality into real estate (i.e. property sale prices). Hence, this has resulted in several studies overlooking several neighbourhood and school quality attributes that can contribute a significant variation in property values.

The key challenge has been differentiating between the effects of other neighbourhood factors and the influence of school quality (Wulsin, 2009). This is observed from several housing valuations undertaken that do not prioritise value input factors (both financially beneficial and adverse) to residential selling prices. To overcome such financial burdens of properties being under-valued during valuations for prospective homebuyers and/or home-sellers, this paper sheds light on the effects of public services on property values. The empirical study will be realised through investigating properties within the school catchment area (based on a 4-7km radius using a fixed-boundary approach).

1.3.2. A Structured Proxy: Defining School Quintiles and their Quality

Subject to the SASA of 1996, the initial objective was to provide an educational structure that is constant for governance, funding and organisation (Dass, et al., 2017) of educational facilities and resources. Public education is subsidised by the government through a pro-poor funding model. This denotes that former black schools obtain more government funding than previously white schools (known as, former model C schools). The government funding model categorically creates five different schools called quintiles. The objective is to determine the funding amount each school receives. The schools that do not charge school expenditures

are declared no-fee schools in the lower quintiles 1, 2 and 3 (see figure 1). Quintile 1 schools receive more capital per pupil for non-personnel and non-capital expenditure than learners in higher quintile categories (Strassburg, et al., 2010). Conversely, schools in quintiles 4 and 5 obtain minimal financial support and therefore can charge school fees.

Purposively, for this report to untangle the impact of quality schools and appropriately measure the capitalisation value added by educational facilities on housing selling prices, the following proxy will be based on higher quintile schools. As asserted above, quintile 1 to 3 schools have substandard educational resources and facilities. The basis of their quality in terms of the school's education, facilities and resources reflect a somewhat non-conformity to section 21 functions. Therefore, it is not viable to estimate the value-effect of low-quality schools on housing values situated in a similar location. Comparatively, adopting quintile 4 and 5 schools will contribute a higher significance in measuring the effects of both quality school and other unobserved neighbourhood characteristics on property selling prices. The prevalent selection category of quality schools to act as an efficient and effective proxy for this study is based on schools (particularly SGB's) that are registered and adhere to the section 21 functions, and responsibilities of the SASA 1996 Act.

Figure 1: A Quintile Model for Education Funding - The State's Approach

Source: Dass, et al. (2017)

As a framework proxy adopted from the section 21 Act, the quality of Jeppe Boys High school is grounded on the terms that the school is registered for section 21 functions. Jeppe Boys High is a fee-paying school, which is where SGB's real powers come from, by the school

having sufficient money to carry out their section 21 obligations (Acta Academica, 2007:39). The school independently raises funds to take advantage of these awarded powers and also decision-makings on how the capital should be distributed. There is no funding assistance provided by the government to Jeppe Boys High to allocate for management purposes and facilities maintenance. This includes maintaining basic services such as electricity, drainage and water supply. The Department of Basic Education also contributes minimally towards any form of improvement or development opportunities for the school.

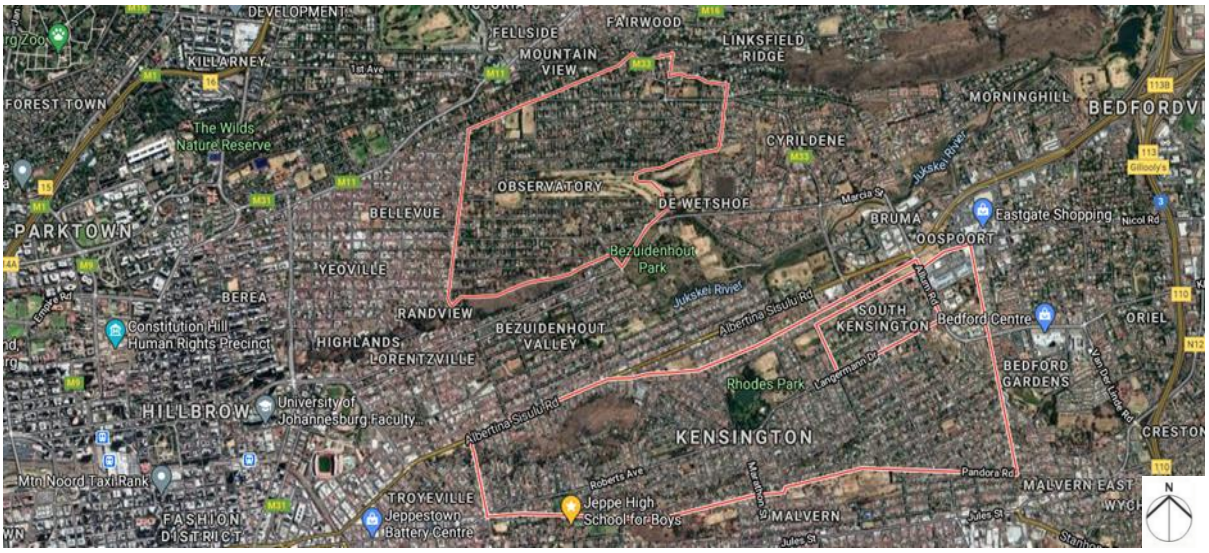
The school is required to use funds from fees to perform day-to-day operations and maintenance of school buildings, grounds and facilities. One of the areas for SGB's in terms of financial management is to improve and maintain educational facilities and school buildings. In regard to section 21 schools, there are limited funds to execute these prerequisites. Jeppe Boys High is a fee-paying school that needs more capital support to enable sustainable fruition of facility improvements. Emphasis on the maintenance of the school premises, timely-repairs and reforming the bureaucracies are of paramount importance (Freire, 2000:24) for the functionality and management of section 21 schools within region F (CoJ). Providing support to this, school environments are major contributing determinants on the rollout of learners optimally learning and developing. Educational facilities that are in dire conditions and dilapidated state (property-related, such as arson) impend the well-being and state-of-mind of students, further posing health risks to the staff and students. As such, providing quality education and resources (for pupils and staff) holds paramount significance as highlighted in this paper.

1.4. Delineating the Study Areas – Dynamics behind Property Markets, Locational and Environmental Amenities

Map 1 delineates the local geographical areas, namely, Kensington and Observatory situated within the jurisdiction of CoJ used as sampled case study areas for the empirical findings in chapter 4. The CoJ is subdivided into 109 wards of the total 420 wards in the Gauteng region (De Wet, et al., 2008). The case study areas are situated in ward 66 (and a part of Kensington South also falls under ward 61) of region F. Observatory covers an area of 66ha and Kensington covers an area of 5.21km^2. This region is saturated with a pool of lower-and-medium income residential spaces coupled with a few older and heavier industrial nodes, with a limited number of new developments in these areas (CoJ, 2018). Both sampled areas are well-located with ease of access to the N1, M1 and M2 freeway system. The land use composition of this region varies from manufacturing, industrial areas in the northeast and residential in the northwest, the south and west. As Nobel et al. (2001) points out, efforts in

the area of mitigating urban decline has led to the creation of other developmental projects and social housing. This includes the infrastructure investments anchored around the train station in Jeppestown as it is situated within the JHSB school district. This was an exodus of business investment to decentralized nodes four-to-five decades ago leading in the CBD degrading with the spread of urban blight into some areas in close-proximity (CoJ, 2018).

Map 1: Aerial illustration of the sampled case study areas in region F's housing market, CoJ



Source: Google maps (2021)

Despite the underlying dynamic shifts behind property market variations in CoJ, it is evident that the spatial layout structure in this city's development space has resulted in areas that are fragmented and segregated spatially, socially and economically. These consequential outcomes are highly attributed to the heirlooms of apartheid planning which also contributed to the reconfiguring of urban landscapes resulting in urban decay (HSRC, 2014). This is also reflected by the Gauteng city-region embodying a trichotomy of isolated and varied spatial areas consisting of cities, towns and urban nodes (Thantsha, 2016). This cluster within the city region covers about 2% of the South African land mass and it is associated with pervasive development despairs.

Disjointed land-uses constrict opportunities of realising linear sustainable growth patterns that strengthen the development direction envisioned for the Gauteng city-region. However, for over a decade a plethora of strategic spatial interventions have proliferated the CoJ geography as an effort of both public and private sectors (Lightstone, 2020). This was intended to significantly alter these traditional structures through developing integrated spatial structures. The stimulus created by these regenerative strategies has provided valuers with an

opportunity to quantify the market values associated with various locational and environmental amenities and dis-amenities. Incontestably, Yadavalli et al. (2014), Hummel (2017) and Wen et al. (2019) argue that locations present both costs and benefits that are incurred as an effect of being in close-proximity to neighbourhood attributes and a myriad of other local determining factors.

Notably, in this transaction market these annotated property determinants have exponential value-laden effects that remain unaccounted for, particularly on their general influence on property valuations and sale price variations. There is a trade-off that homebuyers face amid the costs or local taxes they have to pay for public goods and the quality of local services provided (Wulsin, 2009; and Black, et al., 2011). For example, Yadavalli and Florax (2014) state that this can be related to the bid-rent models in which land prices are bid up based on the willingness of people to pay for proximity to city-centres compared to households that prefer suburban spaces. Therefore, with region F many attractions can explain such outcomes because there has been constant strategic planning, developments and improvements in the quality of local public goods and neighbourhoods. This gives reference to the inherent place-based improvements by the government and policies to ensure radical economic growth, integration and diversity in local areas.

1.5. Problem Statement

Scholarly debates related to the notable variations in property selling prices caused by the school quality parameters when capitalised into hedonic pricing frameworks have been distinguished as highly polarised. In the international body of literature, some empirical studies (also see Nguyen-Hoang, 2011; Khachiyan, 2013; Wild, 2016; Hummel, 2017) attempting to reveal the true value of school quality to explain property price trends overtime produce estimates that are plagued by a myriad of hedonic issues or produce average coefficients. To date, the uncertainties associated with some of these hedonic findings is due to the cumbersome nature of the datasets or school inputs and outputs used in the multi-variate regression analysis. Furthermore, the lack of a standardized proxy to measure the added value of school quality has resulted in scholars, such as Ka-wai (2010) and Wild (2014), constructing bias and average hedonic coefficients. The most identified contributor is the failure to incorporate the full-range of relevant school quality attributes in their hedonic regression models, particularly, when conducting property valuations.

With regards to research conducted on this specific topic in South Africa, there are no credible studies depicting the relationship between school quality and property sale prices. However,

there are notable literatures conducted on hedonic methods that focus on an array of issues experienced in the African emerging markets, and demonstrate the responses of prices to market indicators (see section 2.8.2). This provides the report with an opportunity to quantify the impact of the school quality variable and examine the variations in property prices in the short-run and long-run - contributing to this topic and expanding knowledge of key differences in price changes. The findings of such a study would benefit participants in the real estate market of making informed property decisions based on reliable empirical evidences. An in-depth analysis investigating the impacts of the quality of model C schools using a well-defined hedonic price framework in the CoJ property market is yet to be seen.

What is also evident on this topic is that in the construction of property price indices there is a noticeable consensus of model failures in minimising the endogeneity and spatial dependency problems (Chiodo et al. 2010; Wilhelmsson, Ismail, and Warsame, 2021). Wilhelmsson (2019) further noted that such problems are usually caused by reversed causality or hedonic function selection bias. These issues can also be cured by using the Treatment Effect model (see Wilhelmsson, 2021). With regards to the reviewed hedonic papers, the school quality variable has been exponentially dismantled with a voluminous amount of dummy variables employed to act as proxy representatives of the actual determinants that have a significant effect on housing prices. A close examination of the empirical evidences ascertains that, there is no conclusive stance taken on the appropriate functional form to employ, that adequately predicts the relationship between property sales and the school quality variables. Therefore, this shows that emerging markets cannot simply infer the research findings from developed countries to make sound decisions on the conditions of their local property markets.

The thesis employs the traditional hedonic pricing framework to explore the Kensington and Observatory property market. There are several practical contributing reasons that influenced the decision behind the selection of the study areas. In regards to region F (CoJ), there is a plethora of attractions and other fundamentals that can explain property price trends because there has been constant strategic planning's, developments and improvements in the quality of local public goods and neighbourhood's post-2010. This makes inference to the ongoing place-based improvements by the government as broadly inscribed in development policies to ensure radical economic growth, integration and diversity in some of the local areas. To uncover the values of these factors, the research suggests that policy be directed to availing a database of macroeconomic and microeconomic determinants that are fundamental in explaining property price variations in the South African property market.

The study aims to capture and provide comprehensive and up-to-date nascent evidences on price statistics of properties within the Region F's real estate market. The construction of such

local indices will enable the locals to grasp an overall picture of price development trends in the regions property selling market. Since quality schools are mainly situated in affluent urban environments, discarding the control of selected predictor variables has resulted in an overestimation of the school quality effect on housing prices (Wild, 2016). Hence, the econometric analysis in the study will aim to measure the hedonic equation by employing multi-level analyses to construct the empirically suitable form of the hedonic function. The research argues for the use of robust multivariate regression analysis techniques, with the potential of addressing the sources of hedonic problems such as, but not limited to, omitted variable bias, multicollinearity, and heteroscedasticity identified in most of the previous empirical studies investigating property values. The research further tackles two other problems in hedonic modeling that are empirically ignored in previous studies examining school quality. The study explores the following:

- The first notable problem is *spatial autoregression*. Housing prices are dependent on each other and residing in close-proximity to high-quality homes can, solely, increase the housing price of a homeowner (Wen, et al. 2019). Considering that if this effect occurs and spatial lagging of the property sale price is disregarded, the estimated parameters will presumably be biased.

- The second issue is regarding *spatial correlation* identified in the error terms in hedonic models, and if overlooked could create larger standard errors (and subsequently smaller t-statistics) than anticipated (Wen, et al. 2019).

It can be observed that research focussed on exploring and analysing a display in behavioural changes from school quality variables across the price distribution has yielded a myriad of critiques. As meticulously highlighted in chapter 2, the methods used in estimating school quality are criticised for being distinguished as weak in establishing long-term housing market price movements in time series. Hence, one of the paper's contributions is to bridge the gap identified in previous literature focusing on this topic by augmenting that knowledge in the South African context. The study seeks to investigate the measurable variation in property selling prices when a model C school quality variable is employed in the hedonic dataset. To achieve the study's purpose, the researcher illustrates the research methodology and the appropriated model approaches and regression techniques employed to measure this highly untapped relationship in African Urban Economic studies.

1.6. Significance of the Study

The major practical contribution is measuring the added-value of school quality, locational, structural composition, and neighbourhood quality attributes in property sale prices. Currently there are still empirical challenges evident in literature of the generated valuation reports of property transactions in the open market that rely on bias information. The role property markets play cannot be overemphasised (Owusu-Ansah, 2010) particularly in the socioeconomic development of many countries, as a result, providing reliable and robust property price indices remains fundamental. Therefore, hedging the risk of unnecessarily capitalizing insignificant factors as identified in literature that has resulted in the under-or-over estimation of property market transactions still needs prioritisation. The research extends on this knowledge and contributes to better understanding the effects of school quality and the true attributable value representative of this neighbourhood attribute on property sale prices. In so doing, constructing a standardized proxy for school quality using relevant attributes of model C schools was crucial for the research. Moreover, addressing issues in this subject area of endogeneity, standard omitted variable bias, heteroscedasticity, multicollinearity, spatial autoregression and spatial correlation would provide academics, policymakers and market participants with reliable information that can be used for property market analysis. The research further examines household's preferences of local public goods, demand and supply, and the willingness to pay for the composition of the attributes offered in local areas that could provide valuable spin-offs for market participants, particularly for the analysis in chapter 4.

1.7. Primary Research Question

The research question is purposed to address the lack of understanding of how and the extent to which the measure school quality effects property sale prices, with a consideration of other unobserved neighbourhood characteristics.

"What are the effects of school quality and other unobserved neighbourhood characteristics on property sale prices?"

1.7.1. Secondary Research Questions

i. What is the marginal propensity of households to pay a higher housing premium to gain access to a better-quality school?

ii. What are the most appropriate factors of school quality that have more effect on property market values?

iii. What other unobserved neighbourhood characteristics can be capitalized in property sale prices?

1.8. Research Aim and Objectives

The aim of the research is to estimate the effects of school quality and other unobserved neighbourhood attributes on property sale prices. The realisation of the research aims requires undertaking the subsequent objectives:

i. To identify and provide a review on the potential areas where there is a paucity in research or knowledge on school quality and property price indices.

ii. To apply the hedonic dataset to the different models and define the appropriate functional forms by which property-transacted prices are accurately examined.

iii. To empirically estimate the effects of a model C school quality and other unobserved variables that cause a variation in property transacted sales.

iv. To apply the best-fit techniques based on the employed regression models to ensure that the hedonic problems (omitted variable bias, multicollinearity and heteroscedasticity) encountered in previous studies (in this subject area) are cured and under control.

1.9. Assumptions of the Study

This study makes two main assumptions:

- The quality of former model C schools has an effect that can be capitalised in property sale prices; and

- Research conducted on this topic is only generalizable in the South African context, particularly in region F of the City of Johannesburg.

1.10. Research Report Structure

The structure of the research report contains five chapter, this includes:

- Chapter 1 outlined the introductory sections where the background, research problem, aims and objectives of the study are discussed.

- Chapter 2 provides a review of the key elements regarding property sale prices in previous literature. The main property attributes are presented that has been included in prior studies, as well as discussion on the conceptual framework that is linked to the empirical findings. This chapter further provides a synopsis of the employed hedonic price model and outlines the processes followed of hedonic regression modeling.

- Chapter 3 introduces the research design and the employed hedonic pricing techniques. It further discusses the sources of the data and the limitations of raw dataset in this study.

- Chapter 4 analyses and discusses the empirical findings. To support the analysis this chapter draws a comparison with the literature findings and highlights the value of the remedial actions applied to aid the hedonic problems encountered in previous studies.

- Chapter 5 summarises and concludes the study. It provides conclusions, possible discrepancies with previous literature, policy implications, limitations, contributions of research knowledge and the potential areas for future research.

2. Chapter Two: Literature Review

2.1. Introduction

This chapter identifies and discusses the key concepts relevant to the research to purposively link the theoretical component to the empirical part of the research. It provides an outline of the main debates, arguments, discussions, ideas, and theories (where applicable) to the construction of property sale prices, locational and environmental determinants. This assisted in providing a sense of direction for the study and to identify the existing gap in knowledge that this work aptly addressed.

2.2. Determinants of Property Values

This section of the literature attempts to understand the concepts through reviewing the primary determinants (location and structural attributes) that contribute to changes in property values and the discourses regarding the topic area. It then assumes a position of the identified environmental components that are covered in the scope of addressing the issues in the research (see section 1.5). This includes a focus on school quality. At the end, it will be followed by a brief discussion on the gaps identified in the body of literature knowledge the research filled.

2.2.1. Location

The value of a property depends on a myriad of variables, which are associated to both the physical characteristics of the property and its geographical location. The centrality of your housing location has a financial bearing on the value of that property. As Fernández-Durán et al. (2011) notes, each specific location represents varying values in attributes that can include distance, crime, traffic noise, accessibility, socioeconomic conditions and proximity to local amenities such as good quality schools. In addendum, this implies that location has an influence on the decision of choice and asking price of each house (Fernández-Durán et al. 2011). Residential properties situated within the jurisdiction of a large metropolitan city tend to attract higher pricings during appraisals as opposed to homes located in the suburbs (Gibbons and Machin, 2007). The backside issue is the land size and their highly associated costs in the city area reasoned by the benefit of gaining access to ample amenities in close-proximity as opposed to commuting long distances to gain access.

Location also includes employment opportunities and other recreational and leisure benefits (Gibbons and Machin, 2007). Neighbourhood attributes such as schools and other community

services or amenities are contributing determinants towards the value of a property. Environmental factors also have wide implications that can either raise (such as good quality schools, health centres, public transport) or reduce (such as property-related crime) a property's price. Therefore, location incorporates a myriad of amenities/attributes that provides a residential property with a marginal cost that can be capitalised into the property's value.

2.2.2. Structural Attributes

The attributes of a property towards its implicit value is widely determined by various determinants. In terms of this section, the study provides that structural attributes are also significant, for example, square footage is seldom used to determine a home's appraised value (Hu and Wang, 2016). This is based on the expectation that it will have a substantial effect on the property selling price (Wild, 2016). Bearing out the arguments in previous studies, it can be revealed that homebuyers in different price assortment value this variable (Ziets, 2008). In his study, this is identified by the significant variation amid the coefficients at the highest and lowest quantiles whereby the marginal value of a square foot for highly priced houses is two and a half times the marginal price per square foot for the lowest priced houses (Ziets, 2008). Virtually, traditional methods such as the Ordinary Least Square (OLS) model or other methodologies that account spatial autocorrelation often are found to overstate the price of an additional square foot for lower priced houses but devalue the effect on higher priced houses. The quantile results give significant insights to the various association that the explanatory variables have with a house's selling price, including variables, such as, bathrooms, lot size, square footage, and floor type have a significant effect on the increases of selling prices (Ziets, 2008). Moreover, other variables that constantly have a relative effect on house selling price across various value assortments – which can include distance to city centre (CBD), exterior siding, and garage (Fen, et al., 2010).

2.3. Environmental Determinants of Property Values

For decades, economists have understood the value-adding nature of neighbourhood characteristics in which that specific property is located as well as its structural attributes (Clark and Herrin, 2000). These variables are considered to be, amongst others, important determinants of housing value. Research in this subject matter has elicited strong interest and public debate. Arguably, school attributes have remained key input variables in measuring property value coupled with location and structural attributes. Other neighbourhood determinants or issues such as distance, crime rate, air quality, proximity to sub-centres and employment centres, and environmental hazards. The CPEP (California Public Education

Partnership) published an opined driven poll the organisation conducted indicating that most residents in California value improvements in public education more than mitigating social costs (crime) and environmental costs (transport noise) (Clark and Herrin, 2000). Analytically, this survey also suggests that those local residents are more willing and prepared to pay more taxes in prioritising school quality improvements above other environmental quality issues. Considering the derived results from various research authors (Tita and Greenbaum, 2006; Gibbons and Machin, 2007; Gibbons, 2011; Wild, 2016; Fleishman, et al., 2017; and Wen, et al., 2019), it is somewhat a mammoth task to contrast or generalize their results. This draws from the argument that results tend to be more inclined to context-specificity and proximity of the property from a specific environmental determinant of property value. The section below further highlights the various debates and deliberations in research as a supplementation of the property value determinants provided above.

2.3.1. The Effects of School Attributes on Property Values

As put forward by Aliya et al. (2016), housing prices should adjust accordingly to the accounted school quality differentials, when assuming that the prospective homebuyers consider variations of school quality and are mobile within neighbourhoods of various district attendance zones. The quality of school has been long capitalized into property values in similar prior studies (Wen, et al., 2019), to which home values increase. Ceteris paribus, when a good school quality characteristic changes (Clark and Herrin, 2000). Moreover, controlling for various structural attributes provides in literature (Downes, et al., 1997; Leech, 2003; Davidoff, et al., 2007; Clapp, et al., 2008; and Wild, 2016) that the school attendance zone in which the house is situated is also an imperative determinant of housing sale prices. Surveys and studies conducted by the above-mentioned authors suggest that most households choose to relocate to different locations based on a myriad of neighbourhood factors.

As one of the essential determinants that causes a marginal variation in housing selling prices, quality educational facilities, has elicited scholarly enthusiasm (Wen, et al., 2019). Firstly, Oates (1969) provided empirical evidence indicating that public goods, such as schools, are capitalised into housing selling prices of which homebuyers are willing to pay high premiums to gain access to good quality schools. The concept Oates provided became widely used by numerous scholars. The housing data was predominantly used as a measure to determine the value of schools. For instance, Black (1999) confirmed that the valuation of school reform needs a quantitative estimation representing the parents' view of school value and the selling price of a property can be utilised to infer this value. Black's study, similar to Oates (1969)

indicated that parents' willingness to pay an added housing price of 2.5% for an increase of 5% in their children's test scores.

Yang et al. (2018) further substantiated by verifying that certain parents particularly those with quality education backgrounds, are more willing to use extra capital to purchase a residential property adjacent to a quality educational facility. Other researchers such as Machin (2014) reviewed numerous literature studies and their estimates. This was to provide a summary denoting that there is a significant positive correlation between housing price and school quality. The quality improvements of educational facilities will result in varying degrees of housing selling price (Bae, et al., 2013; La, 2015; and Zheng, et al., 2016). Unlike majority of the research studies that attempted to measure the average rate of school resources, Chiodo et al. (2010) verified that the correlation between housing prices and the quality of public primary and secondary schools is non-linear. Subject to the research's problem statement, having a linear assumption of test scores can potentially overestimate the premium offered by general schools or underestimate that offered by high-quality schools. The correlation between school's performance and residential prices feeds back into admissions of educational facilities ('also known as the selection by mortgage') of a specified category of children into good schools (Fleishman, et al., 2017). This considers that according to the school's catchment areas, access is offered to children who reside within that school attendance zone which can be viewed to also reinforce school segregation. To a certain extent, good schools drive the value of residential property markets.

2.4. A Review of Previous Literature Studies Relevant to the Research

Arguably, an investigation of the correlation (or lack thereof) existing between housing prices and high-quality schools, requires a characteristic variable to be set initially in order to adequately represent the quality of schools. Wen, et al. (2019) stresses that scholars not accurately expressing educational quality has become a constant concern. Notably, this stems from two common methods used. The initial method expresses school quality with output variables (dropout rate, test scores, and so forth) input variables (student-teacher ratio, per pupil spending, teachers' salaries, and so forth). In western countries, this method is widely used for being considered simple and straightforward in accurately estimating the impact of educational facilities. For example, Brasington et al. (2003) validated that a significant correlation occurs between housing prices and school expenditure in the United Kingdom using a data set of 77,000 housing transactions. Similar summations of results were drawn from other articles (Hilber, et al., 2002; and Dhar, et al., 2009).

More scholarly reports validated that both output and input measures can significantly add value to housing prices, of which the implicit price of input variables is higher in contrast to output variables. By contrast, literature on educational capitalization that was reviewed by Nguyen-Hoang et al. (2011) proved that most studies confirmed the capitalisation effect of school quality, particularly the output measures. In most empirical studies, test scores were considerably used as the common representative output measures. For example, Rosenthal (2003) asserted that a rise in test scores results in an increase in housing selling prices with an estimated price elasticity of approximately 0.05. In addendum, John et al. (2010) confirmed that a 10% rise in the test scores (GCSE) will increase housing prices by 1-3%. There are other scholars that provided similar conclusions (see Davidoff, et al., 2007; and Clapp, et al., 2008). Considering the abovementioned, Downes et al. (1997) arguably affirms that homebuyers value both input and output measures as equally important.

The second method of measuring quality of educational resources is widely adopted by Chinese scholars that uses subjective scoring or school rankings. This method instinctively differentiates the school's quality to contrast their varying effects on housing prices (Wen, et al., 2019). An example of this was corroborated in Shi and Wang (2014) article, providing that key public primary and secondary schools have a significant effect on the selling value of residential properties. Considerably, a drop-in ranking of a primary or secondary school will lead to a 0.90% decline in housing values. This similar method was also used in Wen et al. (2014) research paper. Numerous existing literatures affirms that the quality of both school types project a higher value-added effect on housing prices. Amongst others, this can be reasoned by the 'nearby enrolment policy' that was enactment to ensure that the enrolment undertaken in both school types is aligned with the school district where the pupil's household is situated.

