EBS Universität für Wirtschaft und Recht

EBS Business School

A Quantitative and time Sensitive Evaluation on the Impact of the new Doha Metro on Land Values:

Block Abstract:

Purpose: The purpose of this research is to evaluate the impact of the newly build Doha Metro on property values. We compare values at two significant points during the production process of the Red Line: Announcement and construction phase.

Methodology: To our knowledge this is the first quantitative study to use panel data collected from the Doha Ministry of Municipality on property values. We use attributes established in Hedonic Pricing Models and spatial characteristics (distance to Metro, health care, education, CBD). We amend the econometric model by comparing impact and direction on property price at two significant points in time using statistical methods.

Findings: Our results highlight the nature of complexity to find the right model and significant attributes to explain for it. We find that the impact on property price does not follow the effect of positive impact due to proximity to rail. Secondly, we find that property values in Doha are highly related to construction phases.

Originality and Value: In the meta-analysis on “Impact of Railway Stations on Residential and Commercial Property Value” Debrezion et al (Debrezion, 2007) finds mixed results ranging from negative to positive impact on property values due to spatial characteristics, temporal effects, and methodology. Based on our findings we argue that time effects should be added to the model. For future research we see potential conducting a literature research/meta-analysis clustering results according to regional effects (Europe, Amerika, Asia) and according to greenfield versus brownfield metro projects.

Key Words: Property Value * Hedonic Pricing * Metro * Value Add

Prepared by: Dunja Demmler-Schunk Dipl.-Ing. – M.Sc. Real Estate

Adresse: Akazienweg 7, 63263 Neu-Isenburg

Table of Contents

1 ECONOMETRICS:

The following contributes by amending existing research on the impact of *new* rail infrastructure on land values in Doha, Qatar. Using Econometric tools and a first new data set. We asses possible impacts on land values due to proximity to Metro Stations in Doha, Qatar. We use three distinct distance bands to research the impact. The hedonic pricing model (HPM)is extended with two new variables. To our knowledge this is the first quantitative approach to changes in land values due to proximity to rail (Metro) for Doha, Qatar.

The research is unique as it is based on a Geodata – set obtained from the Ministry of Urban Planning and the Ministry of Municipality during two distinct construction phases (announcement of Metro and soft launch) of the Doha Metro.

Studies on the impact of public infrastructure on real estate is comprehensive. It is relevant to differentiate studies based on their researched parameters. Not only do they differ in the applied Method (Data collection and definition), statistical method (linear regression, Difference in Difference Method) but also in their findings (no correlation, strong correlation).

In a meta-analysis by Debrezion, Pels & Rietveld on the "Impact of Railway Stations on Residential and Commercial Property Value" they find variations in the findings in respect to impact magnitude and direction. (Debrezion, 2007)

Our study is based on a comparison of land values. We use land values to control for biased appraisal, as the property value is determined by structural attributes, land use, land rates and the location. Real estate (commercial and residential) is always determined by its unique set of attributes which makes it is hard to control for, adding possible bias.

Quantitative research history on the topic of land value uses internal and external characteristics attributable to differences in land value. The literature provides two approaches analysing land value in urban locations.

In his paper "Urban Economic Theory" Fujita offers to approach the differences in land value based on a set of constraints. (Fujita, 1989)

The second approach aims to explain the value by identifying the characteristics of land. These characteristics affecting the value can be physical and environmental as in the study by Bowes and Ihlanfeldt (Bowes & Ihlanfeldt, 2001) and characteristics can differ from study to study. We adopt the later approach and have added two location attributes, distance to health care and distance to schools in our model.

We arrive at our statistical model by assessing different specifications of access and a household´s willingness to pay. We draw on findings of Alonso´s Willingness to Pay Model (WTP) (Alonso, 1964 Reprint 2013), Alonso´s Bid and Rent Function and the Alonso-Muth-Mills model which is based on a monocentric city with its population living around a central business district (CBD) (Christy, 2008)

Sherwin Rosen extends the model with a Bid and Rent Function and establishes the Hedonic Pricing Model (HPM). The HPM is a pricing model accounting for the internal and external factors which affect land values. (Rosen, 1974)

We incur characteristics and attributes used in HPM and expand the function by adding "access to education" and "access to medical care" to the function to develop our model and perform a regression analysis.

Robert Cervero and Jin Murakami suggest an increase in land values due to proximity to public infrastructure. They measure the impact, among other variables, according to the plot distance to metro station using concentric distance bands. (Cervero & Murakami, 2009)

We adopt three distance bands for our model 0-500m, 500 – 1.000m and further than 1.000m from centre Metro station to measure possible changes in land values.

In a young nation like Qatar, without any history of public infrastructure but individual car traffic running subsurface, the public shows mixed reactions to this novelty. Hence, we expect the announcement of the new metro in Doha to have the reverse effect on land values, i.e., a drop in land values at close quarters (500m) to Metro stations.

We dissent from collecting data before and after the construction of a metro but *during* announcement of the new metro and during soft launch. We choose these time bands as we expect a deviation from the expected increase in land values.

We contribute to the discussion, that changes in land values may be correlated to very specific times during urban development. A new infrastructure system is constructed over a substantial time-period and might show a correlation to project maturity.

1.1 Introduction

1.1.1 Qatar Rail: The transport innovation in Doha, Qatar

The State of Qatar is a country in West Asia and is one of seven independent emirates located on the Arabian Peninsula. The Peninsula is made up by Saudi Arabia, Oman, Yemen, Kuwait, Bahrain, and United Arab Emirates (UAE) which they share a border with. Qatar is the home of 2,88 Mio. residents of which approx. 2,3Mio. live in Doha, the capital of Qatar. (Crystal, 2021)

Since its independence in 1971, Qatar has grown substantially both in numbers as well as economically. Qatar´s roadmap and pegged requirements to transform into a “advanced society capable of achieving sustainable development by 2030” are set in the Qatar National Vision 2030 published in 2008. Based on the Vison 2030 to reduce the carbon footprint and as a precondition holding the FIFA 2022 World Cup, an integrated public transportation system is introduced. (General Secretariat for Development Planning, 2008)

One pillar of the integrated public transportation plan is the planning and construction of a new rail, light-rail, and Metro System. To archive the infrastructure goal Qatar Railway Company a state-owned company is established in 2011, responsible for the design, construction, commissioning, operation and maintenance of the entire rail network and systems (Qatar Rail, 2021).

After the initial study and the planning assessment procedure, construction on the Metro system started in 2012.

During Phase I, three metro lines (Red, Gold, and Green) opened with 37 stations and approx. 300 km track length.

(Qatar Rail, 2021)

The line under study will be the "Red Line" composed of 17 Stations and 40km track, running from north to south. It connects major infrastructure, social and economic hubs, such as the Qatar International Airport, the Central Business District (CBD), Katara (Centre for the arts) and Lusail (residential hub).

(Qatar Rail, 2021)

The opening of the new Metro is a milestone in the history of Qatar. Up to this point Qatar could not offer any rail-based transport to its 2,6 Mio. population of which 96% live in an urban environment. (Planning and Statistics Authority, 2021)

How did the planning and construction effect land values in the proximity of Metro stations?

1.1.2 Background Theories and area of research

In the following theories "public transportation" and "accessibility" are quoted as factors influencing land and property value. Can we draw on the following theories to understand how to predict the impact of rail proximity on land values?