Wen et al. (2019) for instance provides that as a result, housing prices reflect a significant educational district effect. As an example, Leech (2003) substantiated and proved this argument by adding that in the United Kingdom (U.K Coventry) the housing price can be improved by a good school district approximately by 16%-20%. Further studies in this area concluded that compared with location, distance, property-related crime, shopping convenience, floor, or building age, homebuyers are willing to pay an additional price of about 27%-39% for a high-quality school district (Jayantha, et al., 2015). Analytically it indicated that high quality schools have become an essential indicator of measuring a household's willingness to pay extra for high grade school in a competitive school district. Notably, the presence of school district housing may possibly widen the income gap further between the poor and the rich. Housing assortments based on affordability and educational accessibility

within a certain distance can widen the disparity of providing equal education between rural, urban and metropolitan schools.

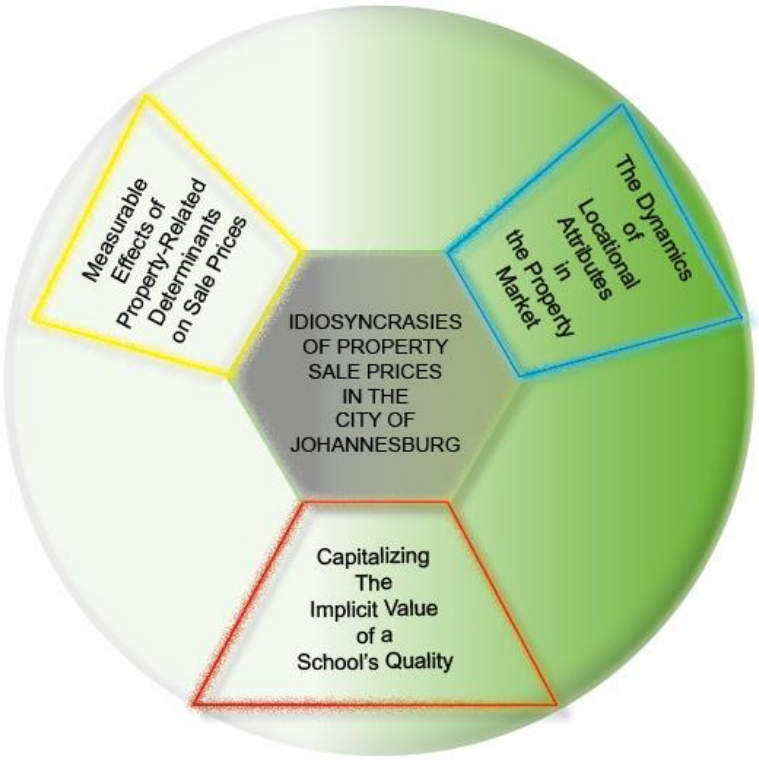
2.5. A Framework on the Development of Key Theories and Concepts

This section of the research mandates clarification of the key theories and concepts driving the research. This is in order to unveil where a knowledge-gap exists to occupy with the findings and outcomes of the collected data. Wilhelmsson (2000) identifies the fundamentals that influence demand in properties and the research adopts three of these factors, namely, locational amenities, structural attributes and environmental attributes. The primary conceptions provided below are assessed and authenticated further in the subsequent sub-sections. Purposely, this section is to assist in answering all the questions (both the main and sub-questions) outlined in Chapter 1.

Vaughan (2008) and Maxwell (2013) support this, as it is an imperative part of the research design. Both authors affirm this by asserting that it also assists in framing a conceptual framework that informs the research based on preceding debates, assumptions, theories and popular beliefs. As asserted by Vaughan (2008) a conceptual framework can be in either a narrative or graphic format. Therefore, the research representation of the conceptual model employs both styles of formatting. This section chronologically unpacks three residential value determinants (dependent variables) and property price (independent variable), which comprises of the following ideas:

- The Dynamics of Locational Attributes in the Property Market
- Capitalizing The Implicit Value of a School's Quality
- Measurable Effects of Property-Related Determinants on Sale Prices

Figure 2: Conceptual Framework Model



2.5.1. The Dynamics of Locational Attributes in the Property Market

Most authors are familiar with the axiom that values of real estate "are determined by three basic characteristics – location, location, location" (Hayes and Taylor, 1996: 7, pp. 2-9). The general presumption by economists, researchers, developers and planners is that locational access and proximity hold prominent significance as attributes, which contribute highly to property price changes (Seth, 2017). In this study, location presents both cost and benefits that are incurred as an effect of being in close-proximity to neighborhood amenities and dis-amenities. This component forming part of the property sale price is widely credited for the impetus and imperative role it has in South African developmental frameworks and policies. Purposively, policymakers, economists and authors employ this attribute to address the decadal issues highlighted in most literature studies provided in this chapter. Biased regressions, autoregressions and the inability of monetizing determinants such as location and other environmental attributes (e.g. transport noise, crime) were the culminated outcomes in most findings. Moreover, this became a mammoth task to uncover these determinants 'real capitalization value' due to the use of a wide varied data and samples of property attributes and neighborhood characteristics. Fortunately, locational attributes elucidate much of the variation in property prices (Hayes and Taylor, 1996). For example, locations in close-

proximity to recreational areas (i.e. city parks), good quality schools, and transport hubs command a higher premium in contrast to locations near city dumps, which accrue a discount. Similarly, locations with high levels of property-related crimes and higher property taxes tend to have lower real estate values. Housing prices also fall in communities where the commuting time and distance to the CBD (Central Business District) increases and rise as the quantity of air pollution declines. In support of Tiebout's (1956) and Oates (1969) assertion that individuals or households make property residential decisions based on the inter-or-intra-jurisdictional variations in property taxes and provisions of public services.

2.5.2. Capitalizing The Implicit Value of a School's Quality

Examining the extent to which educational quality has influence on property sale prices as a locational determinant has somewhat been flawed. This is evidently indicated by most homebuyer's preferential differences in specific school attributes they have a higher propensity and willingness of paying a higher premium for when making housing decisions. According to Bayer et al. (2007), not all educational attributes appear to be determinants of school quality. They further accentuate that few households deem certain school characteristics such as changes in student body characteristics or school expenditure as a property indicator. Gibbons (2011) and Gibbons et al. (2012) provide evidence alluring to this literature debate indicating that the same school variable homebuyers pay a premium is similar to the variable economist's associate educational quality with. This is namely labelled as the marginal effect of the school on pupil's overall performance.

From this research, a number of policy implications can be drawn relative to these flawed arguments and discourses. Analytically, policies that influence school effects can have a substantial impact on residential property prices. Undoubtedly, it presents a questionable view on policy initiatives or analyses that equates educational quality with school expenditures. For example, the study by Bayer et al. (2007) conducted in Dallas revealed that as far as home-buyers are concerned in that focus area researchers have been on target in identifying policy reforms directed at reducing the marginal effectiveness of schools.

According to several authors on educational quality (Hayes and Taylor, 1996; Weimer and Wolkoff, 2001; Reback, 2005; and Clapp, et al. 2008), it is suggested that either – high property sale prices may lead to higher quality schools (particularly public schools), or – highly ranked schools may presumably influence property sale prices to increase. The first argument presumes that households with higher income are more inclined to seek election into school representative boards or seek to increase educational quality through input on school expenditure and employment decisions. Assuming this is factual, then higher than average

property prices could increase educational quality. The second argument denotes that potential residents may be enticed by the education quality of a highly ranked school with a consequent bidding up of property sale prices in this school service boundary. Many new households that arrive in such school district zones actively seek information from local perceptions and online on school quality, report cards, test scores, demographic composition of the school, composite or summarized statistics for the school to make school and housing decisions.

2.5.3. Measurable Effects of Property-Related Determinants on Sale Prices

A real estate property may be viewed as a bundle of goods. Each place is somewhat attached or interconnected with a variety of value-laden structural attributes, neighborhood amenities and dis-amenities and other amenities, such as accessibility, environment, schooling, and the labor market. Real estate properties embody a collection of desirable characteristics such as location, shelter and comfort. Hayes and Taylor (1996) assert that, the premium homebuyers are willing to pay should simultaneously correspond to the prices they are willing to pay for its housing determinants or attributes. There are several factors and variables that help determine the value of a particular property. Some of these property determinants form part-and-parcel of the physical building structure (features, design, size and age) (Gibbons, et al. 2014). Others are attributed to neighborhood characteristics, such as, the quality of environment (air quality) and condition of surrounding housings, and the status quo of crime rates (Rosen, 1974). Some are directly attached to the land such as accessibility to certain important locations and lot sizes. By viewing a property as a sum of its attributed characteristics, a hedonic housing model generates valuations of an individual's willingness to pay for each attribute.

The concept of using land values to value place has a long background tale and grasping this insight reveals that the value attached to a plot of land unveils a homebuyer's propensity of the demand for that land in that particular environmental setting (Gibbons, et al. 2012). Just to add mud in the water, pure land values are rarely observed (Gibbons and Machin, 2007). Gibbons and Machin further argue that, instead the value associated with place needs to be detached from the overall cost of the physical property structure from the land it is built on and the view of nature in its peripheral surrounding. This type of underlying issue can be addressed using modern statistical approaches such as the multiple regression analysis method (Gibbons and Machin, 2007). This advances the disentangling of housing expenditures into various attributed components (contributing to the value of the residential property) coupled with the characteristics forming the property and its location (Gibbons, et al. 2014). However,

research in economics underpinning this have preferences of adopting the theoretical work based on the demand and supply of an optimum bundle of goods. This particularly draws from the instable economic groundwork cemented by Oates (1969), Rosen (1974) and Black (1999) arguably. Observantly Rosen method of approach outlines a market equilibrium whereby its 'consumer choice over bundle of goods' - such as housing equates to opting for an optimum composition of goods. For example, this entails transport accessibility, property size, location, and school quality.

The initial and commonly used understanding from this theory is that in equilibrium the marginal benefit of improving, for example, any component of a bigger residential property for a provided consumers income and preferences must be offset by the utility expenditure of the additional cost involved (Fernández-Durán, et al., 2011). According to Hu and Wang (2016), this presents a rationale for using the cost on a residential property to monetize the probable benefits associated with its observable characteristics. This can be realized by attempting to measure the amount of change in housing cost with marginal changes in a single observable attribute (ceteris paribus), and then this can presumably be interpreted as its implicit price or the marginal inclination to pay for that attribute. The hedonic price function has become that well-known measure to trace the locus of the correlation (or lack thereof) between property sale prices and the amount of its composite attributes (Wen, et al., 2019).

The second perception denotes that homebuyers in such property markets, in terms of their income and preferences are very heterogeneous as are home-sellers (Fernández-Durán, et al., 2011). This indicates some level of heterogeneity. Therefore, signifying that the value of any attribute associated with a property is not deemed unique even in single residential property markets. Although the value of educational quality may be very high in high-quality districts, homebuyers in that market area are probably more inclined to pay a higher premium for additional improvements in the academic performance of their children. Conversely, the price linked with an improvement in educational quality may presumably be lower in market areas where the quality of the school is sub-standard. For instance, this can be reasoned by several home-sellers and homebuyers who may not have children are most likely place little value on the educational quality in that area of the market (Epple, 1987). Hence, the hedonic price model is highly viewed as being non-linear. In other words, the implicit prices indicate a variation over the distribution of neighborhood and housing characteristics. The shape and slope of this correlation is anchored by the relative number of buyers (i.e. consumers), home-sellers and property developers (i.e. suppliers) in different areas of the market. This is reasoned by the fact that market forces determine the direction and pace of such dynamic shifts.

2.6. Positioning the Research in Previous Literature Studies

For the research, the pertinent schools of thought discussed in this chapter are essential to the analysis of the findings. The results were also considerate of the dynamics behind the role played by microeconomic and macroeconomic fundamentals and other factors that explain property price changes in the short-run and long-run (Owusu-Ansah, et al. 2021). Although this is beyond the scope of the research, there is a need for more empirical studies to uncover the true values of shadowed or under-utilised amenities, dis-amenities and dis-utilities that explain property price indices overtime. The general literature findings identified in previous studies on school quality and property sale prices cannot be convincingly used in drawing conclusions for other countries because of context-specificity and differences in property market behaviours.

For the research, appropriating a review of the hedonic method and approaches, selection of the appropriate functional form based on the secondary data used, and addressing the problems encountered in previous studies on hedonic modeling was beneficial. Hence, the research formulated a standardized proxy of relevant school attributes to represent school quality and avoid producing bias information and large standard errors and subsequently smaller t-statistics in the hedonic models. The research tackled the major hedonic issues that most studies on school quality and housing prices failed to resolve, namely, spatial autocorrelation, spatial correlation, omitted variable bias, heteroscedasticity and multicollinearity, which are discussed below (also see section 2.9). The research moves a step further to run diagnostic experiments and test the validity of the empirical findings in order to contribute accurate estimations and knowledge in the local markets on the effects of school quality and other property fundamentals against property sale prices. With the exception of SA, the effects of quality model C schools on the property market is less understood, particularly at the local regions. Hence the research aimed to provide insight on this gap concerning what most literature archives have misrepresented and the relative differentiation that occurs on the ground using Kensington and Observatory as case study areas. In summation, the research recommends that a database with more information on real estate markets be established to enable future research to produce reliable and valid estimations that can be useful to policymakers, investors, real estate developers, households and banks during property valuations.

2.7. A Comparison of Empirical Methods Based on Previous Studies

Commonly, measuring the capitalization effect of quality education or schools is done using four methods, or a combination of these methods used simultaneously. This includes the (1) hedonic price model, (2) instrumental variable method, (3) fixed-boundary method, and (4) difference-in-difference (DID) method.

The hedonic price model is the most widely used to measure the implicit price of heterogeneous goods (Fletcher, et al. 2004). The theoretical underpinnings of the hedonic method are derived from Lancaster's (1966) consumer theory and Rosen's (1974) equilibrium model. By keeping neighbourhood characteristics, location and structure the same, this model has been to determine the implicit value of educational variables. However, the omission of variables is one, amongst other, possible concerns associated with the hedonic price model. Hence, the research study is purposed to address the omission of school or neighbourhood variables that has led to argumentative discourses of scholars providing bias results. Clapp et al. (2008) contrasted the coefficients of the quality of school (using test scores as an input factor), and affirmed that there are possible overestimations of the effect of test scores on the price of housing when unobserved variables are omitted.

The fixed-boundary method is commonly used to prevent the biased over-or-under estimation caused by omitted variable bias and obtain robust results when investigating the effect of educational facilities on the housing price. This method intuitively sets a boundary to ensure all noted characteristics (excluding the educational traits) of residential properties on both paralleled sides of the boundary line are similar (Wen, et al., 2019). Weimer et al. (2001) conducted such a study, by using the school district boundary, the enrolment scope, and the jurisdiction to control the effect of unobservable variables such as public services, high school characteristics, and the student race. The obtained results confirmed a significant positive correlation between housing price and primary schools. A similar study by Fack and Grenet (2010) was conducted by setting a school district boundary to investigate the impact of public and private schools on residential property prices. Respectively, houses on the right and left sides of the boundary belong to a private and public school district. The study determined that housing prices are boosted by public schools significantly, whilst the public-school capitalization effect decreases as an outcome of a private school that is newly constructed within the same district locality. Moreover, Dhar et al. (2012) adopted the fixed-boundary model to control the impact of certain economic factors (such as real estate tax and school expenditure), housing market conditions, and structure characteristics.

Findings conclude that the omitted variables can result to biased measures being obtained and an overestimation of the capitalization effect of test scores. Therefore, this further confirms the conclusion provided in Clapp et al. (2008) research article. Certain researchers found that neighbourhood factors such as, but not limited to, school quality, are endogenous variables (Wen, et al., 2019). To acquire unbiased results and eliminate this effect, instrumental variables were used. Downes et al. (1997) inspected the change effect in test score on selling prices of residential properties in Chicago from 1987–1991. In the study, they used proportion renting, test scores, student expenditure, per pupil-assessed value, and the housing proportion of the tax-base. Moreover, every four years the public schools in Britain are required to undertake an Ofsted inspection and Rosenthal (2003) utilised the results of the inspection as an instrumental variable to prevent the effect of omitted variable bias. This includes certain economic and social factors. However, selecting an adequate instrumental variable is a mammoth task and its explanatory capacity remains controversial (Wen, et al., 2019).

The difference-in-difference (DID) is widely adopted for analysing the alteration of the capitalization effect of educational resources pre and post some determinants. This includes a specific event (construction of a new primary school facility, school ranking reassessment, and school district restructuring) or policy. For instance, Imberman et al. (2013) conducted this study in Los Angeles (LA) using the DID model to investigate the capitalization effect of freshly value-added information that has been recently published. The findings confirmed that no significant correlation exists between housing prices and value-added information, thereby, showing that such information is not necessarily of value to homebuyers. Ries (2010) employed the DID model to investigate the changes in housing prices pre and post the redesign of the school, and also proved a positive and significant effect of quality schools on housing prices. In addition, Wen et al. (2017a) determined that the ‘zero school choice’ policy promotes the capitalization effect of the compulsory school. Arguably, this policy implemented in Hangzhou, September, 2012 is deemed prohibitive of school choice and specifies that the admission of junior high and primary school pupils be firmly in accordance with the school district. In this study, the empirical findings indicated that this policy cannot result to real justice but can possibly exacerbate the separation of housing areas in accordance with inequality of educational and income opportunities.

2.8. Methods Relevant to the Research Based on Literature Findings

Figure 3: Empirical Methods Relevant to the Research

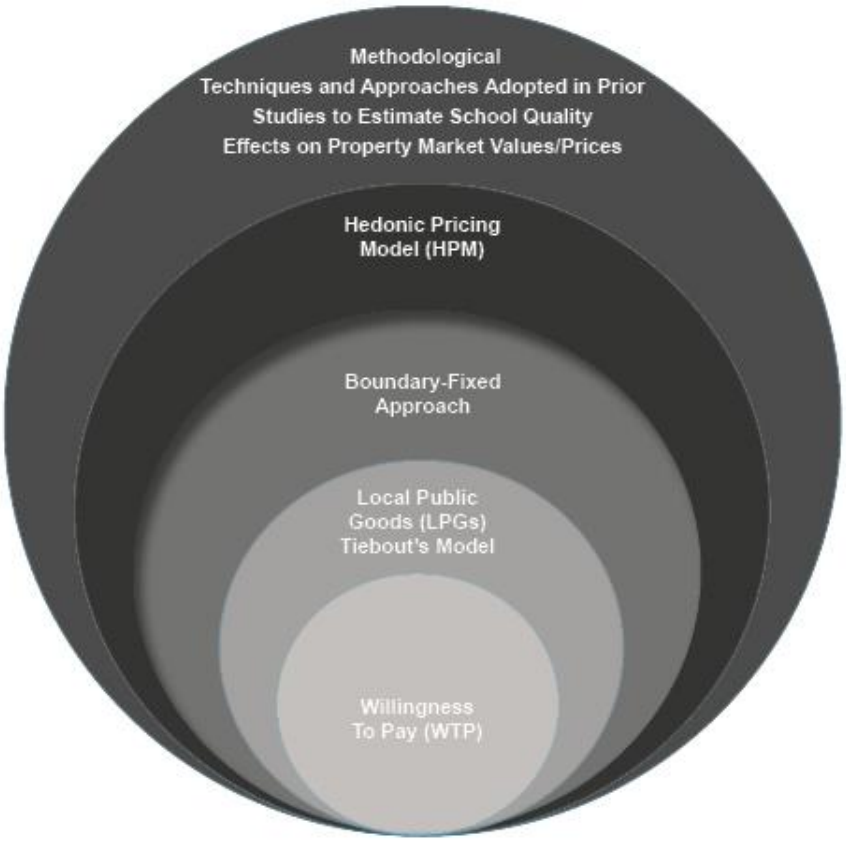


Figure 3 is an infographic illustration of the relevant models that were reviewed and interpreted in the current body of literature on hedonic modeling that are deemed relevant to constructing the research methodology for this study. The following section identifies and discusses previous methods and techniques employed in previous literature. In this thesis, the nuances behind the identified empirical models provide a hedonic framework to estimate and examine the implicit values of property-related parameters and their role in price changes. This section provides justifications why employing the best-fit instruments based on empirical findings is crucial to help in answering the research questions and objectives. Wulsin (2009) puts forward that existing literature indicates that although economists have employed numerous approaches to enable a measurement of the quality of a school, there is still lack of evidential consensus as to which technique metric is most viable when factored in housing prices. Hence the research aimed to overcome such difficulties. The four main concepts employed in the research include the following:

2.8.1. Hedonic Price Model (HPM) – A Review of International Studies

A history of empirical studies have widely used the hedonic price model as an ideal method of quantifying the value added by good quality schools to the price of property values. In this model, as noted by Sirmans (2005), a buyer's demand for a property is considered an aggregation of their demand for each single attribute of the property. Therefore, the hedonic method identifies a dwelling as a composite good, regressing its determinants on sale prices uncovers the marginal contribution of each attribute in the housing value (Owusu-Ansah, 2012). The research uncovered that using a direct comparison between two dwellings situated in a similar location, sharing equal attributes, the model held that if the quality of a good school is a valued variable to a property then the value of the house in the better school catchment area will reflect that value (Wild, 2016). The traditional hedonic model equation generally takes this format:

$$P = f(SA, N, LA) \qquad \text{(Equation 1)}$$

Which is expressed as, House Price = (Physical Attributes, Other Factors).

Equation 1 implies that the price of the residential property is a function of its physical attributes (location, bathrooms, square footage, age and various amenities) including other variables such as good school quality and external (neighbourhood) factors.

As a backdrop, the Willingness to Pay (WTP) concept can be collectively adopted along with the Hedonic Pricing Model in regard to the dynamics around observed and unobserved differentials. On that same note, hedonic housing prices can be viewed as 'implicit prices of an attribute' (Hummel, 2017), which can be observed from values of variated goods and the exact quantity of attributes associated with them (Rosen, 1974). Significance to the research arises through the consideration that this theory is not limited to a specific market, but the housing market that is appropriate for the hedonic pricing model. This market focus considers that the hedonic analysis enables the study to estimate the price for a particular characteristic (such as school-quality), because the housing market is ideally seen as an implicit market as a result of all goods (housings) which are traded or transacted in bundles. As put forward by Hummel (2017) (adopted from Sheppard, 1999), 'the demand for such housing goods are based on the attributes the houses embody. In addition, each attribute 'bears the utility to the consumer' and that enables a price for each attribute of the house (Wildeboer, et al. 2017). Significantly, the hedonic pricing model deconstructs the characteristics of the composite good (house) into implicit values for each characteristic.

Hummel (2017) argued that the derived marginal price obtained from a hedonic regression or function does not necessarily reflect a precise measure of a household's willingness to pay for a property with certain attributes. Arguably, in the real estate market, property valuations is the one of the prime indicators of demand and supply interactions. Maddison (2001) further accentuates that the hedonic model is expected to uncover demand parameters for the household represented provided it occurs on the ground of the restrictive condition of homogenous predilections. This considers that properties with homogenous attributes of socio-economic and income and different supplies will be reflected by the coefficients of the hedonic model which indicate the marginal willingness to pay of households (Hummel, 2017).

Debatably, scholars such as Maddison (2011) assume the hedonic model can provide a perfect equilibrium, which would require no transaction costs and perfect information. This supports the arguments made earlier relating to biased results. Wen et al. (2017) substantiates this stating that if equilibrium is not reached, or does not hold, then the implicit values obtained from the hedonic price model will reflect biasness. Moreover, disequilibrium issues usually appear as a result of increased variance in the estimated results rather systematic biasness. Wen et al. (2017) further argues that most criticisms relating to the hedonic price model stems of the derived estimates and not that it's based on the interaction of demand-supply and does not necessarily influence the hedonic price function. Although it considers that supply is not affected by individual consumer decisions (Wen, et al., 2017).

According to Freeman (1979), another detectable issue raised relates to the adjustment speed of the real estate market, which denotes changing the condition of supply-and-demand. It further substantiates that in the case of an incomplete adjustment situation, the observable marginal hedonic price will be incapable to accurately measure the marginal willingness to pay of the household. This is reasoned by that when there is an increase in the demand for an attribute, the marginal implicit price can underestimate the true value of a household's willingness to pay. This occurs since willingness to pay is not accounted or capitalised into market transaction costs that influence marginal hedonic prices. Until this is realised, the potential utility gains exceed the moving and transaction cost threshold (Freeman, 1979).

2.8.2. The Application of the HP Model in Local Studies

As astutely observed, there is a paucity of empirical research in urban economics about hedonic pricing models in the burgeoning African local property markets. The impetus to understand the intrinsic dynamics of the local markets has been overly stressed by property analysts in African cities.

Even though the MRA models employed are relatively similar, the variables adopted in the construction of property price indices in the national property models differ from the local models (Owusu-Ansah and Talinbe Abdulai, 2014). For this thesis, to use the SA national and the CoJ regional property price indices for market analysis and for investors to make informed decisions on the Kensington and Observatory local property markets, as Owusu-Ansah (2010) argued, may 'devalue the true appreciation' of property prices in those areas. He further highlighted in his study that making such local housing decisions may be deemed inaccurate because of the vast dynamics behind economic and geographical variations. For example, the house price levels in Kensington and the Observatory local housing market may be comparatively higher and more volatile in contrast to the entire Region F and the City of Johannesburg's house price levels. A notable review of a body of hedonic literature studies indicate that the response of prices to different variables varies across spatial localities. To increase market efficiency in African property markets, further examinations are ad-hoc on the application of hedonic model forms (i.e. linear, semi-log, and log-log hedonic models) to contributing to a better understanding of key idiosyncrasies in these burgeoning markets. In general, the semi-log and log-log specifications are commonly used for the empirical analysis, which is also supported by previous studies (see for example, Watkins, 2002; Wilhelmsson, 2009; Owusu-Ansah, 2014) which have also employed this functional form. Even though the existing methods to produce property price indices are voluminous, literature has not firmly concluded as to which methods have the highest performing accuracy in emerging and developed property markets.

Dodds (2010) paper examined 203 residential property sales in the region of West Rand (Gauteng) along with their representative attributes to produce a theoretical valuation of the observed housing transactions. The research used a formulated multi-variate regression model to predict sales from 1996 to 2009, a total of a 13-year sample period. The composite property index accounted for real house price growth, inflation, and locational determinants (which were sub-divided into locational quadrants used as dummy variables). The study further created an ordinary least square equation to run regressions on a screened dataset to reveal the relationship between the independent variables and the property values in the West Rand housing market. The actual sales and predicted values of the sampled property size were plotted on a relative fair value frontier against each other. Dodds mentions that the value/price chart reveals that the model identified 62 residential properties that were over-priced, and 32 properties that were under-priced based on the standard error lower and upper bound distribution. The noticeable discrepancies between the predicted values and actual sales price is the mismatch in pricing because of several factors that were not included in the

model. Despite the model's standard error, the applied regressions have practical contributions based on the high degree of fit revealed by the R^2.

Du Preez and Sale (2014) conducted an economic research on a suburban area called Walmer, Nelson Mandela Bay (Port Elizabeth, South Africa) to examine the effectiveness of employing assessed housing values as a proxy for actual housing sale prices. The study employed a hedonic price framework, namely, the Seemingly Unrelated Regression (SUR) model to run estimations and build test statistics to compare the estimated parameters in the two equations they used of the assessed value and the actual sale price. The accuracy of the model in producing reliable housing indices was miniscule. The null hypothesis of their study was rejected for the assumption that the coefficient estimates for the hedonic independent variables (i.e. house attributes) generated by the model that employed the assessed values are very similar to the actual sales price estimations. The empirical findings deduced that there are clear variations amid the distribution of assessed values and actual sales prices, with the assessed values, on average, being higher than the actual sale prices (Du Preez and Sale, 2014). The study postulates that hedonic practitioners should evade datasets from hedonic analysis that do not reflect the true market conditions.

Yacim and Boshoff (2019) study examined the extent and nature of urban housing price changes in Cape Town (South Africa). The study used data of 3,142 single-family property sale transactions in Cape Town in their hedonic analysis, with observations of a two-year period (from 2012 - 2014). They employed the weighted and ordinary least square estimators within the hedonic framework to measure the impacts of data-quality. Both models were specified using 12 housing characteristics including locational coordinates (x, y) to investigate the spatial attributes of properties within the study areas. The results indicate that the geographical weighted regression (GWR) outperforms the OLS method with predictions of the Cape Town property market for house valuations. They revealed in their hedonic analysis that the GWR showed more potency of correcting spatial heterogeneity in property transaction analyses. They further argue that the automation of value-based property valuations for taxation, such as the use of computer-assisted mass appraisal (CAMA) have gained prominence to become acceptable as a hedonic valuation technique to realize vertical equity and uniformity. Recessive dormant

Another notable publication conducted by Mnisi and Karam (2020) employed the hedonic OLS model, the linear and log-log functions using the STATA method to determine the change in value of bonded properties in close-proximity to the Fleurhof (Johannesburg) affordable housing development subsidised by the Integrated Residential Development Programme (IRDP). The observed case study areas that includes, Florida (Roodepoort), Orlando West

(Soweto) and Meadowland East (a suburban area in Zone 1, Soweto) to examine, in particular, whether there is a dormant depreciation effect of the IRDP housing development on the values of these properties. To uncover this effect, the period of observations used in the regressions were dated from 2001 to 2017 (a total of 16-years), in which their total observations accounted for 1 100 housing units that were tested for statistical significance. The regression output reveals that an approximated 87% of the variation in the values of these bonded properties is explained by the independent variables (i.e. locational amenities, neighbourhood characteristics, and structural attributes) in the model (Mnisi and Karam, 2020). In relation to the age variable which was absent in this thesis, but, instead employed a constructed dummy variable as a proxy, their paper regressed this parameter and reveal a beta coefficient value of -0.68 which decreases housing values. They show that an increase in the age of a property per annum, results to a decrease of 0.68 units in the property's market value. Evidently, their papers negative age coefficient buttresses the axiom that the age determinant is recognised as a depreciating variable in hedonic modeling. These bespoke nature of the multi-variate regression models truly obey the standard depreciating effect of the age indicator within the property's value, based on their physical condition (i.e. structural characteristics). However, this also depends on the maintenance or negligence of the property, which must be accurately valuated on the grounds that reflects the true market value attached to a property using reliable local datasets. It should be noted that this impact would be accurately captured when the age variable uses a matrix of data on the current physical condition and the age of the observed housing.

The empirical study conducted by Owusu-Ansah, Anim-Odame and Azasu (2021) examined the key macroeconomic determinants of house price trends in Accra (Ghana) using the vector error correction model (VECM). They observed 4,722 valid residential transactions (from a total of 5,103 transactions) in the construction of the quarterly hedonic price indices. To ensure suitability for their statistical analysis, the validation of the hedonic data required the removal of observations with missing variables. The study captures house price movements by constructing the 'quality-adjusted hedonic price indices', in which quarterly time dummies were included and the log of house prices was regressed on the macroeconomic variables in the hedonic dataset (also see Owusu-Ansah, 2018). They employed a dataset comprising of quarterly time series data with variables and indicators that approximated to a 25-year period. The study asserts that hedonic data quality is key to circumvent issues of producing spurious results. The paper further noted that to ensure the dataset used is suitable for the empirical modeling, various tests for stationarity and multicollinearity are necessary, particularly in time series analysis (Owusu-Ansah, Anim-Odame and Azasu, 2021). The results show that price movements in Accra are mainly driven by Gross Domestic Product (GDP) growth, interest

rates and unemployment; however, these significant variables do not have a long-term correlation with housing prices in the long-run.

The reviewed studies above provide nascent evidence about the application of the MRA models where property companies, investors and banks could benefit from data collection and recording of accurate property price indices in the African market. The most noticeable opportunities would be the determinations of reliable selling prices, the time and cost savings concerning the duration to prepare property valuation reports.

2.8.3. Boundary-Fixed Method

The boundary-fixed approach was initially adopted to attempt isolating the effect of school quality from the influences of neighbourhood attributes (Bogart and Cromwell, 1997; and Black, 1999; and Machin, 2011). This regression method attempts to alleviate the impacts of endogeneity (as widely deliberated in literature) prompted by unobserved neighbourhood attributes by viewing differences of house values across school admission boundaries. The principal argument for identification is that a good school quality's effects on property values can be identified from discontinuities at attendance zone boundaries as the model assumes that properties and local services should be similar on both sides of the boundary (Nguyen-Hoang, et al., 2011).

Numerous studies have adopted the use of the boundary-fixed effect, of which if employed can help in enabling a study to define these effects on property sale prices. This becomes more beneficial when included in the hedonic regression to control for neighbourhood traits that are unobserved and shared by the local residents in that school catchment area. However, the research does not assume that this approach will account for all unobserved neighbourhood variables. For example, Gibbons et al. (2012) arguably puts forward that discontinuity in property values across school attendance boundaries presumably emerges from various differentials, other location-specific services, or omitted geographic characteristics (for example, a rail station). The research is not inconspicuous of the issue of estimations of school's capitalization possibly being biased when households (situated on either side of the boundary border) do not share a similar school. As argued by Cheshire and Sheppard (2004), school attendance boundary-zones are not restricted to future developments ('non-permanent boundaries') which create ambiguity about the prospective location of a house (with children who need access to good quality local schools) in the long-run. Such complexities are largely predominant along school attendance boundaries, of which the school quality variable will only be capitalized partially into the property prices proximate

to boundaries (Nguyen-Hoang, et al., 2011). It can be assumed for the research that in the valuation of properties near boundary zones undergoing substantial new construction (change due to growth) experience a discounting of school quality. To substantiate the argument, Brasington (2012) puts forward that lower capitalization of school quality as a variable in, for example, areas of new construction emerges due to the housing supply having a high elasticity.

2.8.4. A Preference of Local Public Goods (LPGs) - Tiebout's Model

The adoption of this model is based on the idea that many household decisions of relocating to different locations stems from various neighbourhood factors or determinants. This conceptual framework is specifically associated with the behavioural observations Tiebout conducted in his 1956 research paper - which presented what became to be known as the 'Tiebout model' which laid groundwork for decadal empirical studies on house prices and the preferences for local public goods (LPGs) (Wulsin, 2009). According to the model, which states that people are more prone to 'vote with their feet' when making residential location decisions to reside in as a resultant of specific neighbourhood characteristics such as school quality (Yadavalli and Florax, 2014). In addendum to the above, it is identified that neighbourhoods with the most valued services such as good quality schools, ceteris paribus, become more expensive to inhabit.

2.8.5. Willingness to Pay (WTP)

The research adopts the Rosen 1974 theoretical approach to determine the house price premium when capitalized with good school quality. Purposively, considering the cumbersome nature of this empirical model – will assist in identifying how neighbourhood characteristics (for example, local public amenities) and property attributes influence housing prices within local markets (Black, et al., 2011). The model is centralized around how a market equilibrium in which preference of a consumer over housing equates to choosing an opportunity cost of a bundle of services such as housing variables (for example, size of the house) and neighbourhood amenities (for example, good school quality). This relatable concept will help supplement the existing conceptual groundwork to give a narrative direction for the research. In addition, this approach also provides justification for employing the use of expenditure on a property to monetize the benefits attributed to its observable characteristics. Therefore, given the probability to determine how much property (such as housing) expenditures alter with marginal changes in a single attribute (such as good quality school), ceteris paribus, then this

can be assumed as the willingness to pay for that particular school quality, or supposedly its implicit price (Black, et al., 2011).

2.9. Discussion on the Main Hedonic Modeling Issues

In the research, it is imperative to understand the main sources of hedonic problems, namely, omitted variable bias, heteroscedasticity and multicollinearity. This section highlights the underlying contributors and how they should be addressed when identified in hedonic modeling.

2.9.1. Omitted Variable Bias, Multicollinearity and Heteroscedasticity

Omitted variable bias is a common source of hedonic problems. Wooldridge (2006) and Greene (2008) note that the HPM produces bias estimators when several unobserved property attributes are likely excluded from the model. Generally, this arises when the specification of the hedonic model omits a fundamental attribute (predictor variable) in the regression modeling (Long and Wilhelmsson, 2020). This means the functional form in the model is poorly defined and the measured parameters will presumably be bias. As seen in literature the issue of including a multitude of explanatory variables in the model, omitting a single important characteristic will produce bias coefficient estimates. Wen et al. (2019) also noted this issue. To an extent, it is still a mammoth task to collect a database containing all the important observed and unobserved property attributes. This includes both macroeconomic and microeconomic variables that have an effect on sale prices. Furthermore, literature does not provide a specific technique of addressing the issue of omitted variable bias. Therefore, in the research employing two different model specifications to separate the omitted variable that indicates signs of high correlation (either in a negative or positive direction) with the dependent variable or other independent variables is a crucial effort.

Multicollinearity occurs when some predictor variables in a multiple regression model indicate a very high correlation level. Fletcher et al. (2004) and Wilhelmsson (2019) refer to multicollinearity as a phenomenon seen when several linear dependencies exist between the independent variables. The presence of multicollinearity can be identified in numerous ways in the hedonic regression models. For instance, the bivariate correlation tool can also be used to observe if there are high values between explanatory variables, which indicates the presence of multicollinearity. In addition, Wilhelmsson (2000) notes that when there are large variations in the measured coefficients after the omission or addition of an independent variable then presumably there is problems of multicollinearity. However, imperfect

multicollinearity is more common than perfect multicollinearity. For example, this can occur between the erf size and the number of bedrooms. In the research, ceteris paribus, if the erf size increases, the number of bedrooms will also increase. The estimated regression coefficients may be less reliable, so to examine the strength and direction of the variable coefficients these independent variables can be separated between two models with appropriate specifications and analyse the stability of the coefficients.

Heteroscedasticity describes a decreasing or increasing level of variability of the disturbance or error term within a multiple regression model (Curwin, 2013: pp. 479-483). Investigations indicate that the variables used for analysis of the sample is one of the major sources of heteroscedasticity. The most highlighted attributes in hedonic literatures is the age of the property and its external environment, which lowers the property sale price and causes higher variances in the model. Other indicators that cause heteroscedasticity include outliers, measurement errors, omitted variables, and the use of inappropriate functional form in the model specification. This is most prevalent in cross-sectional data and in time-series data that deals with long periods of data, which have an increasing variance over time (Curwin, 2013: pp. 482-486). Although the problem of heteroscedasticity can be mitigated by using a larger sample size, it may however result in changes to the estimated coefficients and the standard errors of the estimated parameters (Wilhelmsson, 2019). There is a paucity of literature studies that provide robust standard procedures to test for heteroscedasticity in the hedonic models, and ensure the *t*, and F statistics are valid, and the standard errors are adjusted.

2.9.2. Selection of the Hedonic Functional Form

Selecting the appropriate function in hedonic modeling is key when one wants to examine the implicit prices for the property attributes. This is noted as one of the major issues in the outcomes of regression coefficients that are plagued by endogeneity problems that have a great impact in biasing the results. For instance, Woolridge (2009) argued that the functional form selection is not of significance to every hedonic study, because some of the predictions can be recognised as unbiased even when the regression model produces coefficient estimates that are bias. Researchers (see, Dodds, 2010; Du Preez and Sale, 2014; Owusu-Ansah, 2018; Yacim and Boshoff, 2019; Mnisi and Karam, 2020; and Owusu-Ansah, Anim-Odame and Azasu, 2021) have employed and investigated a myriad of hedonic functional forms in their property price analysis, of which specifications also includes the:

- *Linear model*; this is also known as the level-level model in which both the independent and dependent variables are expressed in levels.

- *Semi-log model*; this represents the semi-logarithm model, whereby the independent variables are not transformed to a natural logarithm form, but only the dependent variables is transformed.

- *Log-log model*; this model uses the independent variables in a transformed log form and the dependent variable is expressed in absolute terms.

The linear hedonic function is considered to be restrictive, which confines the robustness of the empirical analysis, which is also supported by previous studies (Woolridge, 2009; Dodds, 2010; Ludick, 2019). However, it should be considered that the accuracy of the parameter estimates is not based on the choice of the functional form because the regression estimates are not subjected to some criterion on which the best fits models are tested. For the empirical analysis, the researcher uses the untransformed property characteristics observed in the hedonic dataset but only log transformed the dependent variable (transacted prices). The benefits of employing the semi-logarithm function is that it allows efficient analysis of the estimated parameters.

2.10. Literature Gaps

The predominant issue pervasively detected in prior studies attempting to quantify the effect of good school quality on property prices is separating the school's net effect. This specifically focuses on determining the level to which a change in the link between property sale prices and good school quality feeds back to parent's willingness to pay a higher premium to reside in close-proximity to a good school. In that same breath, relevant prior studies identify the difficulty in indicating the etiological correlation between property sale prices and good quality schools. This also owes to the existing relationship (or lack thereof) between neighbourhood attributes and good quality schools (good schools are situated in good quality neighbourhoods) (Black and Machin, 2011). This considers that the correlation between property sale prices and school quality indicators is measured without controlling for differentials that reflect neighbourhood attributes (for example, socio-economic level) of which the results can be potentially be obscured by noise upward. This is exacerbated by presence of the main hedonic problems during modeling, which includes, but not limited to, omitted variables bias, heteroscedasticity, multicollinearity, spatial autoregression and spatial correlation as identified in the research. The aforementioned are also considerate of the flaw in the capitalization of specific school qualities into house prices as a result of inadequate data utilized (Nguyen-Hoang, 2011). The failure to control for certain variations can result in biased estimations of the extent of school quality capitalization into house prices. Exploring these

under-searched dynamics and contributing to areas where further knowledge is required should indicate how the results may have an effect on housing choice.

2.11. Conclusion

The significance of property price indices cannot be overstressed, as it has stimulated scholarly interest to conduct empirical studies that investigate the performance of the various regression methods. This section culminates the literature chapter and extrapolates analytical nuances on the hedonic models and their production behind the theoretical and practical valuations of property prices. The chapter has discussed several property price indices in the international and local markets of Africa. The schools of thoughts on regression methods assisted the study to curate a comprehensive methodological framework that includes the hedonic price model, boundary-fixed method, local public goods approach and the willingness to pay method. The identified composite of methods accounted for exposing the true appreciation nature of the school quality variable when capitalised into property prices.

Although the existing approaches and techniques to construct price indices are voluminous, scholars have not firmly concluded as to which methods have the highest performing accuracy in the burgeoning African markets. With regard to the literature suggestions, it can be noted that the accuracy of methods profoundly depends on the comparisons of the techniques and the nature of the dataset used. The research debunks and accentuates the intrinsic value and shortcomings of these underlying index methods employed in the examination and analysis of property indicators. To date, the development on multi-variate regression analysis methods (i.e. hedonic models) have gained prominence and precedence overtime in the valuation of property variables.

The notable review of the empirical studies in the African body of literature that provide nascent evidences on the use of hedonic models are significantly documented and cannot be bypassed. Although their contributions deviate from the main-focus of this thesis, it is worth noting that a vast majority have extended fruitful insights that accurately depict regression coefficients in the construction of local property price indices. These academic provisions simulate and capture an array of housing prices and their responses to different variables across numerous geographic localities (see, Dodds, 2010; Du Preez and Sale, 2014; Owusu-Ansah, 2018; Yacim and Boshoff, 2019; Mnisi and Karam, 2021; and Owusu-Ansah, Anim-Odame and Azasu, 2021). These scholars support the statistical validation of the hedonic dataset to circumvent spurious regression estimates and expand the spectrum of authenticated knowledge in African cities. The researcher practically considers that price movements are driven by various macroeconomic and microeconomic indicators in the short-

run, but in the long-run most of these significant variables do not have a long-term relationship with prices. The overall market forces that determine the trajectory and momentum of these annotated dynamic shifts in the national, regional and local levels are spatially transverse. Hence hedonic deliberations astutely buttress the need to increase market efficiency in these mushrooming local areas to ensure hedonic practitioners report information that reflect the true market conditions. The next chapter presents the hedonic models and data employed in the construction and interpretation of the research findings.

3. Chapter Three: Research Design and Methodology

3.1. Introduction

This chapter presents a detailed enumeration of the research design and techniques employed in the research. The study makes use of secondary data sourced from Lightstone Property (Ltd) of sales transactions of residential and commercial properties (closed-corporations) as recorded in the deeds registry (dated from 2010 to 2020). The type of data obtained aided in filtering down the choice of method, sampling size and techniques to enhance the quality of the results and their associated discussions. To enable the research to measure the attributable implicit-value associated with the quality of model C schools and other unobserved local characteristics on property sale prices requires multiple regression models. This included the Statistical Package for the Social Sciences (SPSS) and R Studio software's, and the Hedonic Price Model.

3.2. A Quantitative Approach

The research method employed in the study is quantitative. A quantitative approach emphasizes objective estimates and the mathematical, statistical or numerical analysis of collected data (Babbie, 2010). As put forward by Thompson (2015) and Saunders et al. (2016), often a link between a quantitative research and positivism exists, which is the paradigm of the research. According to Kivunja and Kuyini (2017), this may be reasoned by the specific nature of quantitative studies that undertakes collection, analyses and interpretation of primary or secondary data as a method to uncover a particular phenomenon or significant correlation between the underlying variables.

3.2.1. Research Model Specifications

The research identified the relevant variables, namely, housing prices, structural attributes, and locational attributes that are best suited for estimating the effect of school quality. The paper investigates the extent to which these selected variables are capitalized into property sale prices. To achieve this, the study adopted the hedonic pricing model (HPM) as the best-suited method. HPM researches are largely missing at the local level in contrast to the OLS model; hence, this paper aims to fill in the knowledge gap in the local regions of the South African context. The choice of method is based on several advantages associated with the method, which includes the model's versatility and ability to be adapted to consider a myriad of correlations (or lack thereof) between market goods and other determinants that form the composite of property sale prices (Hummel, 2017). As a backdrop, the above review of the

HPM is not inconspicuous of the limitations of the model, which will also be highly considered when employing the model, as meticulously outlined in chapter 2.

3.2.2. The Main Econometric Model: Hedonic Pricing Model (HPM)

The HPM assists in determining how a myriad of neighborhood characteristics of different localities influence property values (Structilia, 2020). Traditionally the model is based on a multiple regression model, which considers various value determinants as the independent variables and property transactions as the dependent variable (Bello, 2009: 8). According to the model, heterogeneous goods (real estate properties) can be described by their characteristics, essentially that a good is a bundle of several characteristics (De Hann, et al., 2013). For example, in the housing sector this bundle can be constituted by characteristics of both the location and structure of the property. This considers that individual characteristics have no market; thus, they cannot be traded in isolation of the other property attributes. Hence, the values of these characteristics are not observed independently (De Hann, et al., 2013). Relevant to the research, the supply and demand for a property implicitly determines the marginal pricing contribution of a property's characteristics to the value of the property. For the research, implementing the HPM through regression techniques assisted in estimating those marginal price contributions. In addendum, the HPM can also be used in estimating the willingness of households to pay for the various characteristics attributed to the property.

The traditional hedonic method illustrated in Equation 1, makes the assumption that the sale price (*SP*) is observed as a function of its structural attributes (*S*), local neighborhood characteristics (*N*), and good school quality (*SQ*), which in the research is set as follows:

Model I Equation

$$SP = f(SA\beta, N\gamma, LA\alpha) + \varepsilon_n \qquad \text{(Equation 2)}$$

Where, *SP* is a vector representing the observed logarithm of property sale prices; *SA* is a vector representing the physical attributes of a property; *N* is a vector signifying the local neighborhood attributes; *LA* represents the distance of a property to a locational amenity; and ε_n denotes the error term or the disturbance in the model**.** The relationships between the property sale prices, independent variables, and a property's determinants are described by α, β, and γ parameters.

Structural Attributes (SA) describes the internal characteristic of a property. This includes the physical attributes such as the erf size, the number of the bathrooms and bedrooms, and the number of garages (Wen, et al. 2017). These attributes form part-and-parcel of the determinants that contribute to the sale price of a property. Locational Amenity (LA), measures the accessibility of an urban amenity such as good quality school to the desired targeted subject (Kohlhase, 1991: 2). It should be noted that certain locational attributes of a neighborhood have adverse effects whilst others have beneficial influences on housing prices. In the case of this study, Jeppe High School for Boys has a positive influence on the prices of the surrounding residential units. For the research, considering the direct influence of other unaccounted neighborhood factors that help improve the purchase price of a property is imperative to ensure the provision of non-bias and real transaction costs of a property. This also assists in avoiding properties becoming inconspicuously undervalued during valuations.

Neighborhood characteristics (N), illustrates the social condition and economic quality of the neighborhood. This includes the quality-state of community infrastructures and other amenities, and level of municipal service provision relative to the research. The study used other unobserved neighborhood characteristics, which included quality of schools (Jeppe Boys High School) located near the sampled areas.

For the research, the HPM was employed to help answer the objectives outlined in chapter 1. The objectives state that school quality differentials have a significantly positive correlation when capitalized into the selling price of a property. The approach includes independent variables (based upon data access) of school quality attributes (for example, performance index, test-scores) or dummy variables of a school's district (for example, improvement school district). The approach used in the research follows Equation 3:

Model II Equation

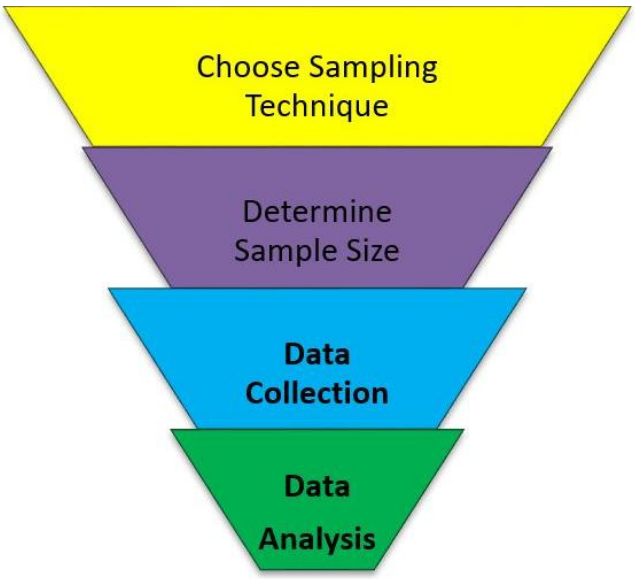
$$\text{Ln}(SP_n) = \beta_0 + \beta_1 S + \beta_2 N + \beta_3 SQ + \beta_4 LA + \boldsymbol{\varepsilon}_n \qquad \text{(Equation 3)}$$

Where**,** LnSP represents the sale price in a log functional form. *S* is a vector for the structural attributes of a residential property. *N* is a vector that comprises of local neighborhood attributes. *SQ* is a vector representing school quality, including independent variables of a school district such as input factors (for example school sporting performances, teacher or expenditure per pupil), output factors (for example, school district report card, proficiency test scores), dummy variables of a school district, and value added. *LA* represents the distance of a property to a locational amenity, and $\boldsymbol{\varepsilon}_n$ is the error term. In Equation 3, the β1, β2, and β3

parameters measure the size of the relationship between property prices and the respective independent variables included (S, N, and SQ) in the HPM. In particular, these parameters measure the additional change in the sale price, that is, the marginal cost which homebuyers have a propensity to pay for the incremental change in the attribute, holding all the other physical attributes constant (Mnisi and Karam, 2020).

3.3. Research Design (Sampling Technique)

Figure 4: The Sampling Process



Source: Alasuutari, et al., (2012)

Figure 4 is a diagrammatic representation of a sampling process this study adopts. Primarily, sampling aims to obtain a representative sample or small collection from a larger population size (Alasuutari, et al., 2012). Taherdoost (2016) further asserts that sampling makes inferences about a specific population target or accurate generalizations about the larger population group (samples like the population) in relation to existing theory. Quantitative researchers typically adopt the probability (more than the non-probability) sampling technique based on theories of probability (Alasuutari, et al., 2012). According to Zikmund (2002) random or probability sampling provides the researcher with more autonomy from bias estimates. However, in terms of time and energy it may appear the costliest for a given level of sampling error (Brown, 1947) hence this study adopted a non-probability sampling approach.

3.3.1. Sampling Technique

In this study, similar to most prior studies have relied on non-probability sampling (Rowley, 2014). Saunders et al. (2016) states that this significantly assists in ensuring reliability of the findings. Asiamah et al. (2017) puts forward that obtaining a large sample size of will ensure avoidance of biasness. Specific to this study, the data was screened for wrong entries, missing values, distribution normality, homoscedasticity and linearity. This resulted in 1177 spurious property transactions (observations) that were not retained for analysis due to missing information that could have caused regression errors such as heteroscedasticity and multicollinearity.

3.3.2. Sample Size

Specific to this study, there is no target population that will be composed. This study used secondary data obtained from different sources. The sample size is based on the property price list sourced from Lightstone Property (Ltd). The number of properties obtained amounted to 3066 in two different locational areas, namely, Kensington and Observatory situated in region F of the JMSA. The data was received on an excel spreadsheet and it was sufficiently large to be used for regression analysis. The sample was reduced to 1889 properties due to missing data in the remaining sample size, such as the number of bedrooms and bathrooms. Therefore, inclusion of such incomplete data would be an arduous task during regression modeling and analysis.

3.3.3. Data Collection Techniques

The hedonic data obtained for the study is secondary data sourced from Lightstone Property (Ltd) containing different property types in those sampled areas, which was the main source of information for the adopted regression models. The data was categorized based on three essential determinants constituting a property's value as assumed by the HPM. This comprises of location, neighborhood characteristics and structural attributes of a property. In the research, the HPM uses the purchase price as the dependent variable and the number of bathrooms and bedrooms, registered date of the property, and the erf size as the independent variables. These variables are preferred by most scholars researching this topic, due to a property being considered as a hedonic good (impacted by both extrinsic and intrinsic attributes). The extrinsic variables include, neighborhood characteristics and locational amenities, and the intrinsic variables includes structural attributes. As at December 13, 2010, to March 3, 2020, there were 3066 registered property sales in the deeds registry in both areas

of Kensington (1684) and Observatory (1382). There were only 1889 valid observations from this sample size, which qualified to be included for the regression analysis.