Relevant for this study is the standard location model of William Alonso (1964), which was extended by Mills (1967) and Muth providing the cornerstone of modern neoclassical urban *location theory*. Alonso is the first offering a full analysis of urban land usage within the orthodox economic theory. (Muth, 1969)

Orthodox economic theory is the study of how humans make decisions in the face of scarcity and is considered one of three types of economic thinking – orthodox economics, institutional economics, and Marxian economics. Applying recognized scientific analysis to the elucidation of our present economic system (including deduction, induction, and statistical methods) (Estey, 1936).

Location theory states that: Accessibility benefits will be capitalized into property values. Suggesting that "any improvement in transportation infrastructure is capitalized into land values in a short–term urban partial equilibrium" (Alonso, 1964 Reprint 2013). Hence *accessibility* is defined as one of the influencing factors for land value.

The intra-urban Alonso-Muth-Mills model suggests that "high prices close to the city centre are offset by short commutes." (Glaeser, 2007). This is especially true for monocentric cities. Most models therefore use *distance to Central Business District (CBD)* as a factor influencing land values.

Location theory suggests that "any improvement in transportation infrastructure is capitalized into land values in a short-term urban partial equilibrium" (Alonso, 1964

Reprint 2013), we use *distance to Metro* measured in three different radii as a factor influencing land values.

Researchers across the world and from varying disciplines study the impact and/or effect of rail systems on property values (see list of references). As they research the impact attributed to different academic disciplines i.e., law, economics, sociology, architecture, to name a few, they measure impact and effect of new Metro lines diverse.

We approach the field focusing on economics and econometrics.

The motivation is to in a first step understand the impact of rail proximity on land values and in the following chapter assessing its implication on a possible value capture scheme to financing rail investments.

1.1.3 Sustainability regarding the research question (Research Questions)

This work draws from both economic as econometric fields to answer the following questions:

Research Question 1 [RQ1]: How does the land value of residential/commercial/mixed use properties change due to their vicinity to a metro station in Doha, Qatar?

Research Question 2 [RQ2]: How does the land value change within set radii (0-500, 501-1000m, >1000m) from the announced and new metro stations in Doha, Qatar?

Research Question 3 [RQ3]: Are the commonly used attributes of the Hedonic Pricing Model (for Real Estate) sustainable in a car driven, new economy such as Doha?

Research Question 4 [RQ4]: Does the assessment during a specific construction phase influence the impact and direction of the result?

Research Question 5 [RQ5] Has the announcement of the new Metro influenced Land values in Doha?

We use the introduced economic models as well as an econometric approach to answer the research questions. In the following discussion the basic econometric approach is introduced and pertinent theories.

1.2 Introduction to Econometrics

Economist Adam Smith and his work "An Inquiry into the Nature and Causes of the Wealth of Nations" (Smith A. , 1776) is referred to as the founding father of Economics.

Economists seek to analyse the forces determining prices. They assume that "prices of the various things must be interrelated; economics therefore ask how such a price system or market mechanism is correlated and what conditions are necessary for its survival" (Blaug, 2020)

The following work resorts to the above-described economic theory and uses econometrics to test the theory using statistical methods. This is done in the framework of an empirical study.

Econometrics is a standalone discipline within the field of Economics. Econometrics aims to combine economic theories, empirical data, and statistical methods. (Wirtschaftslexikon Gabler, 2018).

It was Ragnar Frisch who coined the term econometrics and is quoted as the founding father of econometrics. The Nobel prize winner used statistical methods to describe economic systems.

In 1928 it was Frisch together with Charles F. Roos (a young mathematician) who agreed that there is a need to combine economics, mathematics, and statistics closer together. Charles F. Roos who was a Fellow at Princeton University and Ragnar Frisch who was a visiting Professor at Yale founded the Econometric Society in 1930 (Christ, 1983).

Together with Jan Tinbergen they formulated mathematical formulations of economic mechanisms.

Econometry can be subdivided into the following two approaches to economic problems: theoretical-quantitative and empirical-quantitative. The following work adopts an empirical-quantitative approach using the economic model of hedonic pricing and the statistical method of a linear regression.

1.2.1 Methodological tradition

Following the tradition of quantitative research, this section resorts to a linear regression method to estimate the impact of rail proximity on land values during construction time in Doha, Qatar.

The hedonic pricing function is chosen as a strong quantitative pricing valuation tool. It relies on robust data and significant model specifications. We will discuss the limitations of the model in the following chapters on limitations and bias.

1.2.2 Hedonic pricing model

We refer to the hedonic pricing model as it has shown its strength and robustness in multiple research areas such as consumer and market research (Hirschmann & Holbrook, 1982), calculation of consumer price indices (Moulton, 1996), and tax assessment (Berry & Bednarz, 1975) and the real estate market (Lancaster, 1966).

Lancaster argues that it is not the good itself (plot of land, house, flat) that creates utility, but the relevant *characteristics* of the good. Since a good has multiple utilities and characteristics the aggregated value of each creates its hedonic utility. In his theory he claims that attributes or characteristics can be arranged into groups. (Lancaster, 1966)

The Hedonic pricing model developed by Rosen (Rosen, 1974) derived from the characteristic theory of value proposed by Lancaster (Lancaster, 1966) and seeks to explain a good considering the internal and external factors or attributes to define its price/value. Prerequisite for the applicability of the model is that a good is directly influenced by internal or external environmental factors.

Fig. 2: Foundation of models used for land use and property value analysis
Source: Adapted from McDonald & McMillen[23]

Quelle: (McDonald & McMillen, 2007)

The *hedonic pricing model* in real estate deconstructs the value of a property into its underlying bundle of attributes. These attributes can be building characteristics and external characteristics and define or estimate the sum of implicit prices households are willing to pay. (Monson, 2009), (Wolf & Klaiber, 2021).

Using regression, these components determine the current and future transaction price through analysis of its correlations (relationship) to a dependant variable.

The basic function for this estimation is written as:

$$P_h = f(S_i, N_j, A_k, Q_M)$$

$$P_h = good\ price\ (house, property, land\ value)$$

$$S_i = site, property\ characteristics\ (room, garden, size)$$

$$N_j = neighbourhood\ characteristics\ (distance\ CBD, crime\ rate)$$

$$A_k = accessibility\ (public\ transport, roads)$$

$$Q_m = environmental\ quality\ (noise, pollution)$$

FIGURE 1:

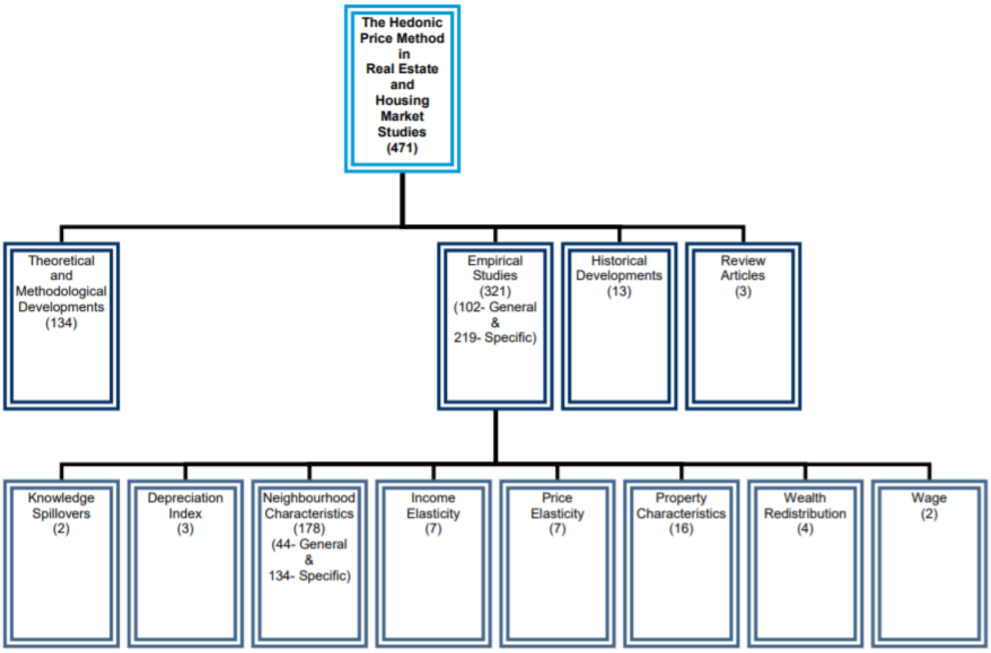
Hedonic methods are often used in real estate pricing to make estimates on property values. The main objective is to find a strong set of attributes which best describe the good and reflect its value best.

The function shown in Figure 1 is used in the property/housing sector as well as determining consumer price indices because of its flexibility to accommodate variable factors as well as stress probable connections between selected factors.

The hedonic pricing model is a valuation which can be translated into a hedonic regression model. We select a set of variables or characteristics best determining a fair value/price. The regression of this bundle of attributes or characteristics in a specified context (here: land value) determines the contributory value of each characteristic separately on the price. We apply econometric tools SSPS and Gretel to determine which attribute has a strong impact on the land value and which can be omitted.

The available literature provides volumes of research with sometimes contradicting findings on the topic of factors influencing land value. This can be attributed to differences in, research methods, nature of data sources, and temporal effects. (Mohammad, Graham, & Melo, 2017 (Agostini & Palmucci, 2008)).

The different approaches are shown in a work by (Herath & Maier, 2010) which groups the papers into its different hedonic pricing methods. Today the actual number of studies rises way beyond the study presented in 2010.



Graph 1. The types of real estate and housing market studies that use the hedonic price method
Notes: Number of papers is shown in parentheses.

Quelle: (Herath & Maier, 2010)

We use the basic function and add our specific set of attributes. These are introduced and discussed in the following chapter on research setting.

Both models draw on data obtained from the Ministry of Municipality in 2019 using Geodata and confidential information from the Ministry of Municipality. The raw data must be treated confidential and will only be disclosed on special demand and permission of the Ministry.

1.2.3 Literature and related work

The history of the hedonic pricing method is described in the previous chapter. Hereafter we draw on literature and related works *directly* related to rail infrastructure and more specific metro infrastructure and the effect on land values.

The effect of rail systems on property values and the range of estimates have been examined by a substantial number of researchers. Mohammad et al. conducted a Meta-

Analysis of the impact of rail projects on land and property values and found substantial differences in the outcomes, effects, and estimates. (Sarah I. Mohammad, 2013)

Most research papers conclude a positive effect on property value due to proximity to rail. Examples for property enhancement was found for example in studies from Laakso (Laakso, 1992), for residential properties in Helsinki, Pan, and Zhang (Pan & Zhang, 2008) for properties in China. Cervero and Duncan as well as Weinberger (Weinberger, 2001) found significant capitalization of about 23% for commercial land located in the vicinity of light rail transit (LRT) for properties in the USA. (Cervero & Duncan, 2002)

Although multiple papers show evidence of value uplift due to proximity to rail infrastructure there is a significant number indicating a negative or no effect on property values. For example, the studies by Bolling, Ihlandfeldt, and Bowes (Bolling, K.R., & Bowes, 1998) and Gatzlaff and Smith (Gatzlaff & Smith, 1993).

To understand if we can correlate findings from developed continental regions to the middle East region, we look at the neighbouring countries. Currently there are no studies from neighbouring middle east countries such as Oman, Kuwait, Saudi-Arabia as they are in the construction phase of new Metro systems. The only research placed in the region is from Mohammad et al. on the Dubai Metro.

FIGUR1 2

WWW.GOOGLE.COM/SEARCH?Q=WORLD+MAP+QATAR+MIDDLE+EAST

Mohammad et al contribute with their research paper on the effect of the Dubai Metro on residential and commercial properties. They found a positive effect on commercial properties in the second distance band (700-900m) from metro stations. They found no significant effect on residential properties. (Mohammad, Graham, & Melo, 2017).

Amir Forouhar and Hanoosh Hasankhani found contradicting effects on residential property and land values. They found an uplift in low-income areas but a significant negative effect on treated land in the northern high-income part of Tehran, Iran. (Forouhar & Hasankhani, 2018)

The study by Berawi et.al. found that proximity to rail transit has an “insignificant impact on the commercial property price compared to other variables such as building size, number of rooms, location, and distance to hospitals.” These results are taken from their study on impact of rail transit proximity on commercial land values in Jakarta, Indonesia. (Berawi, Miraj, & Saroji, 2020)

The observed differences in study results might be due to the differences in the hedonic pricing formula, model, variables (building size, plot value) location, different rational for approach. To stay free from any bias the following research must be conducted open-ended.

1.3 Empirical methodology

The hypothesis is that the land values in direct proximity to the new Metro experience a change in land value. As we have no reference for Doha, Qatar we aim to understand the strength and direction of a possible change. This is part of the greater question: How and how much a potential land value gain can contribute to land value capture to finance new and future public infrastructure in Qatar, Doha.

We look at different land usages (residential, commercial, mixed) and distance bands to understand possible contributions. Is it possible to use land value uplifts to finance the new Metro and additional Metro lines? First, we analysed the three new Metro lines and chose the Red Line as it is the longest and most comprehensive line embracing residential, commercial, and special purpose areas as well as the Central Business District of Doha.

As referenced in the chapter choice of variables and literature, proximity to the CBD has shown to have a positive correlation on land values.

We collected appr. 20.000 separate data points using GIS data to determine the distances of land plots to its nearest Metro station as well as distances to education, healthcare, and the CBD.

The information on land values was granted by the Ministry of Municipality as well as the information on the main usage (residential, commercial, mix use) of the land plot. This was a generous exception, as there is currently no expert Committee or else to obtain such information.