3.3.4. Justification of the Data Period Chosen (2010-2020) – Time Dummy Variable

The period under examination of the sourced property price indices is a total of 10-years from December 13, 2010, to March 3, 2020. An attempt was made to source a dataset for a 15-or-20-year period, and other neighbourhood variables such as traffic noise to include the empirical modeling. However, it was recognized that the dataset obtained for all the necessary hedonic variables was consistently available for only this 10-year period. This also considers that a mismatch in the coverage of the independent variables and the dependent variable can lead to the results being statistically insignificant. It should be noted that the use of certain variables (i.e. neighbourhood characteristics, location dummy variables) have limitations to explaining the variations of property trends. Therefore, this considers that property prices are averaged over geographic boundaries. The reasons that informed the decision to use this period are manifold (also see section 3.6.).

The first practical consideration that informed this decision is that, Observatory and Kensington had their most voluminous amount of property transactions recorded during this period, particularly from 2012 to 2016, in contrast to other preceding years. One of the major significant and practical contributors to these property market movements during this period was the CoJ embarking on a myriad of transit and precinct oriented developments. Integrated developments have become a major attraction for housing demand and supply, which can explain the high number of sale transactions (see Table 5 and Graph 1). Another clear advantage is the researcher's familiarity with the contextual setting of the study areas and the proximity to access these residential markets, which makes them suitable for the research. This also made it possible for the researcher to verify some of the data (i.e. locational characteristics and neighbourhood amenities), and conduct site visits to make general observations of the property attributes and their conditions. This was done to allow the analyses of the findings to be accurate and to justify why the methods used can also be applied to other local property markets.

The second reason that influenced the researcher is that the only available information the Jeppe Boys High School could provide for the school quality (predictor) variable only backtracked to a 10-year period. The choice of period of property transactions was also necessary to ensure a consistent construct of indices during regressions. One of the strengths of this study is the time series data employed and the robust methods used to empirically

explore and highlight the underlying price dynamics in two local property markets. The researcher also had insufficient funds and resources to procure an extended period of transactional data, which could not be obtained in time, and therefore, limited the number of observations. For future research, conducting this study using quarterly data or a longer period of analysis may assist in revealing more uncovered value-laden empirical evidences. This will augment the application and extend knowledge on the groundbreaking hedonic studies that are currently existing in the body of African literature.

3.3.5. Data Analysis and Discussions

The raw dataset of the property transactions and a myriad of unobserved attributes were captured in Microsoft Excel, where it was tested, coded and rectified. The corrected data were further transferred into the SPSS database in which the processing and analysis of data was conducted using the software.

The empirical analysis approach was applied to evaluate the relationship (or lack thereof) between the quality of schools and property sale prices. To achieve this objective, the multiple regression model was used along with the HPM and R Studio with the log of the sale prices representing the dependent variable, and the property-related indicators (including school quality as the main factor investigated) represents the predictor variables. The predictor variables stipulated in table 1 are consistent with the model variables used in prior similar studies, and are used in the research to support the analysis. The use of these selected control variables is reasoned by their well-known interrelationship in the regression analysis using the HPM (Kane, et al. 2005, and Hummel, 2017). The data of properties observed are located within a radius of 4km around Jeppe Boys High School. The measured distances of the sample areas were recorded as 3.3km (Kensington) and 3.5km (Observatory) away from the school. This was to ensure that the case study areas fall within the attendance zone of the school district JBHS is situated, as reference to the Department of Education's standard km radius requirements from a schools in Gauteng.

To add depth to the research, other hedonic datasets were collected to add to the model and clarify the quality of a school based on the data type used in prior studies. These included the pass rate in percentage form from 2010-2019 (Jeppe, 2020), the sporting performance (based on an average rate) of the schools top 4 sports activities (Rugby from 2010-2019, Cricket from 2010-2020, Hockey from 2010-2020, and Water Polo from 2010-2020) (SASchoolSports, 2020). The approach to defining school quality for the research was constructed based on a proxy rating system designed from relevant prior sources that also influenced the inclusion of

distance from each property to the Jeppe Boys High to broaden the horizon of how these properties are influenced. The study used R Studio to analyze the data.

3.4. Description of Model Variables and Sources of Data

Table 1 contains the description of the models explanatory variables used in the research. The adoption of nominal variables were largely informed by the nature of data and the effort to submerge into the details of the information to enable further policy suggestions. For example, the registered year of each property were noted as nominal variable, which depicts a negative relationship due to their high level of correlation.

Instead, for policy formulation and direction it is imperative to assess which predictor variables are positively or negatively related to property sale prices. This same logic is applied to other continuous variables. The relevant sources chosen to collect data for the research emerge from prior studies that attempted to address their research aims, objectives and hypothesis in this same subject area. The most cited sources of data in prior scholarly papers included the following: School District data; Department of Education database; Housing survey reports; GIS/Arc-GIS; and Real Estate Companies or Firms. The table below presents the most relevant property-related attributes chosen to answer the objectives outlined in chapter 1.

The hedonic data for both of the sampled case study areas were obtained from Lightstone Property (Ltd). The dataset was secondary information, which required cleaning and rearranging in an Excel sheet. The purpose of excluding unusable data and re-coding the variables was to ensure that the data contains valid and suitable observations for hedonic regression and analysis. The formatting process was also to ensure that the selected models obey the traditional regression assumptions, and that the specifications of the hedonic functional forms are accurately defined. These approaches were necessary to adjust the dataset for multicollinearity and further test for signs of heteroscedasticity. To maintain up-to-date relevancy, this secondary data is backtracked to a 10-year period from 2010-2020. The significance of these selected variables were substantiated by the hedonic model being a derivative of a property's extrinsic (locational characteristics such as quality of schools and its distance to each property in Kensington and Observatory) and its intrinsic attributes (structural components). Table 1 presents the model variables, which includes the sale price as the dependent variable and the independent variables (bathrooms, bedrooms, erf size, year, school quality and distance).

Table 1: Description of Model Variables

Category	Variable	Description	Measurement Scale	Expected Signage
Dependent Variable	Sale Price	The selling price of the property - based on the current market value	Continuous	
Independent Variables				
Structural Attributes	Bathrooms	The number of bathrooms the property contains	Continuous	+
	Bedrooms	The number of bedrooms the property contains	Continuous	+
	Erf Size	The size of the land plot	Continuous	+
	Year	The year the property was last registered or sold (employed as a proxy for 'Age of the Property')	Nominal	-
Locational Attributes	School Quality	Overall pass rate of Grade 12's in the Jeppe High School for Boys – based on a 10-year period (2010-2020).	Continuous	+
		A proxy of Jeppe Boys High School's sporting performance rates over a 10-year period.	Continuous	-

Primarily the research focus required the author to derive a proxy for school quality. As put forward by Saunder et al. (2016), it is imperative to analyse data when it is in a raw form due to the issue of lack of meaning attached to it. The significance associated with this undertaking was to ensure a provision and presentation of coherent and reliable findings of untapped information in the context of Johannesburg Metropolitan Statistical Area. As a backdrop, essentially the sourced data were tested to establish any exhibition of a normal distribution, which required a parametric method of analysis (Saunders, et al., 2016). It is therefore imperative to prepare the dataset to best-fit the study before it can be employed. On another note, the best-fit proxy chosen to accurately quantify school quality in the study is not limited

and has a higher probability of successfully answering the hypothesis due to using more than a single component (i.e. overall pass rate, sports results and current conditions of facilities) of a school to define its true quality. Including sports and the condition of facilities, assists in substantiating why parents would be more inclined to pay high property premiums in Kensington and Observatory for their children to attend in a school like Jeppe High School, which also includes the following:

1. A well-maintained pass rate average – from 2010 to 2020;
2. A good overall performance in various sports; and
3. Daily maintenance of the schools facilities, including repairs and/or refurbishments.

3.4.1. A Proxy for the Age Variable - Year of Property's Registration (Dummy Variable)

Commonly, the age variable is employed into the HP model to control for the effect of depreciation on property-transacted prices (Owusu-Ansah, 2018). A generic consensus in hedonic literature indicates that it is the most highlighted contributor of causing higher variances and heteroscedasticity in the model. In this study, the age variable was reported absent from the sourced Lightstone's (Ltd) database (obtained from the Land/Deeds Registry database) because it does not contain such information for these specific sampled areas. This was a major pitfall in the dataset. In addendum, local level indices have not yet been produced for both the Kensington and Observatory urban areas.

The cumbersome nature of the dataset obligated the researcher to develop a dummy variable from the registered year of each property sale that was used as a proxy to represent the age of the properties. Adopting the registered year of each property did not deter nor bias the outcomes of the research as it assisted in significantly capturing the impact of the age variable on the value of the properties. The approach was flexible and had the ability to produce reliable estimations of observations from the raw dataset that had missing values. Therefore, hedonic issues of omitted variable biasness were resolved and under control. The advantage of the research was using two accurate hedonic functional forms informed by the dataset of a local level, which had a sufficient number of observations and fundamental variables (physical and locational attributes). The study further compared the accuracy undertaking several diagnostic experiments to verify and ensure reliability in order to produce robust coefficients. Notably, there were no signs of heteroscedasticity being problematic in the hedonic model. This is further substantiated by the low VIFs (smaller standard error terms) the model produced because of the large sample size used.

3.5. Reliability and Validity

The primary key indicators of measuring the research quality are validity and reliability (Saunders, et al., 2016). Reliability measurements are used to assess the stability of estimates administered at various periods using a similar standard, of which similar results should be obtained during each application of the instrument to an invariable phenomenon (Heale and Twycross, 2015). Validity is also applicable in trying to explain how the findings represent the data (Noble and Smith, 2015). Validity measures the level of pertinence found within a research instrument, the level of accuracy of the analysis and if whether the findings are generalizable. In accordance to the requirements of the selected instruments (SPSS, HPM, and R Studio) in the research, cleaning and arranging the data helped reduce error terms during regression analysis.

The research used a large sample sized data from Lightstone Property (Ltd) due to its reputable recognition of being a source of reliable data. Data for pass rate was sourced from Jeppe Boys High School (including the) and sporting performance rates were obtained from SASchoolSports (2020). A comparative analysis was also undertaken of the data and findings produced by prior studies and the secondary data collected for the research report to determine possibilities of variation in implications that may not be explained in other studies. For the paper, it was imperative to determine which instruments were appropriately applicable to the research aims, objectives and the hypothesis being tested, if applicable. As it is put forward by Saunders et al. (2016) reliability lingers on a clarification of the causal effect of variables on each other. Hence, the study used validating statistical techniques of analysis to ensure valid clarification of any effect or correlation.

3.6. Limitations of the Research

The subsequent content provides a list of possible limitations prevalent to the research:

- Secondary Data

- The overall dataset had 3066 property sale prices (observations) with 1177 transactions with missing values (structural attributes) such as the number of bedrooms and the number of bathrooms. This forced a reduction of the dataset to 1889 property transactions to avoid statistical errors during regression modeling. The sample size used was sufficiently large for consideration in the regression analysis. In this manner, the researcher believes valid and consistent findings were obtained. In addendum, the database had no age variable of each property for both sample areas, which is a prerequisite based on the HP model. However, the ‘registered year’ of each property was included during regression modeling

and analysis. Overall, these outlined circumstances limited the study in terms of generalisability to numerous areas within the South African region. Nonetheless, the models adopted in the research can also be used to run similar experiments in other areas.

- Time and Resources are Limited

- Due to a plethora of prior similar studies, the use of school quality has been exponentially expanded in terms of definition followed by criticisms of authors providing biased findings. Hence, the study opted for the use of Jeppe Boys High Grade 12 overall pass rate, sports performance and conditions of facilities (from 2010 to 2020) as a proxy to define the 'quality' of the school over a 10-year period. Therefore, the data being limited to a decade to measure its overall quality is arguably insufficient in contrast to using a 15-year period or more.

- Data for traffic noise could not be obtained. This was a result of limited time and resources to adequately measure noise decibels. This was further exacerbated by the Covid-19 pandemic which limited movement and access, particularly gated-communities such as the one identified in the sampled urban areas (Observatory). In addendum, the obtained listed property data did not provide any useful information that could have allowed a possible measure of the effect of transportation noise on specific properties and their associated prices and structural attributes.

3.7. Ethical Considerations

In terms of research studies, ethical considerations entail adhering to certain legal standards that govern processes of how to conduct a research (Escobedo, et al., 2007). This considers that the primary importance of ethical researches is honesty, particularly if the study engages with human participants (Saunders, et al., 2016). However, the research does not involve participants or require direct contact with individuals or organizations. Primarily, the necessary data were sourced from Lightstone Property (Ltd) in the form of a database containing information (i.e. property sale prices and their associated structural attributes) of properties situated within Jeppestown High School for Boys attendance catchment area (5 km radius). Lightstone Property (Ltd) ensures confidentiality and anonymity of all the sampled households personal information as such information was not provided to the researcher. This helped safeguard the data and treat the information in the strictest confidence.

4. Chapter 4: Results and Findings

The previous chapter has identified the multi-variate regression analysis method and further outlined the hedonic data that was employed in the estimation of property price indices. The chapter provides justified discourses on the practical reasons that influenced why the case study areas were chosen. This includes a synopsis of the local property market under investigation as established in chapter 1 (see section 1.4 and section 4.2). Two different models are implemented in the construction of the property price index for the analysis in this chapter, which are based on the hedonic method. This chapter discusses the results established from estimating the property regression coefficients.

4.1. Presentation of Data, Analysis and Discussions

The subsequent chapter presents the main findings based on descriptive statistics, correlation analysis and regression modeling as per the econometric methodologies identified in Chapter 3. These descriptions are imperative since they provide a synopsis of the underlying variables before they are employed in the regression modeling. The employed techniques were largely informed by the nature of the obtained data. This enabled a reflexive and comprehensive deconstruction of the coefficient results. The study adopts the hedonic pricing model, SPSS software and the R Studio program to conduct a series of hypothesis experiments. Regression results helped uncover the dynamics behind the effect of school quality, structural composition and locational determinants on property sale prices in the sampled areas. The presentation of the results and related discussions is buttressed and situated within existing literature (also see, Annexure A). Evidently, the hedonic price model provides a good fit for the data utilised in the research. Both model I and model II significantly explain over an estimated 26% of the variation in the property sale prices within the combined sampled areas. Only 74% is of the variation is unexplained. The findings unveil that the school quality of Jeppe Boys High has a marginal significance in explaining the cause of variation in property prices situated within this school catchment area. In addendum, the estimated coefficient results of all the other structural attributes (excluding the year of deeds registration) are significant and are in the direction predicted in Table 1.

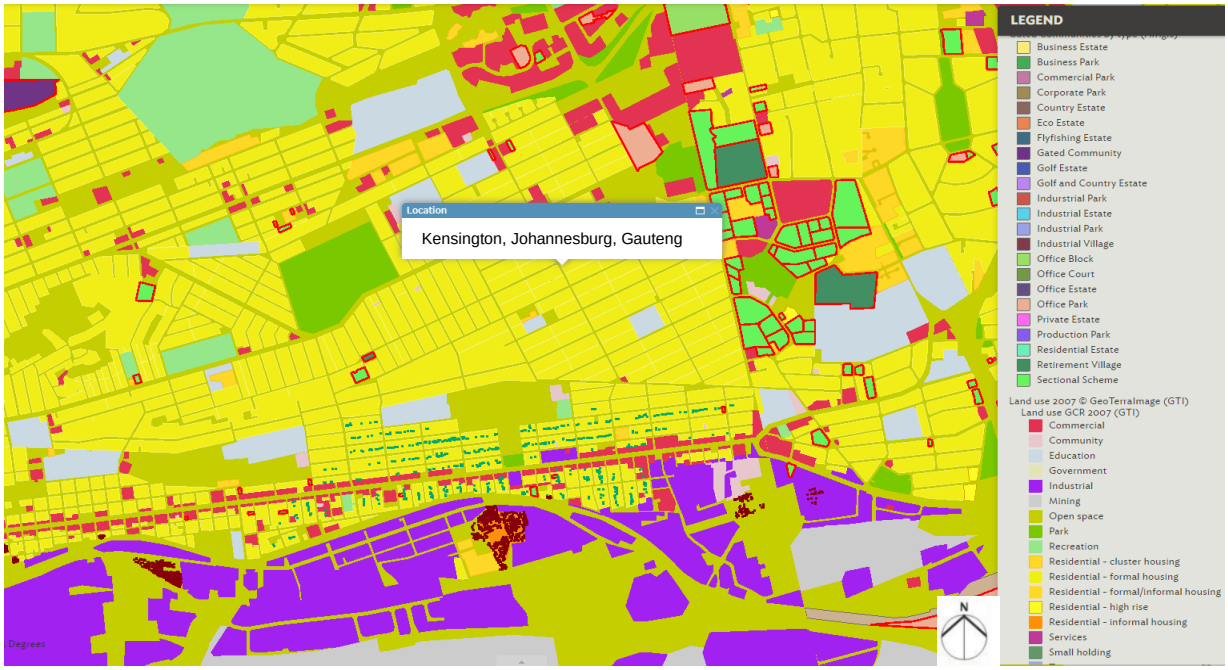
4.2. Spatial Structures of Sampled Areas in Close-Proximity to the Focus School

A review of the case study area's reveals that the household types in the area differ between formal and informal settlements (De Wet, et al., 2008) as opposed to Jeppestown containing the least informal properties in comparison to neighbouring areas. The property types varies from sectional schemes and freehold properties and a great majority are privately owned.

Based on the Lightstone Property (Ltd) title deed registry data, the ownership type in Kensington includes 28 closed corporations, 56 company owned, and 1599 privately owned properties. In Observatory, there are 5 closed corporations, 20 company owned, and 232 privately owned.

The subsequent maps (2 and 3) present the current spatial footprints of both sampled areas as adapted from GCRO (2021). The post-2010 gradient of spatial planning in this particular real estate market, amongst other urban nodes within this region, exhibits a fairly coordinated usage of resources in the built environment space, implying to the supply of land with improvements. Reeds and Sims (2015: pp. 317-321) argue that this is a critical objective realized - to avoid conflicting spatial structures that are restrictive to efficient future developments. Empirical evidence indicates that planning policies such as the National Development Plan and the Integrated Development Plan in CoJ are gradually restoring decaying infrastructures in pivotal areas and integrating segregated spatial structures to enable a flexible land-use assortment that is sustainable for future developments.

Map 2: The Land-use Layout in Observatory

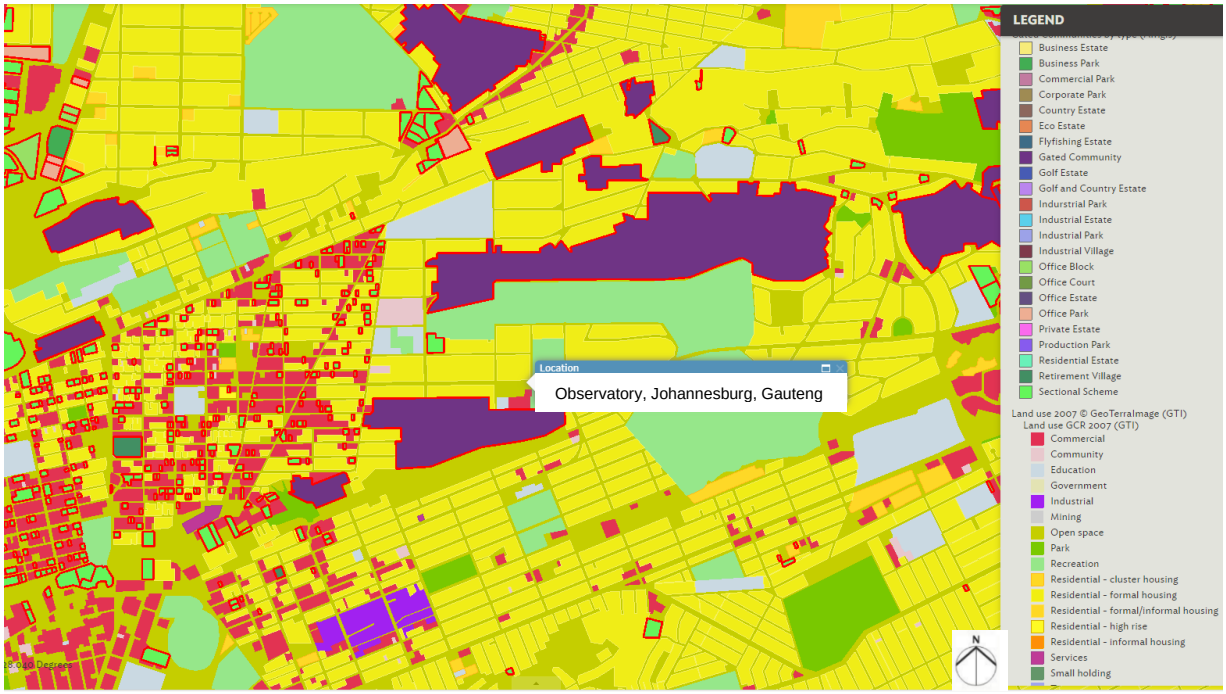


Source: GCRO (2021)

The education district in this school catchment area is saturated with formal residential housing, high-rise residential areas, commercial nodes, office blocks, sectional schemes and few gated communities amongst other existing land uses. Thereby, the location provides consumers with a bundle of goods and services that have measurable capitalization effects on the valuation of property sale prices, as a result of their symbiotic interrelation within the

real estate built-space. Although improvements eventually become physically obsolete, it is prudent to establish whether the market value of these amenities is accounted during property valuations.

Map 3: The Land-use Layout in Kensington



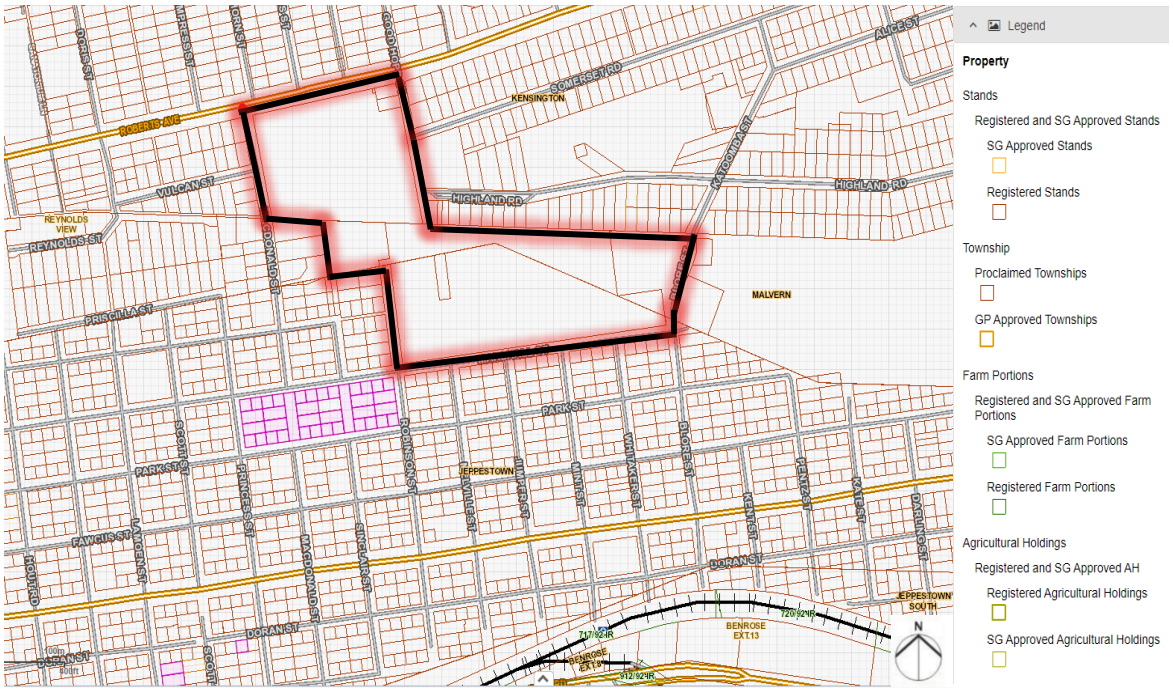
Source: GCRO (2021)

4.3. JHSB - Background Trajectory of the School's Quality

Map 4 exhibits a localised scale of the defined school boundaries of the focus object in the research. The Jeppe High School for Boys is a public educational facility in Kensington, Johannesburg Gauteng. The JHSB School was established in 1890 and is renowned as the oldest public school constructed in the Johannesburg Municipality (TravelGround, 2014). In 1897, the school buildings were purchased by the Witwatersrand Council of Education and later renamed the Jeppestown Grammar School. Under the ownership of the new council major improvements and renovations were imminent to enhance the built-up landscape of the school's former facilities. In 1990, the unique stone buildings were constructed on land donated by the former owner (Sir Julius Jeppe) and the old buildings in Troyeville became occupied by the Prep School (TravelGround, 2014). During the 1986 celebration of the Johannesburg Centenary, the First World War Memorial (dome-shaped) and the main building upon declaration became national monuments. Other distinguished school buildings include

the Oribi House (built in 1912, and known as the oldest hostel), Roan House that was assembled strikingly with its stained-glass windows and the wooden staircase, and Friedenheim stables which currently are used as change rooms (Jeppe Boys Association, 2018).

Map 4: Spatial Footprint – Demarcating the Boundaries of Jeppe High School for Boys



Source: ArcGIS (2021)

The threshold underpinning the observable quality transition over a historical period of the school's facilities and resources demonstrates progressive development. The above illustrations are a basic simulation of the trajectory of the improvements the educational facilities of Jeppe High School for Boys have undergone. Figure 5 exhibits the traditional Main building that was initially used by Jeppe Boys. Figure 6 depicts the extent to which the primitive main building of the school has evolved to its current quality condition. Day-to-day maintenance plans are a prerequisite to avoid building structures and other essential educational facilities in undergoing obsolescence as an act of negligence of their deteriorating structural integrity.

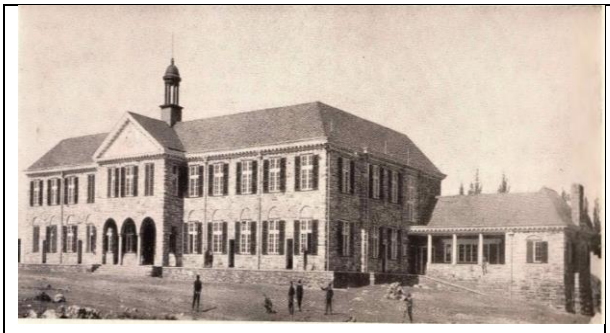


Figure 5: Jeppe Boys in 1911.

Source: African Architect (2018).

Figure 6: Jeppe Boys Main Building.

Source: Jeppe Boys Association (2018).

Figure 7: Improvement and Restoration of the main Gate (Dickie Fritz Gates).

Source: Jeppe Boys Association (2018).

Figure 8: Jeppe Boys Association (JBA) - No.12 Highlands Road.

Source: Jeppe Boys Association (2018).