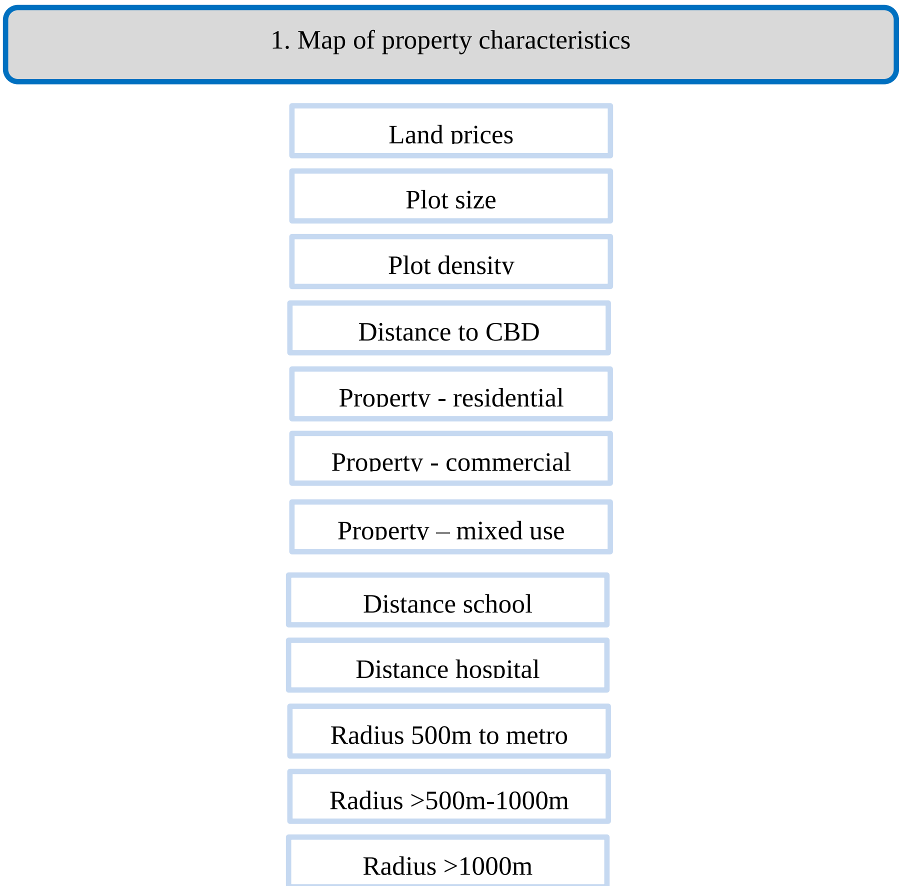
All land plots attributed with a land value are categorized into a distance band via a vector from centre plot to centre Metro Station to evaluate possible correlation of distance on land value.

As this is a new Metro system, we used two points in time to analyse if the market is biased and record negative or positive correlations due to a disruption of the urban fabric. We chose two points in time: the year of announcement and the year of soft launch. The plots are constant at the two assessed periods and create a substantial panel data set.

As The research is conducted using the statistical software SPSS [Statistical Package for Social Sciences] and Gretl [Gnu Regression, Econometrics and Time-series Library] on data obtained by the Ministry of Municipality, Doha to carry out an ordinary least square regression (OLS).

Economically implicit prices are estimated by the first step regression analysis. The land value is regressed on attributes chosen to explain the changes in land values. The land value is the dependent variable, whereas the attributes are the independent variables. An error term is added to complete the function of our model.

1.3.1 Research setting (Research Plan)



2. Estimate contributions to land price HP model

3. Estimate coefficients ß of variables

4. Iteration process – omit variables least correlated

1.3.2 Research Method (Data)

This chapter explains the choice of variables and their origin used in the Hedonic Pricing model. Furthermore, the corresponding relevant descriptive statistics are introduced and discussed. An extensive list of all variables under study is provided in the Appendix.

1.3.2.1 Endogenous (Dependent) Variable

As we are studying the impact of metro proximity on land values the endogenous variable in our model is defined as a price vector. The endogenous variable is defined as being determined in the model by values of the exogenous variables of the model. They are described in section exogenous variables.

The price in our model is defined as price/m² or Qatari Ryal per square metre. We use land values in our empirical study as we do not have to correct for building characteristics (age, amenities, structural, fit-out, apartment size) hence focusing on the effect of the exogenous variables described by location and neighbourhood characteristics. The focus being the effect of distance to new Metro stations.

1.3.2.2 Exogenous (Independent) Variables

The exogenous variables are determined outside the model. They take their values or behaviour as given, observed, or measured. The model includes observed variables which are measured in m such as location and neighbourhood characteristics (distance to education, distance to shopping, distance to medical facilities, distance to CBD) and dummy variables which can take the value 0/1 depending on the definition such as property type (residential, commercial, mixed use). The model also includes a variable for latent = unobserved variables.

Depending on the research question authors used many different factors affecting property values such as physical, environmental, accessibility, historical, economic, social, and land use. As to the relevance of this analysis we have chosen the factors explaining the relationship of metro proximity and land value. The rational of this selection is explained in detail in the following section

1.3.2.3 Accessibility and location characteristics

Accessibility Characteristics are described in HP models by using distance to nearest road, distance to nearest bus station and distance to rail, distance to airport to name the most prevalent.

As part of the Vision 2030 the road network is under permanent development and change. Therefore, this characteristic is omitted from the model.

The new Doha Airport opened in 4/ 2014. During the start-up phase the airport operated concurrently to the previous airport. The new airport is located outside the boundaries of the city. We do not include data on distance to airport as two airports would bias our results.

The bus terminals are an integral part of each subway station. The distance to the nearest station is marginal and is therefore not considered in our model.

1.3.2.4 Distance to new Metro Station

Dewees found the increase to occur perpendicular to the subway station within a 480m (one-third mile) walk to the station. (Dewees, 1976) Further studies found that stations located within a maximum distance of not more than 2 kilometres show a mostly positive effect on property values due to its proximity to rail stations (Agostini & Palmucci, 2008), (Billings, 2011)

Different studies used different catchment zones assessing the impact. The zones vary from 0,5 mile (0,8 kilometre) and within 1mile (1,6 kilometre) in the study by Billings (Billings, 2011) to the catchment zones used by Bowes and Ihlanfeldt of 0,25 miles (0,4 kilometre) and 3 miles (4,8 kilometre) from a station (Bowes & Ihlanfeldt, 2001).

According to the meta-analysis by Debrezion et al. based on studies conducted in the United States and Europe before and after 1990 they found the *mean* impact of a station on property value (rent) for every 250m closer to the station to be 2,61%. (Debrezion, 2007). They also found great variation in the results. They vary from -12,84% to 38,70%. There is also evidence that rents do not follow a linear function.

Mulley and Du show that in their geographically weighted regression (GWR) proximity has shown a *negative* (or no effect) effect in the Tyne and Wear region, United Kingdom. (Mulley & Du, 2006).

For our study we choose three catchment zones. The zones are established by drawing concentric circles with centre station as midpoints. The distance is measured as a direct vector from centre station to centre plot using Geographic Information System (GIS).

Henceforth we establish three distance bands in the following model. Distance band I is 500m to metro station, band II between 500m – 1000m and last band III above 1000m.

1.3.2.5 Distance to Central Business District

Translating the underlying theory that property value increases in response to transportation cost savings to the marketplace, we add the variable distance to CBD. The bid-and-rent model explains for a rent gradient that declines with distance from the central business district for sites that yield equal utility. Hence closeness to CBD is considered more attractive. Early studies by Dewees (Dewees, 1976) found an increase in residential property values between subway stations and residential property value. Fejarang however adds that transport costs and investments in transport reduces the increase in property value. A negative impact on property due to negative externalities is reported in the research by Bollinger C. R., Ihlanfeldt R., Bowes D. R. (Bollinger C. R., 1998). Some research cannot verify any provable effect (Gatzlaff & Smith, 1993).

1.3.2.6 Distance to education and medical care

Studies by (Yan Dong, 2006) and (Wen H., 2014) add the variable Distance to education (school, university, kindergarten) and distance to medical care to the neighbourhood characteristics of the HP model. Studies show that distance to education is valued different according to cultural environment.