Figure 7 demonstrates the restoration and improvement of the Dickie Fritz Gates in order to provide a fitting entrance and gain a competitive aesthetic edge. Subject to the SASA Act 1996, section 21 (16) (1) (a), financing the maintenance of structural buildings within a school's premises is a required function from all registered schools for section 21. Hence the JHSB in September 2017 started a reconstruction project of the gates which cost about R250,000 including a new guard house that was built to replace the old wooden hut which was deemed inadequate and unsightly (Jeppe Boys Association, 2018).

Figure 8 shows one amongst the other new properties (Number 10 and 12 Highlands Road) acquired under the program initiated by the JBA. The Jeppe Boys Association (2018) added that alterations and additions of the new houses were completed to avail for teachers in the

school to use as accommodation. The aim of the program is to build a fund that will enable the JBA to purchase all the properties bordering the school campus as a strategy to expand the school facility. JHSB also has other quality educational and sport facilities, which includes a library, museum, two halls, Astroturf, computer centre, swimming pool (water polo), cricket field, hockey field, rugby field and tennis courts, rowing facilities, basketball courts.

4.3.1. The Quality of JHSB Facilities: Fieldwork Findings

This section is a visual presentation of the fieldwork findings of Jeppe Boys High current conditions of its facilities. Jeppe High is a registered section 21 school mandated to undertake regular maintenance and management procedures regulated by this Act (see chapter 1 section 1.3).

I. *MAIIN SCHOOL BUILDINGS*

Figure 9: Façade of Jeppes Main Hall Building Condition

Figure 10: Landscape of Jeppes Main School Building

Figure 11: Memorial Building

II. *IN-HOUSE BOARDING FACILITIES*

Figure 12: Droste House

Figure 13: Sessebe House

Figure 14: Thabana House (Sable)

Figure 15: Oribi House

Source: Jeppeboys (2020).

Figure 16: In-House Estate Agent's Residential Dwelling

III. *SPORTS FACILITIES*

Figure 17: Sports Office

Figure 18: Rugby and Cricket Field (Dual-Purposed Sports Facility)

Figure 19: Alan Francis Rowing Centre

Figure 20: Hockey Sports Field

Figure 21: Cricket Training Centre

Figure 22: Swimming Pool Facility (for Water Polo)

4.3.2. JHSB Overall Sports Quality and Performance Rankings

I. *<u>RUGBY</u>*

Figure 23: Jeppe Boys Rugby first team. Source: Jeppeboys (2020).

Figure 24: Jeppe Boys Rugby in 2019. Source: Jeppeboys (2020).

The popularity behind rugby as a sport at the Jeppe Boys is reasoned by the long history and tradition of the school, with it being the first and oldest sport there. According to Jeppe Sport (2020), rugby is the biggest winter sporting code at the school. The school can field up to 21 teams in regular interschool rugby games and policy has it that each team plays a single match on a weekly basis during the second term and often carried over to the third term. The first

team competes in tournaments such as the Grey High School Festival in Port Elizabeth and the Easter Rugby Festivals. This includes the U16, U15 and U14 teams also participate at the Grey High and attend the festivals every year. Over the years, emphasis on quality improvements through coaching for all team is mandatory. This is seen by the school obtaining specialists and consultant assistance from the Golden Lions Rugby Union. This contribution has helped the school maintain an average of 22.970 points at 6th place from 2010 to 2019 with a gap of 0.398 points from its closest contender with an average of 22.572 (see, SASchoolSports, 2020). Affies outran Jeppe Boys with an average difference of 12.473 points but this does not discredit their formidable and indomitable competitiveness.

II. CRICKET

Figure 25: Jeppe Boys High Cricket U14A.
Source: Jeppeboys (2020).

Figure 26: Jeppe Boys High Cricket U17.
Source: Jeppeboys (2020).

Cricket at the Jeppe Boys High has also become a very popular sport due to its successful performances over the past decade. The school can field up to 19 teams in regular interschool cricket games. Jeppe encloses four fields including a state-of-the-art indoor practicing facility, 12 turf nets and turf wickets. The activity usually takes place in the first and third terms and the boys over the years have participated in tournaments and festivals running in various parts of the country and in the Johnny Waite Knockout games. Jeppe Boys has a long list of provincial and national honours achieved by the Jeppe cricket players that is associated to the school's long and proud history. In addition, Jeppe currently stands as one of the top 10 schools to obtain a Blue Chip Status in the country by Cricket South Africa (CSA). The boys

from Jeppe enjoyed a phenomenal run considering their ranking positions from 2012 to 2020 ending no lower than 28th on the SA School Sports Cricket ranking list. Their average of 29 is 3.33 points slightly better than Framesby (28th ranking position) their closest contender with an average of 32.33 points (see, SASchoolSports, 2020).

III. HOCKEY

Figure 27: Jeppe Boys High Hockey Team U14A**.**
Source: Jeppeboys (2020).

Figure 28: Jeppe Boys High Hockey Team U16A.
Source: Jeppeboys (2020).

Hockey at Jeppe Boys High has a decadal successive run as a sporting at the school (Jeppe Sport, 2020). The school fields around 16 teams in weekly interschool matches. The hockey team only lost a single match last year to Westville BH (at 2nd position with an average of 23.20 points) and Jeppe Boys at 3rd place with an average of 22.78 points with a difference of 0.42 points from 2012-2020 ranking list (see, SASchoolSports, 2020). The point gap between Jeppe Boys and Maritzburg College (currently standing at 1st position) is 2.44 (SA School Sports, 2020). Last year Jeppe Boys first team successfully continued to play competitively at Malaysia against a men's team and managed to lose a single point. Differentiating between the other existing sports at the school, this hockey team has managed to build its acumen and rapport quicker considering the competitive position of rugby, cricket and water polo. This winter sport is usually played in the 2nd term with all age groups competing in annual festivals, with the first team regularly competing at the national festivals against the top schools in SA (Jeppe Sport, 2020). They are also part of the Gauteng Top 8 tournament. Moreover, on a regular Jeppe contributes numerous hockey players to the Southern Gauteng Provincial teams with a few players having gone to earn national colours (Jeppe Sport, 2020).

IV. *<u>WATER POLO</u>*

Figure 29: Jeppe Boys High Water Polo Games of U16A in 2020.
Source: Jeppeboys (2020).

Figure 30: Jeppe High Water Polo Games of U17A in 2020.
Source: Jeppeboys (2020).

Water polo is amongst the most popular sporting options at the Jeppe High School for Boys with a successive level of competency over the last decade. The school's teams all use the aquatic centers within the premises to undertake both their practice sessions and play matches. These teams are provided coaching and other support by the old boys. The water polo first team has been playing competitively over the years in the most prestigious national tournaments and during each season, the teams go on to play at least in one tour. The most well-known competition hosted every year by Jeppe is the U14 tournament called the Ken Short Festival. According to the SA School Sports (2020), Jeppe Boys current ranking position is 21 with an average of 17.65 points (see, SASchoolSports, 2020) with a gap difference of 15.64 points from Clifton Boys High currently standing at the first position with an average of 33.29 points. Over the past 8 years, the team has only managed to rank in the top 30 of hockey schools in the country.

In summation of the aforementioned, the top 4 sporting codes in the Jeppe School premises (hockey, rugby, water polo and cricket accordingly) maintained a well-rounded average all-together and a top 30 ranking amongst the best sports schools in the country. Quality of a school embeds various elements that propel parents to choose Jeppe Boys and further influence their willingness to pay the current housing premiums offered in areas in close-proximity to the school. This includes suburban areas such as Kensington, Observatory, and Fairview. Therefore, good sports performances such as these over a decade indirectly contributes to the quality offered by the school. The condition of facilities for teaching, staff and sports also contribute to the aesthetic quality the school supplies denoting that quality can

be stripped in various forms depending on the best-fit description and use. The attractiveness of the school places the surrounding property market at a more competitive stance in contrast to other areas in the Johannesburg Metropolitan context with a substandard quality (i.e. education, sports).

4.4. Descriptive Statistics and Associated Discussions

Primarily, the study explores the effect of school quality, location, and property structural attributes on sale prices. A proxy for school quality was necessarily formulated to highlight key factors that strengthen what should represent the quality of a former model C school, such as JHSB, based on secondary data. The property transaction data used (dated from 2010 to 2020) was obtained from Lightstone Property (Ltd), covering properties, which appear on the formal market (deeds registry) in Kensington and Observatory (situated east of region F, in CoJ). This consisted of the sale prices, structural attributes, and the registered year of each property. The secondary data had 1889 valid observations from a total sample of 3066 property transaction sales. The remainder of the provided properties had null values (*N*=1177) for the description on the number of bedrooms and bathrooms. Missing data were removed using the econometric software's (SPSS and R-Studio) to avoid biased statistical coefficients or errors (residuals) during modeling, hence the reduced number of valid observations. However, this did not deter the significance of the coefficient statistics as the indicated valid sample of properties is deemed sufficient.

Table 2 below presents the summary statistics of the variables used in the regression model. As indicated in the table, the average property sale price approximates to R1 614 703.38. Over the observed period, the lowest transaction price is almost R20 000 whereas the highest transaction price is R39 300 000 (the most expensive property in that matter). The standard deviation is almost R8 138 478. A small median (relative to the mean of R1 614 703) shows that the distribution of the property prices in the two study areas is positively skewed. The recorded erf sizes for both combined areas ranges from $0m^2$ to 409 510 $000m^2$, which is evident that spatially the data is not evenly distributed which is cured through log transformation. Each property transacted over the 10-year period in both sampled areas has an average erf size of around $257545.55m^2$. The main variable (school quality) under investigation provides an average pass rate of 97.45%, oscillating from 91% to 99.5% (2010-2020) as the highest performance rate and a standard deviation of almost 2.36%.

Table 2: Descriptive Statistics of Model Variables

Variable	Observations (N)	Min.	Max.	Mean	Std. Deviation
Dependent Variable					
Sale Price	2842	20000	393000000	1614703.38	8138477.692
Independent Variables					
Structural Attributes					
Erf Size	3066	0	409510000	257545.55	8926430.440
Bedroom	2006	1	15	3.17	1.142
Bathroom	1958	1	10	2.16	1.104
Locational Variables					
Pass Rate (school quality)	3066	91.0	99.5	97.445	2.3618
Rugby	3066	22.97	22.97	22.9700	.00000
Cricket	3066	6	54	29.31	17.839
Hockey	3066	22.78	22.78	22.7800	.00000
Polo	3066	17.65	17.65	17.6500	.00000
Valid N (listwise)	1889				

Evidently, the investigated model C school ascertains to improve the value of properties within its school catchment area. The average percentage rates obtained in most prior studies (see Annexure A) of transaction prices are relative and supports the estimated coefficient result obtained in this study (also see Black, 1999; Gibbons, et al., 2007; Chiodo, et al., 2010; Machin, 2011; Wild, 2016; and Wen, et al., 2019).

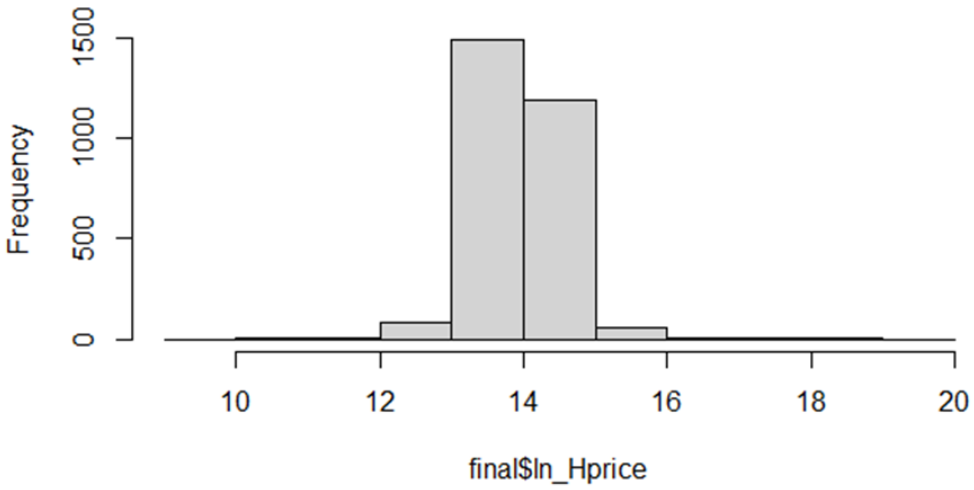
4.4.1. Idiosyncrasies of the Diagnosed Property Sale Price Fundamentals

The dummy registered year in which the properties were registered shows a very strong correlation with pass rate, which initially led to exhibiting problems of multicollinearity. For example, a high correlation between the 'dummy year and school quality' denotes that the regression coefficients are unstable when the standard errors increase. Notably, the issue of multicollinearity was not exaggerated merely because it does not focus on coefficient values but more interested in the sale prices utilised in estimating the quality-adjusted sale price index. Multicollinearity is a well-known issue, particularly during hedonic regression modeling (Wen, et al., 2019) which has persistently brought scepticism and lack of reliability in the literature findings and models adopted by prior studies. To cure the issue that is conspicuously notable in most prior attempts, the researcher opted to use both univariates but in separate models. Hence, model I and model II as stated in section 2.7.1. This method of separating highly correlated predictors helped to build more credibility of the statistical findings by avoiding omitted variable bias during hedonic modeling. Considering the nature of data used, multicollinearity only has gargantuan impact on variables that are highly correlated, and for this study the R^2 is small, denoting that the regression coefficients and statistical analysis are still informative and reliable. To further cure multicollinearity issues, a series of diagnostic tests were conducted using a log transformation approach for the entire data (also see Graph 3, 4, 5 and 6).

Regression modeling has assumptions that the univariates ought to be normally distributed, which was obeyed in this study to elude multicollinearity, which is subjected to transparently mirroring biased-laden results. This was realised by log transforming the variables depicting a skewed distribution to enable a linear correlation between the interlinked property determinants. Drawing from Graph 1, a log transformation of Ln_SPrice was more appropriate than using the original data due to the cumbersome nature in which the information was provided. The sale prices present a normal distribution after transformation. It can be noted that most properties averaging a total of 1500 fall within the price ranges of 13 million to 14 million. The data also indicates that approximately 1200 properties in both areas are sold within the high ranges of 14 million to 15 million. The data further indicates that the most

expensive properties sold between 2010 and 2020 are the Est FH located in Kensington. The least expensive property type in both areas is the Est SS with most of these properties located in Observatory. As exhibited by Graph 1, it is observantly evident that most properties with sale prices over 15 million have the least transaction market. Expectantly, higher priced properties are associated with a lower demand, but it is subjectively dependant on the income level, affordability, and willingness of the potential consumers. During this 10-year period, the city had embarked on a plethora of precinct and transit oriented developments such as the Louis Botha BRT corridor and providing new capital infrastructure in region F (IDP, 2015/16). Integrated developments have become a major attraction for housing demand, which can explain the high number of sale transactions within this property market during this period.

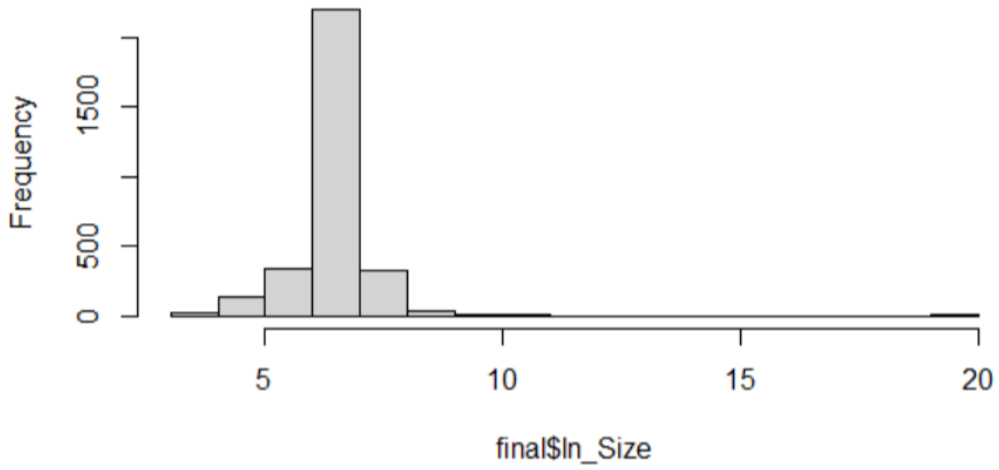
Graph 1: Distribution of the log Sale Prices in Both Sampled Areas



Both sample areas are located less than 6km from these strategic interventions, and such market changes have a direct effect in the rise of housing demand in such areas. A fair census in region F were also enabled by the income growth experienced in CoJ to enter the affordable property market (IDP, 2013/14), hence this observed increase in sale transactions particular in the resale market of residential properties. Surprisingly, the affordable market (of CoJ) exhibited signs of declining transactions from 2016 and 2018 as the buying power in properties dropped decimally, but the conventional and high-end markets saw bulk property sale activity (CAHF, 2018). In addendum, the log transformation of the erf size (see Graph 2) was also appropriate which depicts a positively skewed distribution in contrast to using the original data. Apart from the sale price and erf size that needed transformation, there is no strong motivation given the shape of the graphs of school quality, number of bathrooms and bedrooms to transform these variables (see Annexure B). As revealed by diagnostic measures and theoretical literature, using a log transformed sale price in the HPM to estimate

the effect of school quality (i.e. pass rate including school sports) provides high explanatory power.

Graph 2: Distribution of the Log Erf Size in Both Sampled Areas



4.4.2. The Correlation Nature between Model Variables

The research employs the 'Pearson product moment correlation' to quantify the existing strength and direction of a linear relationship among two quantitative variables. The correlation coefficients range between -1 and +1, with a negative correlation indicating that when one variable increases the other variable will likely decrease, and a positive coefficient indicating that two variables have a tendency to increase together (Saunders, et al., 2016). A correlation of zero exhibits no relationship between two variables (Utts and Heckard, 2007: 166).

Table 3 indicates the correlation results between the predictor variables, which was necessitated to flag any possibilities of multicollinearity. Similarly, Black and Machin (2010) reported in their study that school quality only had a marginally significant effect on property sale prices, which may be explained by a strong or weak correlation between independent variables. Generally, the sale price of a residential property has a direct relationship with the independent variables that constitutes the determinants of a property's price. It should be noted that only correlates deemed as significant from the Hedonic Price Model were considered in the computation of the R-Studio Program, however, no variables were omitted during modeling. Table 3 also helped in detecting any possibilities of multicollinearity, but is not sufficient in confirming this, hence the use of the models in section 4.4.

Table 3: Pearson's Correlation Results

Independent Variables		School Sporting Activities							
		Size	Bedroom	Bathroom	School Quality	Rugby	Cricket	Hockey	Polo
Size	Pearson Correlation	1							
	Sig. (2-tailed)								
	N	3066							
Bedroom	Pearson Correlation	-0,003	1						
	Sig. (2-tailed)	0,880							
	N	2006	2006						
Bathroom	Pearson Correlation	0,017	.583**	1					
	Sig. (2-tailed)	0,460	0,000						
	N	1958	1956	1958					
School Quality	Pearson Correlation	-0,008	0,013	.045*	1				
	Sig. (2-tailed)	0,665	0,548	0,045					
	N	3066	2006	1958	3066				
Rugby	Pearson Correlation	.a	.a	.a	.a	.a			
	Sig. (2-tailed)								
	N	3066	2006	1958	3066	3066			
Cricket	Pearson Correlation	0,012	-0,031	-.077**	0,006	.a	1		
	Sig. (2-tailed)	0,500	0,159	0,001	0,739				
	N	3066	2006	1958	3066	3066	3066		
Hockey	Pearson Correlation	.a	.a	.a	.a	.a	.a	.a	
	Sig. (2-tailed)								
	N	3066	2006	1958	3066	3066	3066	3066	
Polo	Pearson Correlation	.a	.a	.a	.a	.a	.a	.a	.a
	Sig. (2-tailed)								
	N	3066	2006	1958	3066	3066	3066	3066	3066

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
Note: .a Signifies that the univariates cannot be computed due to at-least one of the variables being constant. Meaning there is no correlation.

4.4.3. Analysis of the Strength and Direction of Property-Related Univariates

In Table 3, the correlation between the erf size and number of bedrooms (r = -0.003) and school quality (r = -0.008) implies that there is a very weak and negative relationship between these structural attributes. This explains that when the square meters of a property's erven is smaller the number of bedrooms are also expected to be lower in number, particularly private properties. In addendum, the relationship is also extremely weak due to the correlation coefficient being significantly distant from the 0.01 (1%) and 0.05 (5%) significant levels, meaning the test is statistically insignificant. The correlation diagnostic results that exhibit a statistically positive relationship with erf size include the number of bathrooms (r = 0.017) and sports quality of cricket (r = 0.012). Similarly, both coefficients are distant from the 1% and 5% significance levels, meaning that the relationship is weak. The various sporting qualities (average performances from 2010 to 2020) are included in the variables to augment the overall quality of the Jeppe Boys High School. This is to test their level of influence on parent's inclination to enrol their children in a school with a good sporting profile. Moreover, to evaluate the influence it has on their housing-decisions to reside in close-proximity to the school. However, sports such as rugby, water polo and hockey did not give any valid statistical results or significance due to the data for each named sport having a single constant value.

The relationship tests that exhibit a positive correlation with bedrooms includes the number of bathrooms (r = 0.583) which is statistically significant at the 1% level of significance and shows a moderately strong correlation. This result is not surprising since structural attributes of a property are highly associated and complementary to each other, for example, the number of bedrooms will surely determine the number of bathrooms needed by a household. The number of bedrooms also indicates a positive correlation with the school's quality (r = 0.013) which is also statistically significant at the 1% level of significance. The quality of sports exhibits a negative relationship towards bedrooms with r = -0.031. The correlation test between the number of bathrooms and school quality shows r = 0.045 which is statistically significant at the 5% significance level, demonstrating a strong relationship. The more significant variable in the study is the school quality, which expectantly illustrates a positive statistical relationship with the school's cricket (r = 0.006). However, the erf size shows a negative correlation of r = -0.008, which can be reasoned by that the quality of a school will not influence a parents willingness to purchase a bigger or smaller plot size. This also includes the number of bedrooms and bathrooms a household decides to have is purely associated with preferences and level of affordability and not suggesting that school quality directly influences housing-decisions in both combined sampled areas.

4.5. Presentation and Related Discussions of the HPM Empirical Results

The adopted estimation strategies for the research are based upon the HPM theory. The regression results have provided interesting data for discussions. Notably, the predictor variables used in both models have followed the traditional results of hedonic price evaluation. The models further obey the regression assumptions regarding the nature of the appropriated data during modeling. The adopted model uses the natural log of sale prices in order to maintain a normally distributed data and avoid probable issues of multicollinearity and heteroscedasticity. These are two highly unnoticeable oddities in most prior results of scholarly studies, which has deterred and biased the significance of some of their coefficient estimates. Purposively, curing these issues was ad-hoc in the research to avoid rendering similar modeling problems that are simple by-products of highly correlated univariates in the models, which is meticulously discussed in the subsequent sections.

4.5.1. The Hedonic Regression Method - Model I

As a parametric method, the use of hedonic regression modeling in the research was important when dealing with skewed data variables. In the research, the property sale prices and erf sizes were log transformed (curing variable skewedness) to normalise the data distribution. This regression technique also aids in minimizing, if any, the role of outliers in the sample data. Fletcher et al. (2004) noted that outliers often appear as heteroscedasticity in the model. As also indicated, isolating highly correlated univariates that generally lead to such hedonic issues during model estimations is a valid option. Hence, the use of appropriate functional forms in both model I and model II. Eluding from omitted variable bias the model intrinsically improved the robustness of the coefficient results.

The hedonic pricing model employed estimates the marginal contribution of each property attribute and the school quality (a matrix of dummy variables) to the selling price. The hedonic function used in model I is of the form:

$$LnSPrice = \beta_0 + \beta_1 * \text{lnErfsize} + \beta_2 * \text{Bedroom} + \beta_3 * \text{Bathroom} + \beta_4 * \text{Schoolquality} + \beta_5 * \text{Sports (Cricket)} + \varepsilon_n$$

(Equation 4)

Where:

- LnSPrice is the dependent variable, which represents the property sale price in a natural logarithmic functional form. To make the interpretation of the regression

coefficients easier the study used the log transformation of the sale price. The implied position here is to help minimise the hedonic issue of heteroscedasticity, 'as the percentage adjusts in the value given a unit change in the property characteristic allows for variations in the currency of each property attribute' (Owusu-Ansah, 2010);

- β represents the regression coefficients related with the property and locational attributes (exogenous independent variables), in which;
- B_0 is used as the intercept,
- β_1 is the erf size,
- β_2 and β_3 represents the various structural attributes of the property, namely, the number of bedrooms and the number of bathrooms, respectively;

- B_4 and β_5 is a matrix of dummy variables, which represents a proxy for the school quality variable, in which:
- B_4 is the average pass rate observed over a 10-year period, and
- β_5 is the sporting performance of cricket, which was the only valid sporting variable after running regression experiments); and

- ε_n – represents the error or stochastic term, used to show the deviation of an observed y-value from the sample regression line. The formula is denoted as error = (y-ŷ) (Utts, et al. 2007: pp. 162-163). This helps to capture the relevant attributes of property n that were not included by the matrixes, meaning no issues of omitted variable bias exists.

The first hedonic equation is applied to derive the property price index from the coefficient estimation of the school quality dummy variable. From Equation 4, the log transformation of the property sale prices and erf sizes were used to minimize the risk of outliers in the sample data and cure variable skewedness. The school quality variable and the year variable were regressed in separate hedonic equations to avoid high multicollinearity and endogeneity issues. The issues highlighted can also be cured by using other regression techniques, such as, the Treatment Error (TE) model (also see Wilhelmsson, 2021).

4.5.2. Estimation Results and Associated Discussions

In Table 4, the study reports the results for model I obtained from empirical modeling. The first hedonic regression model is a direct measure of Equation 4. The model illustrates an adjusted R^2 of 0.261, which denotes that 26.1% of the level of effect of the independent variables on property transaction sales is explained by this model. The regression results confirm that the coefficient signs of the predictors are as expected.

Table 4: Model I – Hedonic Price Model Estimates.

Variables	**Coefficient Value**	**Standard Error**	**VIF**
Intercept	9.358***	-0.454	
Log of Erf size (lnSize)	0.275***	-0.013	1.062
No. of Bedrooms	0.008	-0.011	1.554
No. of Bathrooms	0.068***	-0.011	1.528
School Quality (pass rate)	0.028***	-0.005	1.004
Cricket (sports performance)	-0.001***	-0.001	1.007
Adjusted R^2	0.261		

Notes: *** and ** indicates statistical significance at 1% and 5%, respectively.

4.5.3. The Empirical Value Associated with JHSB School Quality on Sale Prices in Both Sampled Areas

The school quality of Jeppe Boys High influences purchase prices significantly, and the price semi-elasticity of the regression coefficient is 0.028 (statistically significant at the p = 0.01 level). This implies that each upgrade in the school quality (pass rate) will result to a 2.8% rise in property sale prices in both areas. Black's (1999) study indicated that parents have a measured willingness to pay an additional housing premium of 2.5% for an increase of 5% in

their children's test-scores. Quality improvements of educational facilities generally lead to changes in property values (Zheng, et al., 2016). Evidently, this fairly low school quality percentage of Jeppe Boys High can be explained by the presence of other good quality schools (kindergartens, primary and high schools) within this same school district and other locational attractions (such as local amenities). However, this is also conditional on where most learners live, in which it is indicative from the findings that there is an ample choice and competition amongst these good local schools. Notably, the enrolment policy at Jeppe Boys High covers areas outside of this school catchment district, namely, Boksburg, south Alberton and Edenvale (Jeppe, 2020) meaning property (housing) sale value may not be very responsive to inter-school variances in quality. Thereby, confirming the hedonic theory arguing that economic agents take housing-decisions by making conversant trade-offs between local amenities and property attributes (Senick, 2019). This pays tribute to Tiebout's (1956) standard urban economic theory as the paper meticulously highlighted in the review of prior literature.

Arguably, the existence of school-district-housing may possibly broaden the income gap further between the rich and the poor. Housing assortments based on affordability and educational accessibility within a certain distance can widen the disparity of providing equal education between rural, urban and metropolitan schools. Alternatively, the research also sheds light that any increase in demand for housing in and around these localities can also be explained by the two sample areas being well-located and not necessarily being in close-proximity to Jeppe Boys High. Distance to and from the school only serves much purpose for households that prioritise travelling time, which has a marginal effect on the selling prices of properties. School transportations are also offered by the school as a premium package (optionally included in the fees) for children coming from these distant areas. The school receives approximately fourteen-to-sixteen thousand applications per annum from all these different locations including neighbouring suburbs (Jeppe, 2020). Arguably, the abovementioned show that demand and willingness differs across space based on what amenities the locations offer, affordability and preferences that maximize their bundle of opportunities. These discussions purposively highlight the discrepancies between the study and other dissimilar prior studies.

4.5.4. Computing the Competitive Rate of JBHS Sporting Profile and Implications on Housing-Decisions

The sporting profile of Jeppe Boys High is variedly distributed. The sports data for average scores were obtained from SA School Sports (2020) as follows, rugby (22.97% at 6th ranking

position provincially and 19th position nationally), hockey (22.78% at the 3rd ranking position provincially) and water polo (17.65% at 21st ranking position provincially). This data was based on constant values (average rate performances) which led to the model not computing their coefficients due to the data not being evenly distributed and recognised as invalid. Therefore, considering Jeppe Boys ranking profile from 2010 to 2020, their level of sporting competency is ordinarily average. Meaning the value-added by sports at the school does not project any effect on property sale prices, housing-decisions or have influence on the inclination of households to enrol their children in Jeppe Boys High and reside within that same school district. The relationship of these sporting variables with sale prices is weak or does not exist. Only the cricket variable provided valid coefficient results indicating a negative correlation of *p*-value= - 0,001 whilst other sporting profiles (rugby, water polo and hockey) were omitted to avoid issues of high multicollinearity during modeling. The cricket coefficient is significantly distant from 1, substantiating to the weak significance and correlation between the univariates. Meaning, every unit point decrease in the cricket average score will result in a 0.1% decrease in selling price. Arguably, it should be noted that sports is not a determinant of housing demand in school districts hence its capitalisation effect on a property's selling price is very minimal. The inclusion of sporting profiles in the quality of a school was purposive when giving reference to the economic agents and the economic theory of consumer's preferential patterns and behaviours, although no significant coefficients could be computed.

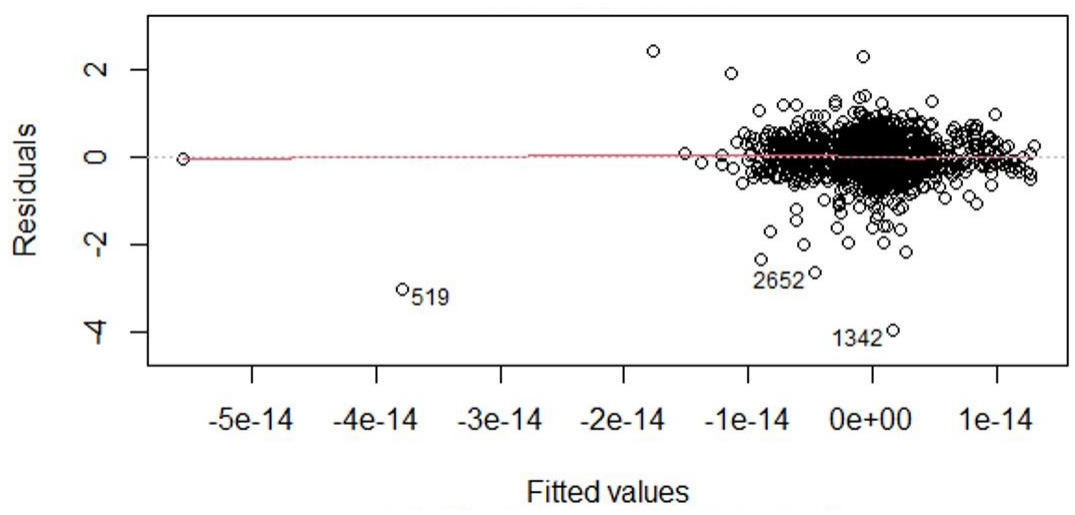
4.5.5. Capitalization of Property Structural Factors in Sale Price Premiums

The predictors regression coefficients, except the number of bedrooms (= 0.008) are statistically significant at the 1% significance level. Table 1 shows the positive coefficients for structural attributes, namely, the 'erf size' (lnSize = 0.275, p = < 0.01) and bathrooms (coefficient value = 0.068, p = 0.01). Authors such as Gibbons and Machin (2007), Ziets (2008), Fen et al. (2010), and Wild (2016) state that an increase in the square metre size of a property's land results in an increase in that property's price. Meaning that a square metre increase in the erf size results to a 27.5% increase in sale value in both areas of Kensington and Observatory. The bathroom coefficient value is given as 0.068, meaning that a unit increase in bathrooms will lead to a 6.8% rise in purchase prices. In addendum, the insignificance of bedrooms could also be reasoned by household's unwillingness to pay a high premium to have more rooms as compared to having a bigger lot size. The relationship between the purchase price and the number of bedrooms is weak and insignificant, thereby indicating that the variable does not contribute to the model, as per the given data.

4.5.6. Diagnostic tests for Multicollinearity and Heteroscedasticity

Remedial actions such as log transformation (and the isolation of univariates between models) were adopted to produce good residuals, which frequently improved the models beyond solving heteroscedasticity. This allowed the researcher to find different approaches to specify the model to reduce the error variance. The research uses the Variance-of-Influence (VIF) to identify severe problems of multicollinearity and heteroscedasticity. The VIFs of all the explanatory variables are less than 5, and none of the estimated parameters are extreme, which can be considered that the risk for multicollinearity problems is considerably low. This implies that the effect of multicollinearity is effectively cured and under control. The regression results prove that there were no signs of heteroscedasticity being problematic in the model, which is furtherly substantiated by the low VIFs the model produced. The standard error terms are smaller in the model because of the large sample size used in the study. In addendum, they were no noticeable irregularities in the data. This indication implies that the regression coefficients are stable. Therefore, confirming that the hypothesis testing in not suspect, further denoting that their statistical significance is highly reliable.

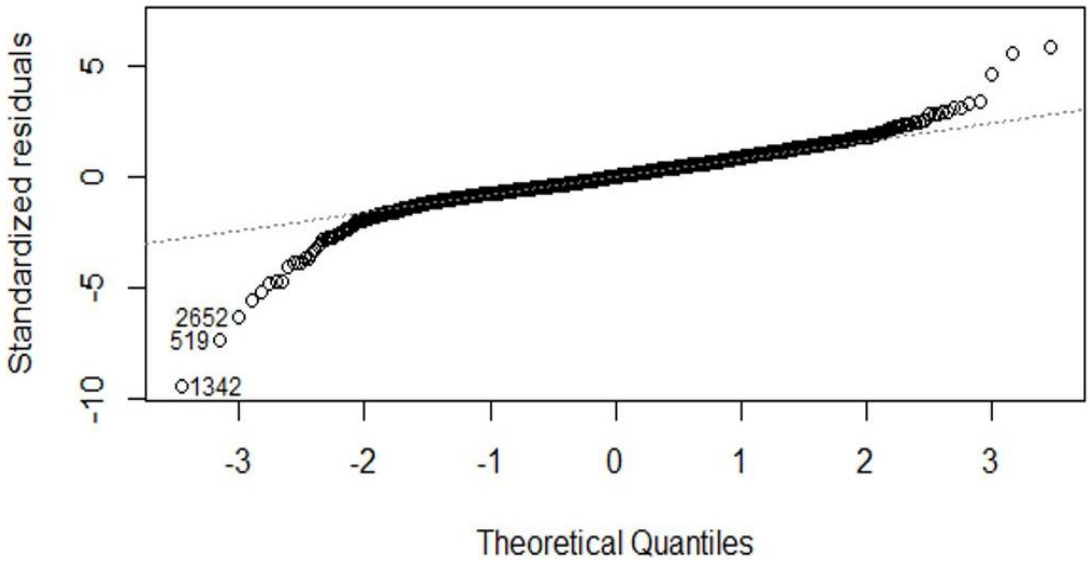
Graph 3: Plotted Residuals (y-ŷ) against Predicted Values (ŷ) for Model I.



The study proceeded to provide substantial and empirical evidence of the models robustness by checking the line of fit of regression (see Graph 3). An examination of the residuals is a predominant technique by which the fitness of the model is checked and it is best presented graphically (Curwin, et al., 2013: pp, 448-449). Generally, fitted values are not correlated with the residuals and are better than observed or unobserved values (De Haan, et al., 2013: pp, 58-62). Graph 3 indicates that the plotted points exhibit a random scatter over the plane and shows no pattern. The residuals are a sample from a normal distribution with several outliers.

The use of regression modeling aided in reducing sensitivities to outliers, informing us how residential properties are clustered across space in both Kensington and Observatory. Graph 3 and 4 display the fit of regression for model I. The residual standard error is 0.419 on 1876 degrees of freedom of which 1184 observations were automatically excluded by the models due to missing data for certain variables. In addendum, the error terms ensured that the number of variables is true to the assumptions of the regression modeling. Graph 3 illustrates that the errors are not related to the predicted ŷ, which is very useful to prove that the adopted models are correct and robust according to the regression equations used during modeling.

Graph 4: Normal Q-Q Distribution of Model I.



Model I included the structural attributes (erf size, number of bedrooms and bathrooms) of all valid property types (SS, EST SS and FH) in both sample areas and the school quality (including sports average performance rates) of Jeppe Boys High. The multiple $R^2 = 0.2628$, meaning that the explanatory variables explain 26.28% of the variation amongst observed values of the property sale price. In addition, the adjusted R^2 is 0.2608 for model I meaning that 26.08% of the variation in the purchasing price is explained by the explanatory variables listed in model I.

4.6. The Hedonic Regression Method - Model II

The hedonic method can be employed in the construction of transacted property prices by using two traditional models, namely the cross-sectional hedonic model and explicit-time hedonic model with the latter regarded as being more robust. This section proceeds to explore

and estimate the coefficient results using the HPM econometric technique. Considering the results, school quality and the year of deeds registration of each property showed an extremely high correlation, which deducing from prior studies can generate problems of multicollinearity. To avoid omitted variable bias and regression coefficient results being unstable when the standard errors increase, both variables were used separately in the two models to allow a comprehensive analysis of the results. Hence, the traditional hedonic model is reconstructed in model II to determine the change in the adjusted R^2 and show the variation in sale prices explained by the models.

Model II is used and described below as follows:

$$LnSPrice = b_0 + b_1 \text{ lnSize} + b_2 \text{ Bedroom} + b_3 \text{ Bathroom} + b_4\ 2011 + b_5\ 2012 + b_6\ 2013 + b_7\ 2014 + b_8\ 2015 + b_9\ 2016 + b_{10}\ 2017 + b_{11}\ 2018 + b_{12}\ 2019 + b_{13}\ 2020 + \boldsymbol{\varepsilon}_n$$

(Equation 5)

Where:

- lnSPrice in both model I and model II denotes the transacted price in a log functional form;

- β_0 is the intercept that represents the structural attributes of the property, which includes β_1 which is the erf size, β_2 is the number of bedrooms and β_3 denotes the number of bathrooms;

- $\beta_4 \ldots + \beta_n$ is a matrix of the time dummy variables which, represents the property registered sales between 2010 and 2020 (10-year period) in both sampled areas; and

- $\boldsymbol{\varepsilon}_n$ is the transaction specific random error term, used to show the deviation of an observed y value from the sample regression line. The employed formula is represented as error = (y-ŷ) (Utts, et al, 2007: pp. 162-163).

From Equation 4, the transacted property price index is derived from the estimated time-dummy variable coefficients. The initial time-dummy variable, '2010', represents the base-year and, hence, it was omitted from the regression equation above (see Table 5). As a result, there was no base-period coefficient estimated for it because the collective inclusion of it along with β_0 would have produced perfect collinearity. The base period variable is generally represented by the zero value and the aggregated price index begins after this initial time-dummy variable. The implicit prices of locational and physical property attributes are assumed to be constant

overtime by the hedonic model, which might be subjective, for instance, the presence of three bedrooms might have more of an impact on transacted property price in 2019 or 2020 than the influence it had prior in 2010 or 2011.

4.6.1. Econometric Applications and Analysis of the Results

This section discusses the hedonic results for Model II obtained from the empirical modeling. The second hedonic regression model is a direct estimate of Equation 5. To ensure validity of the coefficient estimates using the hedonic price model and avoid bias results other regression experiments were undertaken (also see graph 5 of this model validation). Overall, this tests assisted in extrapolating the change in the capitalisation rates of property price determinants (and the adjusted R^2) when statistically insignificant variables are identified. The researcher excluded the pass rate (school quality) and school sports (represented by cricket as the only statistically valid and most dominating sporting activity at Jeppe Boys High, amongst rugby, water polo and hockey) from model II. These variables were only included in model I. The level of variance explained by these explanatory variables when capitalised into the purchasing price of properties is observed as a shortfall in the study (the expected hypothesized coefficient signs of both variables was positive). The coefficient values show signs of multicollinearity when included in the model hence the use of a regression diagnostic test of the results is highly imperative. Seo and Simons (2009) argue that diagnostic results purposively aid in detecting whether the models are properly defined to test the normality based on the HPM residuals. Hence, the research approach elected to separate variables that had a problem of multicollinearity, as their contribution in the regressions was deemed invaluable. Undertaking such statistical processes also aids in exploring the difference in the relationship between other predictor variables (lnsize, year, no. of bedrooms and bathrooms) and the purchasing price of properties in Observatory and Kensington.

4.6.2. Estimation Results and Associated Discussions

Table 5 exhibits the coefficient regression results of model II. In model I, the adjusted R2 value is 0.261. Evidently, as also argued by Wen et al. (2019), the presence of kindergartens, primary and junior schools, colleges and other locational attributes (i.e. accessibility such proximity to retail centres and employment areas) and neighbourhood amenities as seen in and around Observatory and Kensington may explain the influence of parents willingness to reside in these areas and not Jeppe Boys High. Hence, locations present both costs and benefits that are incurred as an effect of being in close-proximity to neighbourhood amenities and dis-amenities (Hummel, 2017). The adjusted R2 of model II is equal to 0.276, indicating

a slight difference of 0.015 from the adjusted R2 value in model I (with all the predictor variables included, except the year due to multicollinearity). Meaning that, 27.6% of the variation in the purchasing price is explained by the predictor variables listed in model II. In addition, this indicates that both models (I and I) can explain more than 20% of the variance of the sale prices in both local sample areas.

Table 5: Model II - Hedonic Price Model Estimates.

Variables	**Coefficient Value**	**Standard Error**	**VIF**
Intercept	11.929***	0.096	
Log of Erf size (lnSize)	0.275***	-0.012	1.069
No. of Bedrooms	0.009	0.011	1.562
No. of Bathrooms	0.062***	0.011	1.544
Year Attribute (dummy variables)			
2011	-0.130**	-0.069	
2012	-0.033	0.066	
2013	-0.034	0.066	
2014	0.071	0.064	
2015	0.134***	0.064	
2016	0.105**	0.063	
2017	0.137***	0.062	
2018	0.192***	0.062	
2019	0.181***	0.062	
2020	0.179***	0.065	
Year Dummy Variable (2010-2020)			1.034
Adjusted R^2		0.276	

Notes: *** and ** indicates statistical significance at 1% and 5%, respectively.

: 2010 is used as the base-year (dummy variable) in modeling the results

Essentially, both models may need more important determinants or variables to be included in the regressions to ensure a substantial variation in purchase prices. As meticulously highlighted in the literature review, this process can be very beneficial because homebuyers are more prone to 'vote with their feet' (Yadavalli, et al. 2014; and Wild, 2016) when making residential and locational decisions in specific neighbourhoods, which gives reference to Tiebout's model. Hence, the research refrains from providing inconclusive assumptions similar to prior studies without an appropriated a thorough examination of the factors driving demand and supply in region F. In response to this underlying issue deterring the reliability of most studies, Gibbons (2011) the hedonic framework makes a provision that the use of multiple-demand-and-supply equations can effectively represent the varied cost structures of home-sellers and the heterogeneous preferences of home-buyers.

4.6.3. The Implicit-Value Linked with JHSB School Quality

In Table 2, the school quality results show that Jeppe Boys High maintains an average pass rate of 97.4%, which might be of a lower quality than other nearby high schools such as Parktown Boys High, Hirsch Lyons Boys High, King Edward VII School and Athlone Boys High. To add more validity and reliability, average ratings of sporting performances were included in the school quality proxy. However, only cricket was statistically significant whilst other sports (hockey, rugby and water polo) had constant values, which were omitted because they neither retracted nor added to the goodness of fit of the regression models. According to Bayer, McMillan and Ferreira (2007) not all educational attributes appear to be determinants of school quality. They further accentuate that few households deem certain school characteristics such as changes in student body characteristics or school expenditure as a property indicator and not necessarily sporting performance, which could explain the above result. Therefore, the research concluded that since no results could be obtained for other sports when regressed with school quality this means that no relationship exists between the variables. The SA School Sports portal exhibits the rates of quality of each sport from 2010 to 2020 against other schools on a national, provincial and local competitive scale, which is where the data was obtained. Overall, Jeppe Boys High's quality ranking could explain the lack of significance of sports in aiding households with decision-making in purchasing a property in close-proximity to the school.

Meaningless to say, few parents base their housing-decisions on which schools maintain a good overall pass rate and the best performances in sports (including quality sportsgrounds) to enrol them in such educational institutes. Moreover, the structural quality or condition of the educational facilities at the school may also be substandard when compared to other

surrounding high schools, justifying the lack of significance on the sale prices of properties. Thereby, indicating that demand and housing-decisions based on locational attributes varies per household and these sampled areas are an epitome of such results. Similar to Jiao and Liu (2010), the price semi-elasticity results of kindergartens, primary schools and varsities to housing prices is generally higher than for secondary schools. Therefore, parents could also be more inclined to pay substantial premiums for a property in these areas due to good location, accessibility and proximity to the city centre, employment opportunities and other neighbourhood attractions instead of basing their preferences on the quality of Jeppe Boys High School.

4.6.4. The Direct Capitalization of Structural Attributes on Purchase Sale Prices

A myriad of structural attributes were examined using the multi-variate regression approach. Results indicate that only the logarithmic erf size and the number of bathrooms are significant at the 1% significance level. The coefficient of erf size (lnSize) in model II is equal to 0.275 with a difference in change of value of 0.001 from model I results of lnSize (0.274). Therefore, a marginal unit increase in price of a square footage is linked with an average of 27.45% increase in the property selling price for both sample areas. In addition, both areas are well-located within the City of Johannesburg, meaning land values tend to increase with close distance to the CBD. However, the equidistance from Jeppe Boys High and its surrounding suburbs to CoJ is approximately 18.7km to 21km and considering their a relative average sale prices, and it can be concluded that for property values in these areas are fairly balanced (in terms of affordability). For example, in Johannesburg (central Sandton) a square metre of land is approximately R39 600 (Businesstech, 2018), meaning the premium value of land and its associated property price should be lower with distant from a metropolitan city or CBD.

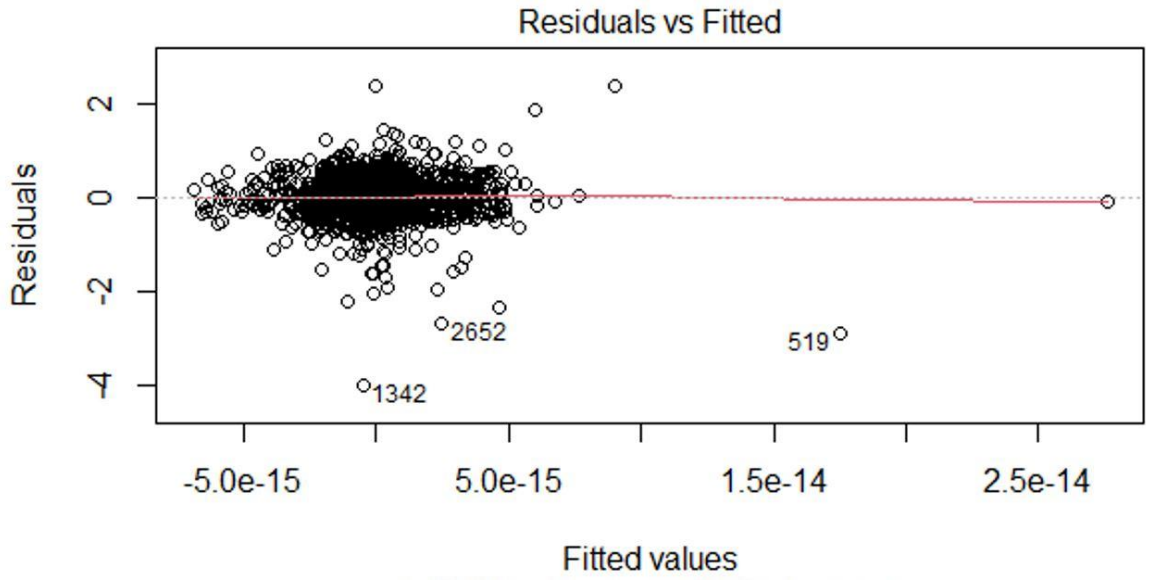
Descriptive statistics show that the most expensive properties are the Est FH with an average price = R2 162 711.83 and highest selling price = R 5 600 000 which are mostly located in Kensington. The least expensive property type in both areas is the Est SS with an average price = R704 545.50 and the highest price = R975 000) with most of these property located in Observatory. Generally, based on the economic theory structural attributes of a property and their associated value should be significantly correlated due to their positive contributing nature as they form part-and-parcel of a property's price determinants. As put forward by Hummel (2017) (adopted from Sheppard, 1999), 'the demand for such housing goods are based on a myriad of attributes the houses embody. In addition, each attribute 'bears the utility to the consumer' and that enables a price for each attribute of the property (Wildeboer, et al.

2017). In table 5 (model II), a unit increase in the number bathrooms will increase the property value by 6.22%.

There is a mixture of both statistically insignificant and significant univariate coefficients in the results. In the HPM, only the year periods from 2012 to 2014 indicate insignificant negative results. However, 2011 exhibits a negative coefficient but it is statistically significant at the 5% significance level. The coefficients from 2015 to 2020 are all positive and significant at the 1% significance level in the model results, with the highest value seen in 2018 and the lowest in 2012. In the latter cases, it can be explained that a unit increase in the year in which a property was registered will lead to 19.2% increase in the purchasing price and a decrease of 3.3%, respectively. The results for the VIF diagnostic test indicates that the coefficient values of all the explanatory attributes are below 5, meaning that multicollinearity in the model is considerably low. This includes the year in which each property was registered (VIF=1.034). In terms of the years, 2010 was used as the base (reference) year as a required econometrics prerequisite.

4.6.5. Transformation of the Selected Subset Explanatory Variates – Model Robustness

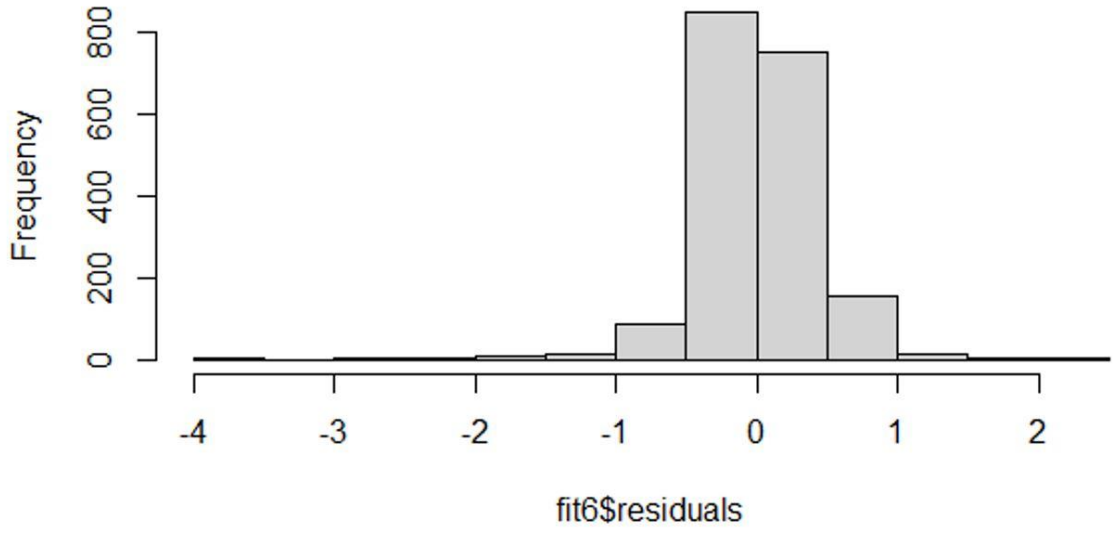
Graph 5: Plotted Residuals (y-ŷ) against Fitted Values (ŷ) for Model II.



Graph 5 and 6 exhibit the fitness of regression for model II. The residual standard error is 0.414 on 1868 degrees of freedom of which 1184 observations were excluded by the model. Model II included the year each valid property was registered, structural attributes (erf size, number of bedrooms and bathrooms) in both sample areas in close-proximity to Jeppe Boys

High. The multiple R^2 = 0.2808, meaning that the explanatory variables explain 28.08% of the variation amongst observed values of the dependent variable. The adjusted R^2 is given as 0.2758 denoting that 27.58% of the level of impact of the independent variables on property purchase prices is explained by model II. The separation of certain explanatory univariates between the two models was to test the difference in the variation of the adjusted R^2 coefficients and the strength of significant and insignificant property determinants. The fitness of the coefficient regressions is shown by the 0.20% increase in the transaction price of a property in Kensington and Observatory is explained variables in the model.

Graph 6: Histogram Plot of Fitted vs Residuals for Model II.



In summation, model II proves to provide robust output results when school quality is isolated from the model substituted by the year each property was registered. However, it should be noted that the age of the property is a prerequisite in the hedonic price model but due to not being able to obtain that data, it became a shortfall. Arguably, prior studies have flagged the use of the 'age' of a residential property because in their HPM and OLS regressions it statistically decreases the selling price by a marginal percentage. On another note, Graph 5 demonstrates that the inferences for regression of model II are robust. Meaning that there are no major outliers or extreme skewness, the regression model works well even though the distribution of the y values are approximately normally distributed. There are not many deviations from the regression line. The conditions of the fit of regression for the preferred model II have been met provided the following coefficient results (see table 5, and SASchoolSports, 2020).