1.3.3 Omitted variables TEXT

1.3.3.1 Crime rate

Bowes and Ihlanfeldt (2001) included crime rates in the model and found high crime rates to have a negative impact on real estate prices situated between 1,6km and 4,8km (one and three miles) from the station. Bollinger Ihlanfeldt and Bowes find that unsafe environment outweighs proximity leading to negative results. This correlation is shown in their study “The impact on Hiawatha Light Rail on Commercial rents” (Bolling, K.R., & Bowes, 1998).

The Qatar tribune released Qatar´s position in the global crime index. The ranking includes 133 countries. Qatar scored the lowest crime rate with 11.90 points out of 100 and ranked first scoring 88.10 out of 100 in reverse of the Crime Index. (Qatar Tribune, 2020)

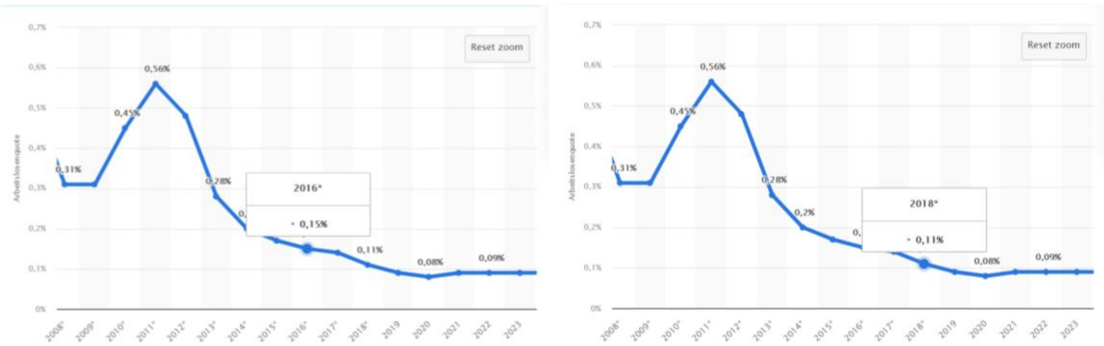
We have two reasons not to include the variable crime into the model First our distance bands follow smaller radii <1,6km and second the crime rate in Doha is the country with the lowest reported crime rate.

1.3.3.2 Unemployment rate

Unemployment rate has shown to be correlated to property prices and land values. The studies show two effects attributed to emlployment/unemployment.

In the study by Ahlfeldt he argues that employment gravity captures proximity to the central business district and also accounts for all employment centres of a multy-central cities effects on residential land prices. Hence overcomming the assumption of monocentric cities. (Ahlfeldt, 2011)

As shown in the diagramms retrieved from the general Secretariat for Development Planning during the years 2016 and 2018 respectively Qatar has an unemployment rate of 0,15% 2016 and 0,11% in 2018. This is due to the fact that access into Qatar depends on being employed.



(General Secretariat for Development Planning, 2008)

Hence a variable for un/employment rate is omitted from the model.

1.3.3.3 Pollution Characteristics:

Pollution is part of the environmental amenities that emanate from the neighbourhood onto adjacent land. These amenities can be good/bad in this context positive or negative impact on the land value. The factor pollution in our context is mainly noise pollution and mainly attributed to train lines running above ground. This negative externality can influence the value of properties along the railway line. The line under study is a metro line running underground, hence is omitted from the model.

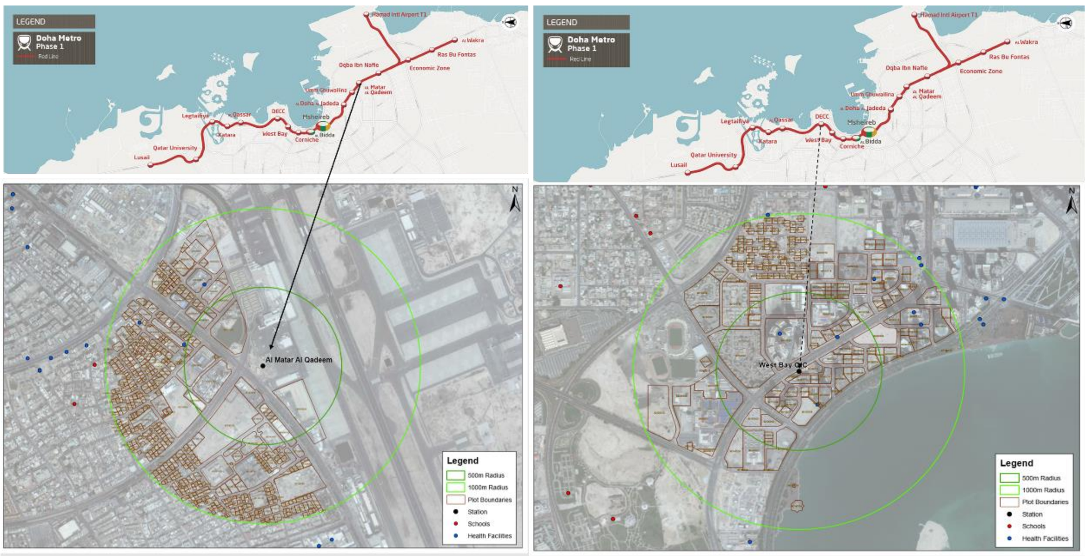
Noise pollution due to construction works is omitted due to manifold construction in the entire city. During the period under study many projects are realized simultaneously, road upgrades, city upgrading, new developments and demolition works. An unambiguous assignment of noise due to construction works on the rail project would not be possible within the model.

1.3.3.4 Time Varying Effects

Referring to the work of Li at al. Time-Varying Models (TVM) are considered a natural extension of linear regression models where a single estimate of each covariate´s effect is provided, whereas in a TVM the coefficients can vary over time (Li, 2017). In our model we draw on metric values such as distance and size or dummy variables. We argue that these independent variables do not vary over time.

1.3.4 Research sample

The research sample is chosen to cover the metro stations of the red line in Doha, Qatar. The centre station is chosen as the centre of the concentric circles with the radius 500m, 1.000m and border of district. Each plot is measured from the centre of plot and is specified by a plot number. The data is collected via Geodata (GIS) and then matched with the relevant price information by the ministry of municipality. The distance to hospitals, schools and Central Business District is measured in the same way. It is measured as the direct distance from centre plot hospital to centre metro station. Respectively the distances to education (schools, kindergarten) and Central Business District.



Cross sectional data – or panel data

The obtained data is collected over two distinct periods in time and refers to the same set of land plots. Our data can be seen as a subset of longitudinal data as the observations are for the same subjects (land plots) each time. Observations with only one or missing information referring to the selected points in time are omitted from our dataset.

1.3.5 Baseline Models

The hedonic pricing model estimates the value or vector of each characteristic attributed with a good or product by comparing market prices among these goods/products with different amounts of the attribute.

The underlying assumption is that a good can be described by a set of heterogenous attributes. Assume a good/product can be described as the sum of prices for each of these attributes.

Using standard hedonic theory (Rosen, 1974) the function is:

$$P_h = f(S_{h,} N_h, A_h),$$

where h the home is defined by its structural attributes $(S_{h,})$, the neighbourhood characteristics (N_h) and the environmental amenities as (A_h).

Written in a universal form as:

$$P = f(x_{1,} x_2, ., x_i)$$

where P is the market price of the good and $x_{1,} x_2, ., x_i$ represent the attributes of which it is formed.