4.7. Conclusion

This chapter has examined a matrix of determinants in the construction of the transacted sales for the Kensington and Observatory local property markets. The purpose of this present study was to investigate the quality of model C schools in region F and their impact on market selling prices in the CoJ. By establishing a suitable proxy of dummy school quality determinants, it was possible to deduce whether proximity and quality of the Jeppe Boys High school spills over into the selling prices of the neighbouring properties. The premise to examine the quintile 4 model C school as the key variable in this study was to advance knowledge in the body of hedonic literature, particularly in the South African Urban Economics, where there is an extreme paucity of studies on this topic. Hence, this study's findings and contributions can be deemed credible to be used as an inference for analysis and a measure of the condition of local property markets, and the underlying unobserved dynamics behind property trends in these areas.

Two different models were used in the estimation of the underlying parameters by employing the hedonic regression approach. The models did not detect any issues of spatial dependence or autocorrelation, which proves the fit of the regression coefficients. As noted in Wen et al. (2018) study, these denoted hedonic issues changed the direction of their omitted predictor variables, which caused high-levels of multicollinearity. In this thesis, the techniques employed as treatments of error terms in the models included separating the hedonic determinants when two or more explanatory variables depicted signs of high correlation. This helped the researcher in specifying the functional form of the hedonic models to minimise the error variance. The variance of influences of the independent variables were all below 5. This infers that the possible bearings of multicollinearity were effectively treated, further proving there were no indications of heteroscedasticity being an issue in biasing the estimated regression coefficients. The actioned processes improved the models beyond solving these hedonic issues, further helping in producing good and reliable residuals.

The estimated regression coefficient representing the price semi-elasticity of the school quality of Jeppe Boys High is shown as 0.028 (statistically significant at the $p = 0.01$ level). Thus, this school influences transacted property prices significantly, which implies that a rise in the school quality variable will improve property sale prices in both areas by approximately 2.8%. Comparatively, other similar empirical studies produced regression coefficients that were only based on output measures, used as dummy variables to estimate the true value added by the presence of a particular school (also see, Rosenthal, 2003; Davidoff, et al., 2007; Clapp, et al., 2008; Nguyen-Hoang et al. 2011; Wen, et al., 2018). For example, John et al. (2010) study

also confirmed that a 10% rise by only using test scores (GCSE) will increase housing prices by 1-3%. As a counter-argument, there is a significant variation between using a matrix of output and input measures to accommodate for all the unobserved variables, thus, both measures of a schools quality should be equally valued. In this study, redefining the input and output measures of school quality and constructing a more enhanced, location-specific, and suitable proxy using determinants of a model C school and accurately estimating their regression coefficients makes this paper a novelty in the SA Urban Economics Research cluster.

Similarly, Zheng, et al., (2016) study argued that improvements in the quality of secondary schools generally lead to a rise in property prices, and are more valued by parents than other schools. In this thesis, evidence shows that this relatively low school quality coefficient of Jeppe Boys High is explained by the presence of other good quality schools (i.e. kindergartens, primary and high schools) within this same school district and other neighbourhood amenities. Thereby, confirming the hedonic theory arguing that economic agents take housing-decisions by making conversant trade-offs between property attributes and local amenities (Senick, 2019). This school of thought acknowledges Tiebout's (1956) standard urban economic theory.

Another notable discrepancy identified between the results in this thesis and that of previous empirical studies is that the hedonic price model is, somewhat, non-linear. Subject to the research problem statement in this study, Chiodo, et al. (2010) study stated that having a linear assumption of test-scores might overestimate the premium offered by general schools or underestimate that offered by high-quality schools. Notably, this assumption is null in the research as shown by the nature of the data, adjusted R^2 and the capitalisation rate of school quality on property sale prices. The shape and slope of this correlation is fixed to the relative number of home-sellers and homebuyers in the local property transaction market of the study areas. The identified issues of non-linearity are always evident in household budgets and can cause measurement errors, which can deter the regression coefficients if ignored as argued in a study conducted by Hummel (2017). Thus, the employed method of treatment in the research is achieved by using a multiple market dataset. The practical consideration behind this is that market forces determine the direction and pace of such dynamic shifts. Overall, the results are buttressed by similar hedonic studies and their related deliberations are fairly situated within the body of current existing literature. As highlighted, the hedonic price models provided a good fit for the dataset used in the regression of coefficients.

5. Chapter 5: Summary and Conclusions

5.1. Introduction

This section provides a summated synopsis of the key empirical findings in relation to the research objectives and problem statement set out in chapter 1. This is supported with an outline of the possible discrepancies with previous empirical studies. The chapter also discusses the limitations of the study and the potential direction for future research. This is followed with a discussion of the research contributions and policy-implications. The last section closes the chapter.

5.2. Summary of the Research and Findings

The research examined the effects of quality schools (section 21 schools) and other unobserved fundamentals on property sale prices within the CoJ (region F) market. Evidently, this investigation confirmed that the school quality of Jeppe Boys High (a former model C school) contributes to the significant variation in the movement of sale transactions in Kensington and Observatory. As highlighted, the findings are linked to the problem statement being addressed in the research and the underlying research objectives. A body of existing hedonic literature significantly buttresses the findings.

5.2.1. Effects of School Quality and Other Property-Related Factors on Sale Prices

The quality of Jeppe Boys High School effects purchase prices significantly, and the price semi-elasticity of the regression coefficient is equal to 0.028 (statistically significant at the p = 0.01 level). This implies that each upgrading in the school quality (pass rate) will result to a 2.8% increase in property transaction values in both areas. Evidently, this can be reasoned by the school district these sample areas are located within, which is comprised of numerous competitive quality high schools including kindergartens, primary schools and colleges. Backtracking to a similarly relevant prior study, the findings reveal that substantial differences exist in the effects of school qualities on property values across the conditional distribution. Differences in the quality of educational facilities will lead to variations in property selling prices (Bae, et al. 2013; and Zheng, et al., 2016). Therefore, the research highlights that school districts in region F (CoJ) have a heterogeneous composite of schools (kindergarten, primary and secondary schools, and colleges and universities) that are in coexistence and exhibit huge variations due to different preferences of local public goods. The general understanding extrapolated from these findings is that in equilibrium the marginal value of improving, for

example, any determinant of a property for a provided consumer's income and preferences will be counterbalanced by the utility expenditure of the surplus cost included.

Generally, the higher the desirability of a property's structural attributes in contrast to other properties (Wen, et al., 2017), the higher the sale price that property reflects these determinants (in monetary value). However, in both model I and model II, the relationship between structural attributes of a property and the associated implicit price was moderately weak, but statistically significant. LnErfsize had the highest and strongest relationship with property sale prices, which explains over 27% of the variation in the selling price. The coefficient results for the number of bathrooms and bedrooms in both models contributed a capitalisation rate below 10%. Meaning each unit increase in the number of bedrooms and bathroom does not add or subtract much to the variance in the housing prices. Moreover, the willingness of parents to pay higher premiums for an additional structural attribute could be less due to the difference in household preferences of property determinants that are more beneficial to maximize their opportunity costs and influence their housing decisions.

The correlation coefficient for model II tested whether the year each property was registered had any marginal effect on the purchase price. Both positive and negative coefficient results were obtained exhibiting a weak relationship, the years between 2011-2013 show a negative correlation and the period from 2014 to 2020 show a positive relationship. The year in which the properties were registered shows a very strong correlation with pass rate, therefore, exhibiting problems of multicollinearity. Generally, this increases the standard errors of these variable coefficients when the model includes multiple univariates that are not merely just correlated to the response variable, but also to each other (Haizhen, et al., 2019). For example, a high correlation between the year and school quality denotes that the regression coefficients are unstable when the standard errors increase.

The VIF for all the factors in both model I and model II indicate that regression coefficients are adequately estimated since the values are not near or above 5, which requires a relatively simple solution. To avoid possible multicollinearity issues, highly correlated predictors have to be removed and regressed separately. This is a well-known issue, particularly in hedonic regressions (Wen, et al., 2019), which has persistently brought scepticism and lack of reliability in the statistical findings and models adopted by prior studies. In the research, using all variables with the year as a dummy variable, the coefficient results for school quality and other sporting predictors supplementing quality do not converge. Arguably, prior studies have flagged the use of the 'age' of a residential property because in their hedonic regression models and OLS regressions it statistically decreases the selling price by a marginal percentage. As with this study, the problem of multicollinearity is not exaggerated merely

because it does not focus on coefficients of the parameters but more interested in the values utilised in estimating the quality-adjusted sale price index (see Annexure C). Undertaking these steps is beneficial if removing one of the highly correlated predictors does not drastically reduce the R^2 result.

5.2.2. Contextualizing Research Discrepancies with Previous Empirical Studies

In the research, the employed econometric methods of investigation were specifically selected due to their significant robust-nature in producing statistically unbiased and reliable coefficient regressions (depending if adequate data is used), in contrast to prior literature that experienced several statistical anomalies in their regression models. Although few hedonic regression modeling challenges still linger in some modern studies, it is of paramount significance that researchers adopt regression methods that are best-fit for the nature of the obtained or sourced secondary data to avoid inconclusive literature findings that question the integrity of models. The research has managed to overcome these relenting heirlooms extensively highlighted in previous similar studies attempting to adequately untangle and respond to this problem statement. This particularly gives reference to issues of omitted variable bias, heteroscedasticity and multicollinearity in the hedonic regression models, which is subject to the type of data and models employed and the underlying objectives as highlighted in the research.

Evidently so, issues of spatial dependence or autocorrelation were not detected in the models, which proves the fit of the regressions considering the results. Spatial autocorrelation also occurs where either similar values (high or low) or different values (high and low) of a certain attribute are likely to group in space (Cheruiyot, et al., 2015). To an extent, this limitation seldom changes a researches direction as noted in Wen et al. (2018) study through their omissions of predictor variables causing high levels of multicollinearity. In the study, remedial actions were promptly undertaken such as isolating the univariates during modeling if two or more explanatory attributes have a high level of correlation. Employing these techniques enabled the researcher to specify the model accordingly to reduce the error variance. Substantially, the VIFs of the explanatory attributes were all below 5 inferring that the plausible impacts of multicollinearity were effectively cured and under control, further proving there were no signs of heteroscedasticity being problematic in biasing the regression results. These actions frequently aided in improving the models beyond solving these labelled issues associated with hedonic modelled regressions. Hence the adopted models helped in producing good residuals (see Graph 3, 4 and 5).

To ensure robust models and estimations these possible spatial effects need to be accounted for during modeling regressions. This is necessary because property transaction prices are depended on one another and residing in close-proximity to high quality homes can, solely, increase the property price of a homeowner. Considering that if this impact occurs and spatial lagging of the property sale price is disregarded, the estimated parameters will presumably be bias-laden (Wen, et al., 2018: p. 74). Moreover, larger standard errors (and subsequently smaller t-statistics) could be created if spatial-correlation in the error terms is overlooked. These results also prove the existence of a positive autocorrelation between property sale prices by using hedonic modeling to ensure reliability in explaining the relationship between characteristic variables and sale prices.

In previous literature, this is still an unresolved matter in regard to the impacts of school quality on residential property values still remains in literature, also considering that school characteristics are deemed as unobservable (Kane, et al., 2005). Therefore, to avoid such persistent issues it is suggestive to consider using best subsets regressions, step-wise regression or specialized knowledge to remove these variables from the data set. The use of the Principal Component Analysis (PCA) or Partial Least Square Regressions (PLA) can also be beneficial in redressing issues of overinflated standard errors that make some predictors statistically insignificant (Heckard, 2007). These forms of resources and programs could serve as useful strategic solutions to potential regression modeling failures in estimating the values associated with environmental, social and economic attributes that can explain a variation in property sale prices. It is undeniably transparent that all these suggestive recommendations outlined in this chapter can be applicable and effectively realisable through ensuring adherence to the requirements of each regression model, as proven in the research. This can be valuable if there is synchronicity and coherence of the issues identified, objectives, nature of the data used to specify the best-fit model in responding to the topic under investigation, which can plausibly help in providing robust and reliable coefficient results as shown in the research.

Conversely, Zheng et al. (2016) study argued that secondary schools were considerably the most valued school type by households of lower-and-higher priced properties, whereas kindergartens and primary schools show the least profound effects. Although the value of educational quality may be very high in high-quality districts, in the research homebuyers in the sampled area’s market are probably more inclined to pay a higher premium for additional improvements in the academic performance of their children. Comparatively, the price linked with an upgrading in educational quality could presumably be lower in market areas where the quality of the school is sub-standard. This is relative to the finding brought forward in the research arguing that several home-sellers and homebuyers in region F’s transaction market

who may not have children are most likely to place little value on the educational quality in that area of the market. Hence, the hedonic price model is somewhat highly viewed as being non-linear (Chiodo, et al., 2010). Subject to the research problem statement in the study, having a linear assumption of test-scores can potentially overestimate the premium offered by general schools or underestimate that offered by high-quality schools. Notably, this assumption is null in the research as shown by the nature of the data, adjusted R^2 and the capitalisation rate of school quality on property sale prices. The shape and slope of this correlation is fixed to the relative number of buyers (i.e. consumers), home-sellers and property developers (i.e. suppliers) in and around the property transaction market of Kensington and Observatory. The identified issues of non-linearity are always evident in household budgets and can cause measurement errors, which can deter the estimations if not accounted for (Hummel, 2017). Therefore, the proposed method of prevention in the research, for example, in a case in which non-linearity signifies endogenous determination of characteristic prices, is realized through using multiple market data. This is reasoned by the fact that market forces determine the direction and pace of such dynamic shifts.

5.3. Research Contributions

With regards to the South African market, one of the important contributions of the research is the use of region F's (Kensington and Observatory as sampled areas) data at the local level to uncover insight on the influences of school quality, demand factors in addition to other fundamentals on property sale prices. In most previous studies on school quality, hedonic issues such as omitted variable bias, spatial autoregression, spatial correlation, multicollinearity and heteroscedasticity remain an unresolved challenge as highlighted. In the research, these issues were cured and under control. The advantage of the research is using two accurate hedonic functional forms informed by the dataset of a local level, which had a sufficient number of observations and fundamental variables (physical and locational attributes). The study compares the accuracy by using hedonic models and undertaking several diagnostic experiments to ensure reliability in order to produce robust results. Using the goodness-of-fit reported by the models is important, however, other applications of comparing the accuracy of various methods is necessary, particularly for future research studies.

As broadly described in chapter 4 and in the review of literature, neighbourhood attributes and local public goods are key determinants of localised sale price variations in Kensington and Observatory. Gibbons et al. (2014) highlight that the implicit sale prices of properties provides shadow prices of non-market goods, which has been immensely recognised in environmental

economics for their contribution into cost-benefit analyses. Arguably, the use of such findings from hedonic regression models to represent the true value of local environmental goods have been overly put into question. Nonetheless, the research has illustrated through the findings that reliable, stable and meaningful capitalization averages of spatial goods such as school quality can explain a sufficient variation in region F's real estate sector (particularly the residential market). This has purposively augmented the widespread gap-in-understanding and doubtfulness in the robustness of hedonic price models to effectively estimate the capitalization effect of environmental amenities (i.e. quality schools) in property sale prices.

Undertaking the research has also helped supplement the minimal responses to the topic in the environmental economics field. Notably, one of the gargantuan obstacles in the SA environment at least, can be reasoned by the lack of resources to avail time-varying details of geographical information on local amenities (such as transport noise, crime rates and the distance to neighbourhood attributes) which would have furtherly strengthened the research. Therefore, future research that focuses on the causal effect behind property price variations can also draw on the patterns of localised policy changes and presumably prove to be very fruitful. Considering such forms of actions is necessary, particularly when empirical challenges of pricing neighbourhood goods without adequate data persists.

5.4. Practical Directions and Policy Implications

It is of paramount significance for property developers, investors, homeowners, home-sellers, policymakers and the private and public sectors to further peruse 'the unknown' in estimating the associated costs of the observed and unobserved fundamentals on property asking prices. This is purposed to determine the unaccounted cause of variation of macroeconomic and microeconomic determinants that can be factored into property sale prices or values. When policymakers ponder investments for educational improvements, a cost-benefit analysis is necessary. Therefore, for policymakers to understand the interactions between these factors and property sale prices can potentially provide evidence-based information for policy formulation that effects activities in property markets. For example, this gives reference to the way in which property markets respond to the school quality indicators used that may cause property sale variations. Since the research has established the influence of these predictor variables, all the relevant stakeholders such as investors will be informed about the fundamental determinants contributing to the variation of sale transactions in both sampled areas.

Based on the knowledge uncovered, property policymakers can impose authoritarian adherence of valuating regulations, strategically planning and developing housing

environments, and provide adequate infrastructures for neighbourhoods that can potentially attract higher property asking prices and stimulate economic growth within the metropolis. In that same vain, valuers and real estate surveyors can develop models that enables them to isolate univariates that have high-sensitivity issues of spatial correlation and spatial autoregression in study areas based on the data of neighbourhood attributes and property statuses. For instance, more structural models can be applied such as the vector error correction (VECM) and the vector autoregression (VAR) model to assist in examining the joint dynamic behaviours of property value variables (Sims, 1980; and Phillips and Ouliaris, 1990). Generally, these methods are prevalent in modeling time-series datasets. The objective behind these recommendations is to contribute to the prevalent gap in knowledge and ensure reliable responses to this subject matter.

Drawing from Haizan et al. (2013) study, such endeavours can provide a theoretical and practical significance and plausibly can be used as a reference point for the articulation of some education policies. This can also provide a backdrop for real estate practitioners that will enable them to find a solid correlation between the findings in the research, analysis of the problematized issue, the adopted hedonic price models and their implementation in property valuation processes. A widespread length of prior literatures demonstrate uncertainties on which appropriate regression techniques to employ (Wen, et al., 2019), inconspicuous that it is highly determined by the nature of the data being used and the underlying objectives. Hence, confusion persists as evidently observed from scholarly debates and arguments questioning the reliability of both the HP and OLS models, as meticulously articulated in the literature review chapter. Although several authors are still attempting to adequately estimate the variation caused by school quality in property price changes, the research is satisfactorily buttressed and situated within existing literature by effectively untangling the effect of school quality (a former model C school) on sale prices instead. The hedonic price model provides a fairly good fit for data utilised in the research, as both models explained over an estimated 26% of the variation in the log of the property sale prices. Hence, the research is considerably beneficial for future decision-makers with the necessary knowledge to elude probable model implementation failures or biasness and standard errors in their regression results.

The empirical discussions demonstrates that government-policy relatively correlates to the hedonic findings of locational amenities such as school quality, which sits amongst the significant area of government policy-design targets and application. This includes other unobserved neighbourhood attributes such as property-related crimes and transport noise, and other improved transport attributes (accessibility) which are significantly capitalised into the values or selling price of properties. In that same view, Machin (2011) extents that there are other useful information on the values attached to local amenities that have proven to be

policy-relevant because of the variation effect observed in these values across places and people (social capital value). To support achieving the abovementioned, it is recommended that policy be directed towards establishing a database for property transactions to enable future studies to conduct research using useful and reliable information. In addendum, by including more relevant variables, future studies can broaden the knowledge spectrum and provide improved and robust modeling techniques and credible results.

5.5. Limitations of the Research

The researcher initially obtained 3066 property sales that occurred between the periods of 2010 and 2020. The study identified 1177 properties without structural variables (missing values) from the raw dataset, for instance, the number of bathrooms and bedrooms. The process of screening for normality, wrong entries, missing values, homoscedasticity and linearity resulted in a total of 1889 valid property observations that were retained for analysis. The process was purposed to avoid potential errors from the SPSS software and R-Studio program during regressions. It further assisted in determining the distribution of the data, variable correlations and their overall significance. However, the study managed to capture a sufficient sample size that is a true representation of the different properties and their associated prices situated in close-proximity to a quality school. In this manner, the researcher believes valid and consistent findings were obtained. In addendum, the database had no age variable of each property for both sampled areas, which is a prerequisite based on the HPM. However, the registered year of each property was used a proxy in replacement of the age variable. However, this did not deter the significance of the coefficient statistics as the indicated valid sample of properties is deemed sufficient, including a myriad of explanatory attributes used in the study.

On another note, the unavailability of adequate data for traffic noise pitted the study in potentially providing a new perspective and extending knowledge on some of the overlooked environmental and locational determinants. To avoid reverting to traditional methods and testing the same hypothesis generally using similar attributes, scholars need a different methodological response to examining specific attributable effects on sale prices. For example, this can consider the effect of property-related crimes on property values. As reference to chapter 2, there are a plethora of different locational, neighbourhood and property determinants that explain a change of variation in sale prices. Failure to obtain data for traffic noise was a result of time and not having access to the appropriate apparatuses to adequately measure noise decibels in and around the sample areas. Moreover, the obtained listed property data did not provide any useful information that could have allowed a possible

measure of the effect of transportation noise on specific properties, their associated prices and structural attributes.

5.6. Recommendations and Direction for Future Research

Based on the abovementioned limitations, the following recommendations should be noted for future research regarding the property transaction market studies. The research analysis demonstrates some practical areas that should be explored to gain valuable knowledge on the interactions between microeconomic, macroeconomic variables and property valuations in the South African market that can also be applicable to other areas. The research also highlights that using national or regional property information can be misleading for any decision-makings related to the local level. It should also be noted that the research refrains from presuming that a high quality school is a key driving factor in household decision-making of location in region F (Kensington and Observatory). It is rather necessitating a broader outlook into the homebuyer's preferential basket of goods, in order to have an enriched knowledge-base of evidence-laden information influencing housing-decisions. Moreover, it highlights that it will crucially depend on the 'willingness and affordability' of parents to pay for goods and services offered within these sampled residential locations. The inclination to pay sizable premiums for local goods such as having access to a good quality school for their children's educational gains lies in the heart of standard urban economic theory (Gibbons, 2011).

Location is one of the key elements of housing value and typically, property transaction prices vary across space. Therefore, another practical area that future studies should explore is the effects of more geographical factors in the construction of property sale prices at the local regions. Upon the availability of more time series datasets in contrast to cross-sectional databases, research should be conducted on examining the long-run relationship of property sale prices and its determinants for other local real estate markets in the CoJ. For instance, researchers should also include, but not limited to, planning and development variables like land availability, bulk service infrastructure, and new constructions in investigating the fundamentals of property values. In regards to region F (CoJ) there is a plethora of attractions and other fundamentals that can explain such outcomes. This can be based on the fact that there has been constant strategic planning, developments and improvements in the quality of local public goods and neighbourhoods. This gives reference to the ongoing place-based improvements by the government as broadly inscribed in development policies to ensure radical economic growth, integration and diversity in some of the local areas within this region. To uncover the values of these factors, the research suggests that policy be directed to

availing a database of macroeconomic and microeconomic determinants that are fundamental in explaining property price variations in the South African property market.

This study further suggests that it is recommendable that more in-depth research is conducted to examine the value-laden attributes (tangible and intangible factors) of a school that should signify 'good quality' as understood by parents, investors and policymakers. Devising a standard proxy that embodies the determinants of a school quality will be complementary to the decidedly indecisiveness in previous studies on the contribution effect of school variables on the overall performance of students. Nonetheless, this is beneficial for decision-makers in assessing policies targeting to improve school variables such as test score performances. The employed estimation strategies and findings provide policy-makers with means to evaluate the gross benefits of policies intended to ameliorate the quality (tests scores, education and sporting facilities) of both public and private schools (fee and non-fee paying). Adding more understanding to the fine grains of this topic, a sheer volume of knowledge on the differences in quality (average test-scores) between schools could also be explained by other policy-laden variables.

As shown in the analysis of the results and literature, considering the coexistence of private and public schools in the same educational district is of paramount significance when designing school assignment policies. There is evidence that tracking the school choice theoretical models, the capitalisation of public schools tends to be attenuated by private school performances in property prices by providing a beneficial outside option to parents (Fack, et al., 2010). Radebe (2015) for instance states that this can take into account the influence of school admission rules (such as school choice and strict housing-based sorting's) on residential and school seclusions, educational inequalities and pupil performances will rely on how the property market integrates such school variables. Notably, discriminative admission requirements have been decriminalised by the enactment of the South African Schools Act of 1996 to allow access to learners of different background and Jeppe Boys High as a model C school is defiant from such policies (Heystek, et al., 2005) and has become an epitome of these acts. This school's admission covers a broader context of including leaners from areas such as south Alberton, Boksburg, and Edenvale.

The enrolment policy of Jeppe Boys High is good with an average intake of fourteen-to-sixteen thousand students per annum, which can also explain some of the demand of housing in this catchment area. Moreover, the more accessible the school the more value it exerts on surrounding properties and influences parents housing decision to reside in close distance to reduce travelling time. The implementation of such a policy in Jeppe Boys High that seeks to break the link amid school admission and place of residence has been crucial to avoid issues

that have previously resulted in the plummeting of intergenerational mobility. Additionally, with the availability of detailed data from the registry data (that could not be provided for the research because of confidentiality) on the characteristics of home-buyers, future research attempts can presumably help in augmenting the gap-in-knowledge on ‘sorting processes or selection by mortgage’ involved in the decision-making of households choosing housing locations, as highlighted in sub-section 2.2.1.

5.7. Concluding Remarks

The research has provided empirical evidence regarding the dynamics of the property transaction market at the local level in region F of the CoJ. The study examined the effects of school quality and the selected microeconomic determinants on property sale prices that explain price variations. Considering the review and analysis in the paper, it is revealed that most predictor variables that form the composite of property’s purchase prices exhibit significant statistical association with property sale prices including school quality in the short-run. This implies that there are other unobserved determinants and other demand factors that play a role in explaining sale price variations, which require further investigation and their effects in the long-run. This academic contribution has extended knowledge that potentially serves as a prospective directory in terms of the effective approaches to analysing property transactions and unobserved fundamentals within the CoJ market. The research has enumerated findings that can also qualify as a benchmark elicited to enable lucrative future decision-makings for households prioritising their location of residency based on access to a good quality school. It cannot be overemphasised based on the adage that it boils-down to an economic agent’s preferences, location, affordability and willingness, which differs per consumer on how they decide to maximize their opportunity costs on property transactions. In summation, the research has extended insight beyond contributing to literature by providing robust and empirical evidences that have significant practicalities and policy implications concerning property valuations in the CoJ real estate market.

References

African Architect, (2018). Johannesburg 1912 – Suburb by suburb research. https://johannesburg1912.com/2012/10/07/164421564/. [Accessed at 23/04/2020].

Alasuutari, P., Bickman, L. and Brannen, J. (2012). The SAGE Handbook of social research methods. *Chapter 8: Quantitative Sampling*.

Aliyu, A. A., Garkuwa, A. I., Singhry, I. M. and Adamu, H. (2016). Impact of school accessibility on residential property values: evidence from existing frontier of knowledge. *Conference Paper.*

Asiamah, N., Mensah, H. K. and Oteng-Abayie, E. F. (2017). General, target, and accessible population: Demystifying the concepts for effective sampling. *The Qualitative Report*, **22(6)**, pp.1607-1621.