The partial derivative of the hedonic price function (with respect to a certain characteristic) $x_{1,}$ equals the marginal price of that characteristic as it represents the marginal willingness to pay. Hence each attribute has an implicit price. (Anderson & West, 2006)

Differentiation of the hedonic price function with respect to a particular attribute yields the marginal implicit price of that attribute, which equals the homebuyer´s marginal willingness to pay. (Rosen, 1974)

Rosen´s theory has little to say about the functional form of the hedonic function. Hence baseline models vary according to research questions, econometrical approaches, Methods, and the set of attributes chosen.

Three empirical models are suggested to allow for more flexibility in the HP model. The following models show a suitable function form specification.

In the initial model follows a linear function. It is limited by the assumption that its marginal price does not depend on the initial level of each explanatory variable. Hence assuming that the relationship between explanatory variables and price is *linear*.

$$P = \alpha_0 + \alpha_1 Z_1 + \varepsilon$$

Under the double-log model or log-log specification, we can measure the changes in explanatory variables relate to the dependent variable in relative terms. The semi log is written as:

$$logP = \beta_0 + \beta_1 Z_1 + \epsilon$$

and the double log as:

$$logP = \partial_0 + \partial_1 \, logZ_1 + \epsilon$$

1.4 Calculations

Using the hedonic pricing model, we estimate the *effect* of selected attributes on land values in relation to the new metro line and its stations. The focus of the research is a possible time effect and the selected set of attributes.

The base model is rewritten as:

$$\ln P_{j,t} = \alpha + \alpha_T T_t + \beta_{nm} NM_j + \alpha_{PC} PC_{j,t} + \ln \alpha_{NC} NC_{j,t} + \varepsilon_{j,t}$$

We measure if the chosen life cycle phase of the project has an impact on land values. The same set of land plots are evaluated at two points in time. Under study is the time of announcement 2016 of the new metro in Doha. Second 2018 soft launch of Metro line. The variable $\alpha_T T$ is a dummy variable taking the value 1 or 0 respectively.

To describe the special impact of the metro on the land value of residential, commercial, and mixed-use plots, the effect of accessibility within three defined catchment zones is distinguished and measured separately.

First zone is defined as the area within a 500m radius to centre of metro station. Second zone as land plot within a 500-1000m radius to centre of metro station. Third measures the impact of plots in a distance greater 1000m but within a district to centre of metro station. The variable $\beta_{NM} NM_j$ is a dummy variable which takes the value 1 if the land plot is within the specific radius or zero if not.

Following contextual variables are included to control for property and neighbourhood characteristics.

The variable $\alpha_{PC} PC_{j,t}$ represents property or land use characteristics. It is a dummy variable which takes the value 1 if it is one of the following: commercial or residential or mixed.

To control for unobserved heterogeneity and confounding in the treatment effect, we include location-specific variables NC. The Neighbourhood characteristics distance to CBD, distance to nearest school, distance to health care, are expressed as $\ln \alpha_{NC} NC_{j,t}$. The distance is measured metric as a direct line from centre plot to centre school, health care and CBD.

The equation is completed with the standard error term $\varepsilon_{j,t}$ reflecting how much sampling fluctuation our statistic will show.

Based on the above baseline model we run the regression analysis to describe the statistical relationship between one or more predictor variables and the response variable.

1.4.1 OLS Regression:

The Ordinary Least Squares (OLS) in statistics is a type of linear squares method. Using the OLS technique, we choose the best fit determinants to predict parameters to describe a linear function of a set of explanatory variables. This is an iterative process.

We have used the function described in the preceding chapter for the OLS stats model.

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten			95,0% Konfidenzintervalle für B		Korrelationen			Kollinearitätsstatistik	
		Regressions koeffizientB	Std.-Fehler	Beta	T	Sig.	Untergrenze	Obergrenze	Nullter Ordnung	Partiell	Teil	Toleranz	VIF
1	(Konstante)	199812911,9	17489970,94		11,424	<,001	165483016,8	234142806,9					
	DCBD	-21227,134	3437,145	-,210	-6,176	<,001	-27973,677	-14480,592	-,210	-,210	-,210	1,000	1,000
2	(Konstante)	281763671,7	22141594,22		12,726	<,001	238303339,2	325224004,1					
	DCBD	-21772,270	3371,455	-,215	-6,458	<,001	-28389,887	-15154,654	-,210	-,219	-,215	,999	1,001
	Xr	-128829715	22017324,35	-,195	-5,851	<,001	-172046126	-85613303,7	-,189	-,199	-,195	,999	1,001
3	(Konstante)	399844285,2	37471688,86		10,671	<,001	326293350,7	473395219,6					
	DCBD	-28106,481	3718,780	-,278	-7,558	<,001	-35405,852	-20807,109	-,210	-,254	-,250	,808	1,238
	Xr	-136725295	21925946,43	-,207	-6,236	<,001	-179762422	-93688167,2	-,189	-,212	-,206	,991	1,009
	DS	-128671,242	33089,767	-,143	-3,889	<,001	-193621,163	-63721,320	-,008	-,134	-,128	,804	1,244
4	(Konstante)	379834207,1	37747687,57		10,062	<,001	305741399,5	453927014,7					
	DCBD	-26340,062	3735,954	-,260	-7,050	<,001	-33673,156	-19006,969	-,210	-,238	-,232	,791	1,265
	Xr	-129579746	21905878,00	-,196	-5,915	<,001	-172577559	-86581933,1	-,189	-,202	-,194	,981	1,019
	DS	-158795,443	34152,397	-,177	-4,650	<,001	-225831,257	-91759,629	-,008	-,160	-,153	,746	1,341
	DM1000	77555978,02	23631692,08	,117	3,282	,001	31170662,07	123941294,0	,146	,114	,108	,847	1,181
5	(Konstante)	380555879,2	38065648,36		9,997	<,001	305838831,1	455272927,3					
	DCBD	-26224,926	3813,722	-,259	-6,876	<,001	-33710,680	-18739,172	-,210	-,233	-,226	,760	1,316
	Xr	-131940512	26838739,78	-,200	-4,916	<,001	-184620855	-79260169,1	-,189	-,169	-,162	,654	1,529
	DS	-157175,169	35787,793	-,175	-4,392	<,001	-227421,135	-86929,204	-,008	-,151	-,144	,680	1,470
	DM1000	77894437,67	23749726,37	,118	3,280	,001	31277355,76	124511519,6	,146	,114	,108	,840	1,191
	Xc	-5570646,902	36546951,42	-,007	-,152	,879	-77306725,1	66165431,29	,072	-,005	-,005	,587	1,705

a. Abhängige Variable: PP16

1.4.1.1 Model Description:

R-value represents the correlation between the dependent and independent variable. A value greater than 0,4 is taken for further analysis.

R-square shows the total variation for the dependent variable that could be explained by the independent variables. A value greater than 0.05 shows that the model is effective enough to determine the relationship.

The adjusted R-square shows the generalization of the results i.e., the variation of the sample results from the population in multiple regression. It is required to have a difference between R-square and Adjusted R-square minimum.