Babbie, E. R. (2010). The Practice of Social Research. 12th ed. Belmont, CA: Wadsworth Cengage; Muijs, Daniel. *Doing Quantitative Research in Education with SPSS. 2nd edition. London: SAGE Publications, 2010*.

Bae, H. and Chung, I. H. (2013). Impact of school quality on house prices and estimation of parental demand for good schools in Korea. *Kedi Journal of Educational Policy*, **10(1)**, pp. 43–61.

Basu, S. and Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *Journal of Real Estate Finance and Economics*, **17:1**, pp. 61–85.

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

Black, S. E. and Machin, S. (2011). Housing valuations of school performance. In Handbook of the Economics of Education (Vol. 3, pp. 485-519). Elsevier.

Bokhs, (2018). The history of BHS. http://www.bokhs.co.za/index.php/ct-menu-item-11/ct-menu-item-13. [Accessed at 08/05/2020].

Bowes, D. and Ihlanfeldt, D. (2001). Identifying the effects of rail transit stations on residential property values, *Journal of Urban Economics*.

Brasington, D. and Haurin, D. R. (2003). Educational outcomes and house values: A test of the value-added approach. *Journal of Regional Science*, **46(2)**, pp. 245–268.

Brasington, D. M. and Hite, D. (2012). School choice and perceived school quality. *Economics Letters*, **116(3)**, pp. 451–453.

Carroll, L. (1989). Alice's Adventures in Wonderland. London: Hutchinson, pp. 63-4.

Cheriuyot, K., Wray, C. and Katumba, S. (2015). Spatial statistical analysis of dissatisfaction with the performance of local government in the Gauteng City-Region, *South Africa. South African Journal of Geomatics, Vol. 4, No. **3***.

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

Chung, I. H. (2015). School choice, housing prices, and residential sorting: Empirical evidence from inter-and intra-district choice. *Reg. Sci. Urban Econ*. 2015, **52**, pp. 39–49.

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.

Clark, D. E. and Herrin, W. E. (2000). The Impact of public-school attributes on home sale prices in California. *Social Science Electronic Publishing*, **31(3)**, pp. 385–407.

Curwin, J., Slater, R. and Eadson, D. (2013). Quantitative Methods for Business Decisions, 7th edition, pp. 447-449.

Dass, S. and Rinquest, A. (2017). Basic Education Rights Handbook – *Education Rights in South Africa.*

Davidoff, I. and Leigh, A. (2007). How much do public schools really cost? Estimating the relationship between house prices and school quality. *The Economic Record*, **84(265)**, pp. 193–206.

De Haan, J. and Diewert, E. (2013). Hedonic regression methods. *Handbook on Residential Property Prices Indices (RPPIs)*.

De Haan, J. and Diewert, E. (2013). Hedonic Regression Methods, in OECD, *Handbook on residential property indices*.

De Wet, T., Patel, L. and Forrester, C. (2008). Johannesburg poverty and livelihood study.

Dhar, P. and Ross, S. L. (2009). School quality and property values: *Re-examining the boundary approach. Working Papers*.

Dhar, P. and Ross, S. L. (2012). School district quality and property values: Examining differences along school district boundaries. Journal of Urban Economics, **71(1)**, pp. 18–25.

Dodds, R. (2010). An Investigation into the Hedonic Price Analysis of the Structural Characteristics of Residential Property in the West Rand. Johannesburg.

Dougherty, (2009). School choice in Suburbia: test scores, race, and housing markets. Trinity College Digital Repository. *Papers and Publications, American Journal of Education **115**. Paper 1.*

Downes, T. A. and Zabel, J. E. (1997). The impact of school characteristics on house prices: Chicago 1987–1991. *Journal of Urban Economics*, **52(1)**, pp. 1–25.

Du Preez, M. and Sale, M.C. (2014). Municipal assessments versus actual sales price information in hedonic price studies: A South African Case. *ERSA working paper, 411.*

Escobedo, C., Ramirez, A., Lujan, G., Serrano, D. and Guuerrero, J. (2007). Ethical issues with informed consent.

Fack, G. and Grenet, J. (2010). When do better schools raise housing prices? Evidence from Paris public and private schools. *Centre for the Economics of Education*, **94(1–2)**, pp. 59–77.

Feng, H. and Lu, M. (2010). School quality and housing prices: Empirical evidence from a natural experiment in Shanghai, China. *Journal of Housing Economics*, **22(4)**, pp. 291–307.

Fernández-Durán, L., Llorca, A., Ruiz, N., Valero, S. and Botti, V. (2011). The impact of location on housing prices: *applying the Artificial Neural Network Model as an analytical tool. Article*.

Fletcher, M., Mangan, J. and Raeburn, E. (2004). Comparing hedonic models for estimating and forecasting house prices. *Property Management*, Vol. 22 (3), pp. 189-200.

Fleishman, L., Fogel, N., Fridman, I. and Shif, Y. (2017). The effect of school performance on property values: a literature review and a case study. *Journal of European Real Estate Research*, **10(3)**, pp. 277-302.

Gibbons, S. and Machin, S. (2007). Valuing school quality, better transport and lower crime: *evidence from house prices*.

Gibbons, S. (2011). Houses and Schools: Valuation of School Quality through the Housing Market.

Gibbons, S., Machin, S. and Silva, O. (2012). Valuing school quality using boundary discontinuities. *CEE discussion papers, CEE DP 132, Centre for the Economics of Education, London, UK*.

Gibbons, S., Mourato, S. and Resende, G. M. (2014). The amenity value of English nature: a hedonic price approach. *Environmental and Resource Economics*, **57** (2), pp. 175-196.

Greene, W. H. (2008). The econometric approach to efficiency analysis. The measurement of productive efficiency and productivity growth, 1(1), 92-250.
Harris, A., Bennet, N. and Preedy M. (2003). Organizational effectiveness and improvements in education. *Maidenhead Open University press*.

Hayes, K. J. and Taylor, L. L. (1996). Neighborhood School Characteristics: What Signals Quality to Home-buyers? *Federal Reserve Bank of Dallas, Economic Review. Fourth Quarter,* **7**, pp. 2-9.

Heale, R. and Twycross, A. (2015). Validity and reliability in qualitative research. *Evidence-based nursing*, **18(3)**, pp. 66-67.

Heystek, J. and Nyambi, M. (2007). Section twenty-one status and school governing bodies in rural schools. *Acta Academica* **39(1)**: pp. 226-257.

Hilber, C. (2007). New Housing Supply and the Dilution of Social Capital, *London School of Economics, Research Paper in Environmental and Spatial Analysis*.

Hummel, N. J. (2017). A Meta study of hedonic pricing models on housing. *Bachelor Thesis Urban Port and Transport economics*.

IDP, (2012/16). City of Johannesburg. Integrated Development Plan: 2014/15 *review*.

Imberman, S. A. and Lovenheim, M. F. (2013). Does the market value value-added? Evidence from housing prices after a public release of school and teacher value-added. *Journal of Urban Economics*, **91**, pp. 104–121.

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.

Jeppe Boys, (2018). Sports. http://www.jeppeboys.co.za/sports/. [Accessed at 21/04/2020].

Jeppe Boys, (2018). Sports and Culture. http://www.jeppeboys.co.za/sport-culture/. [Accessed at 21/04/2020].

Jeppe Boys, (2018). Jeppe High School for Boys Association. http://www.jeppeboys.co.za/wp-content/uploads/2018/08/JBA-Newsletter-001-April-2017.pdf. [Accessed at 21/04/2020].

Kane, T. J., Staiger, D. and Riegg, S.K. (2005). “School Quality, Neighborhoods and Housing Prices: The Impacts of School Desegregation,” *NBER Working Paper No. 11347, May 2005*.

Karlsson, J. (2002). The Role of Democratic Governing Bodies in South African Schools. Comparative Education. Vol. 8, no.3 pp: 327-336.

Ka-wai, N. G. (2010). School quality and housing prices: Evidence from Hong Kong. University of Hong Kong.

Khachiyan, A. (2013). Isolating the School Quality Premium in Housing Markets. *Thesis. University of California, Berkeley Department of Economics.*

Kivunja, C. and Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, **6(5)**, pp. 26-41.

Lancaster, K. (1966). Change and innovation in the technology of consumption. *The American Economic Review*, **56(1)**, pp. 14–23.

Leech, D. and Campos, E. (2003). Is comprehensive education really free? A case-study of the effects of secondary school admissions policies on house prices in one local area. *Journal of the Royal Statistical Society*, **166(1)**, pp. 135–154.

Lokesh, P. N., Manthan, D. J., Muddukrishna, B.S., Bhat, K.M., Bairy, K.L., Udupa, N. and Musmade, P.B. (2013). Informed Consent: Issues and Challenges. *Journal of Advanced Pharmaceutical Technology and Research*.

Long, R. and Wilhelmsson, M. (2020). Impact of Shopping Malls on Apartment Prices: the Case of Stockholm. *Nordic Journal of Surveying and Real Estate Research Special Series, Vol. 5*, pp. 29-48.

Ludick, A. (2019). Determining the Impact of affordable housing development on property prices of adjacent neighbourhoods in South Africa.

Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics*, **18(6)**, pp. 723-729.

McCusker, K. and Gunaydin, S. (2015). Research using qualitative, quantitative or mixed methods and choice based on the research. *Perfusion*, **30(7)**, pp. 537-542.

Mnisi, N. and Karam, A. (2020). The impact of the integrated residential development programme on surrounding property values: case study of Fleurhof, Johannesburg. *Peer reviewed and revised May, 2020, published June, 2020*.

Nguyen-Hoang, P. and Yinger, J. (2011). The capitalization of school quality into house values: *A review. J. House. Econ*. **20**, pp. 30–48.

Noble, H. and Smith, J. (2015). Issues of validity and reliability in qualitative research. Evidence-based nursing, **18(2)**, pp. 34-35.

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** (1969), pp. 957-971.

Owusu-Ansah, A. (2012). Measuring and Understanding the House Price Dynamics of the Aberdeen Housing Market. *University of Aberdeen*.

Owusu-Ansah, A. and Talinbe Abdulai, R. (2014). Producing hedonic price indices for developing markets: Explicit time variable versus strictly cross-sectional models, *International Journal of Housing Markets and Analysis*, Vol. 7, No. 4, pp. 441-459.

Owusu-Ansah, A. (2018). *Construction and application of property price indices*. Routledge.

Owusu-Ansah, A., Anim-Odame, W. K. and Azasu, S. (2021). Examination of the dynamics of house prices in urban Ghana, *African Geographic Review*, **40:1**, pp. 76-91,

Pedersen, D. (1992). RAP, Rapid Assessment Procedures, Qualitative Methodologies for Planning and Evaluation of Health Related Programmes, INFDC, pp. 513-528.

Phillips, P. C. B. and Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, Vol. **58**, No.1, pp. 165-193.

Reed, R. and Sims, S. (2015). Property Development, 6th Edition. Routledge, pp. 317-321.

Ries, J. and Somerville, T. (2010). School quality and residential property values: Evidence from Vancouver rezoning. *The Review of Economics and Statistics*, **92(92)**, pp. 928–944.

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

Rowley, J. (2014). Designing and using research questionnaires. *Management Research Review*, **37(3)**, pp. 308-330.

RSA, (1996). South African Schools Act. Republic of South Africa, *GOVERNMENT GAZETTE. 15 November 1996*, No. 17579, p. 14.

Saunders, M., Lewis, P. and Thornhill, A. (2016). Research methods for business students. *Essex: Person education limited*.

Schooleen, (2020). Top 10 best high schools in Boksburg. Accessed at 08/05/2020. https://www.schooleen.com/best-high-schools-in-boksburg/.

Senick, A. (2019). Welcome to the neighborhood: A spatial analysis of crime and housing values. *Department of Economics. Honors Senior Project*.

Seth, S. (2017). The most important factors for investing in real estate. Investopedia. 30 January. [Online]. Available at:http://www.investopedia. com/articles/investing/110614/most-important-factors-investing-real-estate. asp [Accessed: 05 September 2020].

Schreuder, B. K. (2016). Section 21 schools: norms and standard funding for schools Grades 1-12 – financial allocation for the 2017/18 financial year.

Shani, Z. (2019). Assessing the Competency of South African Corporate Real Estate Manager and the Effects of Educational Path Taken. *The School of Construction Economics and Management.* Proposal 2.

Sims, C. A. (1980). Macroeconomic and Reality. *Econometrica* **48(1)**, pp. 1-48.

Sirmans, S. Macpherson, D. and Zietz, E. (2005). The composition of hedonic pricing models. *Journal of real estate literature*, **13(1)**, pp.1-44.

Strassburg, S. (2010). More than getting through the school gates. Barriers to participation in schooling. *Findings from the access to education study*, Volume 3.

Thompson, K. (2015). Positivism and Interpretivism in social research. Revise Sociology. [online] Available at: https://revisesociology.com/2015/05/18/positivism-interpretivism/.

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

Uren, T. R. (2007). Original school building located on ERPM Golf Course, photograph, p. 14.

Uren, T. R. (2008). Conservation of Heritage Buildings in Boksburg.

Walliman, N. (2011). Research methods: The basics. New York, Routledge.

Watkins, C. (2002). The Definition and Identification of Housing Submarkets, *Environment and Planning, A Review.*

Wen, H., Zhang, Y. and Zhang, L. (2011). An empirical analysis of spatial effects of housing price based on the spatial econometric model: A case study of Hangzhou. *System Engineering Theory and Practice*, **31(9)**, pp. 1661–1667.

Wen, H., Zhang, Y. and Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in Hangzhou, *China. Habitat International*, **42(42)**, pp. 155–163.

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

Wen, H., Zhang, Y. and Zhang, L. (2019). Quantile effect of educational facilities on housing prices: Do homebuyers of higher-priced housing pay more for educational resources?

Wild, J. (2016). The Impact of School Quality on Property Prices.

Wildeboer, A., Van Haaren, J. and Van Oort, F. (2017). Neighbourhood attractiveness and residential property prices; the impact of restaurants as local consumer amenities that foster encounters. *Empirical evidence from Amsterdam, Urban, Port and Transport Economics, Erasmus School of Economics*.

Wilhelmsson, M. (2000). The Impact of Traffic Noise on the Values of Single-family Houses, *Journal of the Environmental Planning and Management* **43**, pp. 799-813.

Wilhelmsson, M. (2009). Construction and Updating of Property Price Index Series. The Case of Segmented markets in Stockholm, *Property Management* **27**, pp. 114-139.

Wilhelmsson, M. and Ceccato, V. (2018). Does crime impact real estate prices? An assessment of accessibility and location In: Gerben J.N Bruinsma and Shane D. Johnson (ed.), *Oxford Handbook on Environmental Criminology Oxford University Press*.

Wilhelmsson, M. (2019). Energy Performance Certificate and its Capitalization in Housing Values in Sweden. *Building and Real Estate Economics, Royal Institute of Technology*, **114** (28), Stockholm, Sweden.

Wilhelmsson, M., Ismail, M., and Warsame, A. (2021). Gentrification effects on housing prices in neighbouring areas. *International Journal of Housing Markets and Analysis.*

Woolridge, J. M. (2009). *Introductory econometrics - A Modern Approach* (Thomson, South-Western, Mason, OH).

Wulsin, J. (2009). An analysis of the effects of public school quality on house prices in Durham, North Carolina. *Honors thesis. Economics Department. University of North Carolina at Chapel Hill*.

Yacim, J.A., and Boshoff, D.G.B. (2019). Vertical equity and uniformity in house price valuation. *African Real Estate Society (AFRES). Urban Real Estate Research.*

Yadavalli, A. P. and Florax, R. J. G. M. (2014). The effect of school quality on house prices: A meta-regression analysis. *Department of Agricultural Economics, Purdue University*.

Yang, L., Wang, B., Zhang, Y. and Ye, Z. (2018). Willing to pay more for high-quality schools? A hedonic pricing and propensity score matching approach. *International review for spatial planning and sustainable development*, **6**(1), pp. 45-62.

Zheng, S., Hu, W. and Wang, R. (2016). How much is a good school worth in Beijing? Identifying price premium with paired resale and rental data*. The Journal of Real Estate Finance and Economics*, **53(2)**, pp. 1–16.

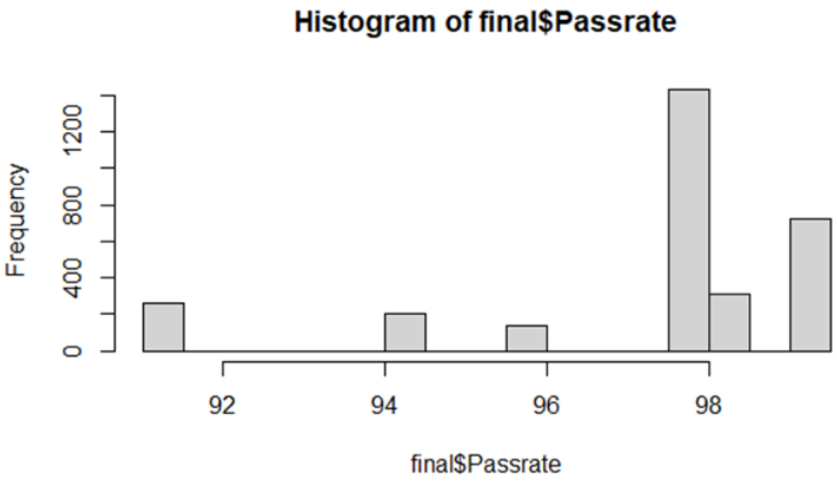
Zietz, J., Zietz, E. N. and Sirmans, G.S. (2008). Determinants of house prices: a quantile regression approach. *The Journal of Real Estate Finance and Economics*, **37(4)**, pp. 317-333.

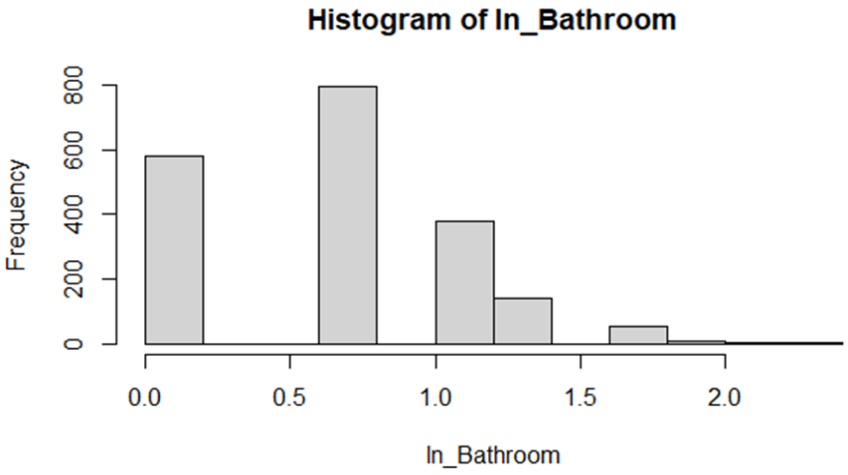
Annexure A: Published Studies on Estimating the Effects of School Quality on Housing Prices

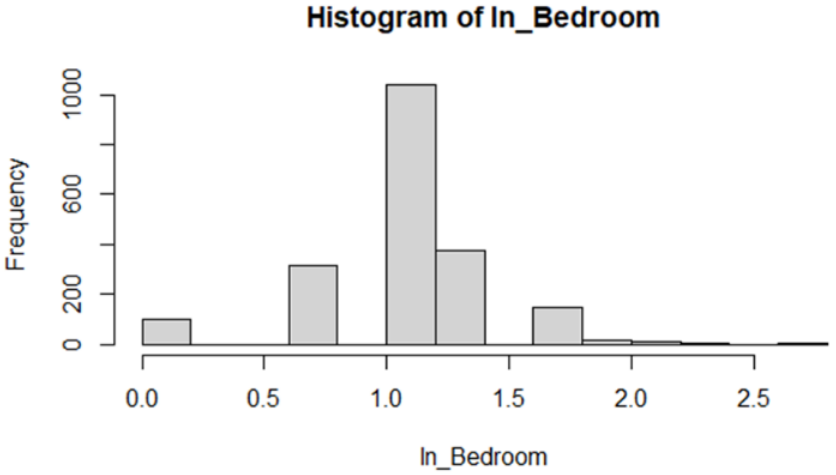
Studies with Boundary Fixed Effects (and Other Controls)	Context	Percentage Point Increase in Housing Prices
Black (1999)	Suburban Boston, MA, 1993–95	2.1
Kane et al. (2005)	Charlotte-Mecklenberg, NC, 1994–2001	2
Bayer et al. (2007)	San Francisco Bay area, CA, 1992–96	1.8
Clapp et al. (2008)	Connecticut towns, 1994–2004	1.4
Dougherty et al. (current study)	West Hartford, CT, 1996–2005	1.9

Source: Dougherty, et al., (2009). School choice in Suburbia: test scores, race, and housing markets. **Note:** See these papers above for details on these individual studies and other controls above.

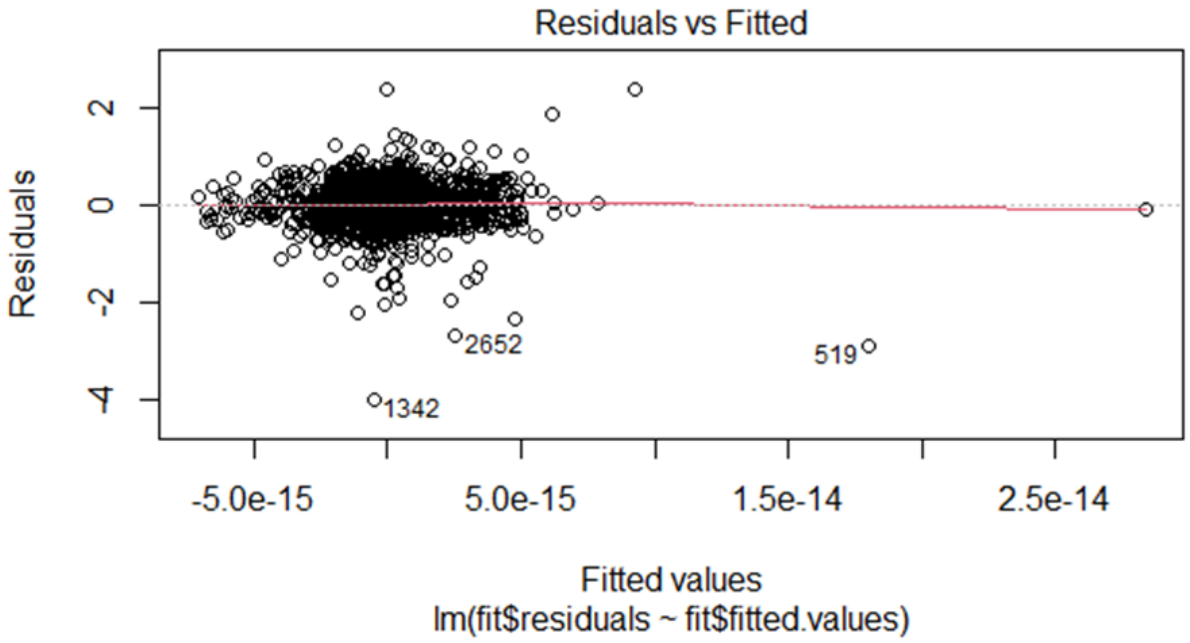
Annexure B: Distribution of the pass rate (school quality), bedrooms, bathrooms.

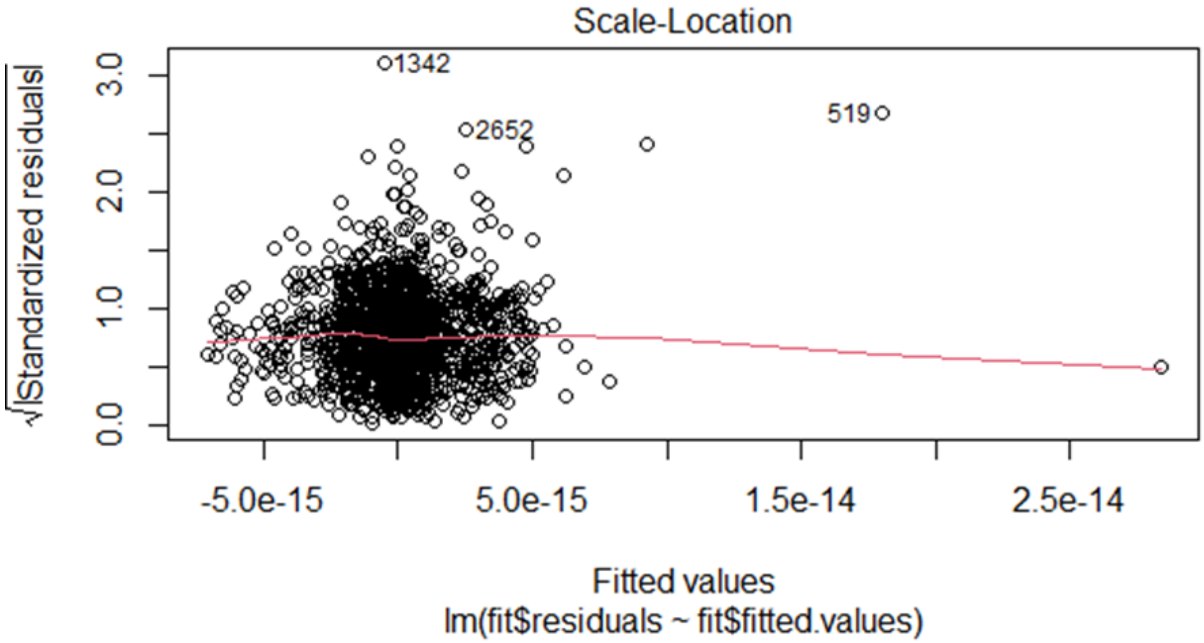






Annexure C: Diagnostic Testing for Possible Multicollinearity Issues





Appendices A: Clearance Certificate

School of Construction Economics & Management

University of the Witwatersrand, Johannesburg -PO Box 20, Wits 2050, South Africa • Tel: +27 (0)11 717 7652/77669 · Fax: +27 (0)11 717 9729 Email:CEM@wits.ac.za

SCHOOL OF CONSTRUCTION ECONOMICS AND MANAGEMENT RESEARCH ETHICS COMMITTEE

CLEARANCE CERTIFICATE	**PROTOCOL NUMBER CEM/20/10/TWS**
PROJECT TITLE	The effects of school quality and other unobserved neighbourhood characteristic's on residential sale prices.
INVESTIGATORS	Thantsha William Seremi 721032
SCHOOL/DEPARTMENT	**SCHOOL OF CONSTRUCTION ECONOMICS AND MANAGEMENT**
DATE CONSIDERED	18/09/2020
DECISION OF THE COMMITTEE	Approved conditionally with respect to the declaration.
EXPIRY DATE	8th October 2021
DATE 09//10/2020	**CHAIRPERSON** Prof. Kola Ijasan

cc: Supervisor: : Professor Samuel Azasu

DECLARATION OF INVESTIGATOR

To be completed in duplicate and **ONE COPY** returned to the Secretary Michelle Kirsten at the CEM reception desk.

I fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I undertake to resubmit the protocol to the Committee. **I agree to completion of a yearly progress report.**

Signature

09 / 10 / 2020
Date