ANOVA[a]

Modell		Quadratsumme	df	Mittel der Quadrate	F	Sig.
1	Regression	3,761E+18	1	3,761E+18	38,141	<,001[b]
	Nicht standardisierte Residuen	8,164E+19	828	9,860E+16		
	Gesamt	8,540E+19	829			
2	Regression	7,006E+18	2	3,503E+18	36,955	<,001[c]
	Nicht standardisierte Residuen	7,839E+19	827	9,479E+16		
	Gesamt	8,540E+19	829			
3	Regression	8,415E+18	3	2,805E+18	30,097	<,001[d]
	Nicht standardisierte Residuen	7,698E+19	826	9,320E+16		
	Gesamt	8,540E+19	829			
4	Regression	9,407E+18	4	2,352E+18	25,533	<,001[e]
	Nicht standardisierte Residuen	7,599E+19	825	9,211E+16		
	Gesamt	8,540E+19	829			
5	Regression	9,409E+18	5	1,882E+18	20,407	<,001[f]
	Nicht standardisierte Residuen	7,599E+19	824	9,222E+16		
	Gesamt	8,540E+19	829			

a. Abhängige Variable: PP16

b. Einflußvariablen : (Konstante), DCBD

c. Einflußvariablen : (Konstante), DCBD , Xr

d. Einflußvariablen : (Konstante), DCBD , Xr, DS

e. Einflußvariablen : (Konstante), DCBD , Xr, DS, DM1000

f. Einflußvariablen : (Konstante), DCBD , Xr, DS, DM1000, Xc

1.4.2 Empirical Results

The empirical results are interpreted using P-values, T-Testing and R-squared methods and its statistical interpretation.

The model fit is described among others (such as ANOVA or DOE) by the R-squared values. The R-square can be described as the percentage of the response variable variation that is explained by a linear model. It is a statistical measure of how close the data points are to the fitted regression line. In theory, if a model could explain 100% of the variance, the fitted values would equal the observed values. Hence all data points would generate the fitted regression line.

The R-square cannot eliminate possible bias of the coefficients and predictions. To further test the explanatory power, we will check the residual plots.

The P-value can contribute to understanding the degree of conclusiveness of the chosen hypothesis in relation to the null-hypothesis of a population. A low p-value (Alpha-Niveau $p < 0,05$) indicates that we can reject the null hypothesis. A low p-value is meaningful to our model because it shows that the predictor variable is related to changes in the response variable. In our model the results estimate the p-value of the ANOVA table to be below the tolerable significance level of 0,05 therefore, our results are considered significant. Hence there is a possibility rejecting the null hypothesis in our further analysis.

We also evaluate the F-test (named after Sir Ronald Fisher) measuring the ratio of two variances. It shows the overall significance of our regression model.

Our first regression investigates the dependent variable property price per m²and the independent variables described in our model.

We can interpret the correlation between the dependent and independent variable represented by the R-value in our regression. Any value greater than 0,4 is taken for further analysis. We observe a high R-square of 50,52%, hence linear model chosen shows a good fit for our set of observations. All predictors, except distance to school, have a significant effect on the dependent variable.
Interestingly there is a low significance between plot price and distance to schools. The p-value is 0.8171 and above the cap of 0,5. We will test if distance to school remains not significant and can therefore be removed from the model.
The other chosen variables show a high significance and will remain in our model. We can see a higher t-value and therefore a higher significance if the plot is within the second distance band to Metro station.

Modell A: 1-18300 (n = 18140)
Removed observations: 16
Dependent Variable: l_Ppm2_16

	Koeffizient	*Std. Fehler*	*t-Quotient*	*p-Wert*	
Const	11,5296	0,271673	42,44	<0,0001	***
DM500	−0,170170	0,0422529	−4,027	<0,0001	***
DM501	−0,490686	0,0423268	−11,59	<0,0001	***
Xm	0,681648	0,0418059	16,31	<0,0001	***
Xc	0,489926	0,0507264	9,658	<0,0001	***
l_DS	0,00700454	0,0302753	0,2314	0,8171	
l_DH	−0,126869	0,0179259	−7,077	<0,0001	***
l_DCBD	−0,136090	0,0164666	−8,265	<0,0001	***

Mittel d. abh. Var.	9,849326	Stdabw. d. abh. Var.	0,654604
Summe d. quad. Res.	172,3691	Stdfehler d. Regress.	0,462447
R-Quadrat	0,505220	Korrigiertes R-Quadrat	0,500923
F(7, 806)	117,5725	P-Wert(F)	1,3e-118
Log-Likelihood	−523,2209	Akaike-Kriterium	1062,442
Schwarz-Kriterium	1100,057	Hannan-Quinn-Kriterium	1076,880

Since a dependency between price and plot size seems to be almost trivial, we will keep plot size but will change the dependent variable to log of price for 2016.

Modell B:
Dependent Variable: LNPP16
Log of property price 2016

	Koeffizient	*Std. Fehler*	*t-Quotient*	*p-Wert*	
Const	11,7631	0,326473	36,03	<0,0001	***
DM500	−0,166068	0,0423554	−3,921	<0,0001	***
DM501	−0,491257	0,0423118	−11,61	<0,0001	***
Xm	0,688889	0,0421650	16,34	<0,0001	***
Xc	0,506864	0,0523824	9,676	<0,0001	***
l_DS	0,00671784	0,0302637	0,2220	0,8244	
l_DH	−0,125980	0,0179318	−7,025	<0,0001	***
l_DCBD	−0,148135	0,0189301	−7,825	<0,0001	***
l_m2	0,979975	0,0155434	63,05	<0,0001	***

Mittel d. abh. Var.	17,21394	Stdabw. d. abh. Var.	1,552535
Summe d. quad. Res.	172,0145	Stdfehler d. Regress.	0,462258
R-Quadrat	0,912221	Korrigiertes R-Quadrat	0,911348
F(8, 805)	1045,717	P-Wert(F)	0,000000
Log-Likelihood	−522,3826	Akaike-Kriterium	1062,765
Schwarz-Kriterium	1105,083	Hannan-Quinn-Kriterium	1079,008

We find that distance to Metro in the second distance band (between 500m and 1000m) to have a higher statistical impact on property prices than within the 500m radius to metro. The drop in price per unit changes -0,16 within the 500m distance band and -0,49 units in the second distance band. Property prices have a higher negative impact if located further away from new metro stations.

The impact of distance to schools (education) has no statistical significance on property prices. This could well be explained due to the social fabric in Qatar. Due to the Qatar census Household size in Qatar reaches an average of 7.4 people in 2016. Most of the household's life in flats 42,9% and 40,5% reside in palaces and villas. It is usual to employ or "rent" a driver. This is reflected in employment field and number. 65.397non-Qatari work in transportation, i.e., 5.4% of the total number of employees. (Ministry of Development Planning and Statistics, 2016, Analysis of the Results of Population, Housing and Establishments Census , 2021). Hence the time dedicated to drive to and from school might not be a property selection criterion for Qatari Citizens.

Property utilization seems relevant. Residential properties are too small in numbers in the selected areas hence is removed from the model. Distance to medical care has higher impact shown in its coefficient, than distance to the CBD. This must be investigated further as the models imply a strong correlation between distance to CBD on property prices.

The corrected R-square is strong with 91%. Such a high percentage might be reached due to a very high significance for the predictor m² and must be acknowledged as such.

The means of dependent variable are small with 17.21 and indicates a great price homogeneity.

The following model will elaborate on the same model but with data collected in 2018.

Fehlende oder unvollständige Beobachtungen entfernt: 16
Abhängige Variable: LNPP18

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	11,8847	0,348203	34,13	1,34e-158	***
DM500	-0,174581	0,0451747	-3,865	0,0001	***
DM501	-0,516414	0,0451281	-11,44	3,36e-028	***
Xc	0,570663	0,0558691	10,21	4,09e-023	***
Xm	0,710250	0,0449715	15,79	3,79e-049	***
l_m2	0,977313	0,0165779	58,95	2,54e-294	***
l_DS	0,000257263	0,0322781	0,007970	0,9936	
l_DH	-0,134701	0,0191254	-7,043	4,04e-012	***
l_DCBD	-0,169458	0,0201901	-8,393	2,12e-016	***

Mittel d. abh. Var.	17,06660	Stdabw. d. abh. Var.	1,583435
Summe d. quad. Res.	195,6754	Stdfehler d. Regress.	0,493026
R-Quadrat	0,904006	Korrigiertes R-Quadrat	0,903052
F(8, 805)	947,6149	P-Wert(F)	0,000000
Log-Likelihood	-574,8363	Akaike-Kriterium	1167,673
Schwarz-Kriterium	1209,990	Hannan-Quinn-Kriterium	1183,916

The dependent variable is the log of price in 2018. The significance for distance to Metro in the 500m band remains significant. The change the coefficients show that the negative trend for land prices continued in 2018. It is slightly stronger for properties in the 500m-1.000m band. Also, the negative trend for properties in closer range to the CBD has increased.

In model C we use the difference of the logarithm property prices 2016 and 2018 to investigate the impact of our variables on the difference in fallen property prices.

ModelC
Abhängige Variable: D_lp_18_16
Dependent Variable: Difference in log price 2018-2016

	Koeffizient	Std. Fehler	t-Quotient	p-Wert	
Const	0,121591	0,0390479	3,114	0,0019	***
DM500	−0,00851320	0,00506595	−1,680	0,0933	*
DM501	−0,0251569	0,00506073	−4,971	<0,0001	***
Xm	0,0213611	0,00504317	4,236	<0,0001	***
Xc	0,0637992	0,00626523	10,18	<0,0001	***
l_DS	−0,00646057	0,00361971	−1,785	0,0747	*
l_DH	−0,00872053	0,00214475	−4,066	<0,0001	***
l_DCBD	−0,0213231	0,00226414	−9,418	<0,0001	***
l_m2	−0,00266179	0,00185907	−1,432	0,1526	

Mittel d. abh. Var.	−0,147332	Stdabw. d. abh. Var.	0,065937
Summe d. quad. Res.	2,460751	Stdfehler d. Regress.	0,055289
R-Quadrat	0,303821	Korrigiertes R-Quadrat	0,296902
F(8, 805)	43,91391	P-Wert(F)	1,56e-58
Log-Likelihood	1206,192	Akaike-Kriterium	−2394,384
Schwarz-Kriterium	−2352,066	Hannan-Quinn-Kriterium	−2378,141

We find that proximity to metro station is favourable to plots in the second distance band 500-1000m as the drop in land price is higher in the first zone 0-500m. Mixed use is affected less by the price reduction than purely commercially used buildings. Closeness to medical facilities is favourable over proximity to education and CBD.

We find that greater distance to the CBD reduces drops in land prices. This is very unusual and might indicate the transition of Qatar from a Mono Centric City with into a city with subcentres. As the metro stations where built new focal points where developed. As an example, the Station Katara is nestled in the middle of the new centre for culture and arts. Another example the new station Qatar University with its new facilities thus creating multiple centres.

The results of the regression will be verified by a visual inspection of the residual plots

1.4.3 Checking for Assumptions

In statistics you can test the assumption of normal distributed residuals. To see if the standard errors of the coefficients of a population over a specific time is constant or non-constant. If the fitted values increase, the variance of the residuals also increases. By producing residual plots, we check for heteroscedasticity (also defined as the absence of homoscedasticity). A distinctive shape such as a fan or cone shape is produced by heteroscedasticity and can be verified by a visual inspection.

(Building on the work of West, Finch and Curran (West, 1995) kurtosis of the residuals. density functions with a skewness of not higher than two and an excess kurtosis of maximum four can still be considered as sufficiently normally distributed.)

Diagram Modell 4:

Aufgenommene/Entfernte Variablen[a]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	DCBD, Xm, DM500, DH, DS, DM501[b]	.	Einschluß

a. Abhängige Variable: LNPP16

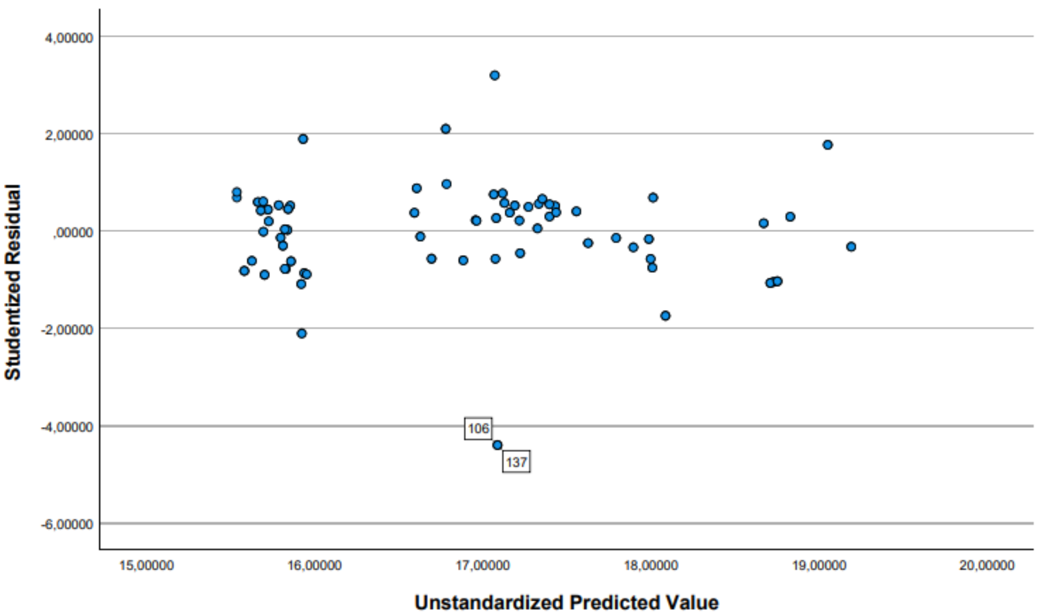
b. Alle gewünschten Variablen wurden eingegeben.

Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,677[a]	,458	,433	1.120717949

a. Einflußvariablen : (Konstante), DCBD, Xm, DM500, DH, DS, DM501

b. Abhängige Variable: LNPP16



In the above Model the R-value is 0.677 which is greater than 0.4, hence taken for further analysis. The adjusted R-square and Adjusted R-square minimum is required to have a difference to account for a possible generalization of the results. Here we have an R-square of .458 which is not far from the Adjusted R-square of .433, which is good. The R-square value is .458 and therefor below the significant value of .5, hence the model does not explain the total variation of the dependent variable by the chosen independent variables. We adjust the model as shown in Model 6.

Diagram Modell 6

Regression : Modell 6 1

Aufgenommene/Entfernte Variablen[a]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	DCBD, Xm, m2, DM500, DH, DS, DM501[b]	.	Einschluß

a. Abhängige Variable: LNPP16

b. Alle gewünschten Variablen wurden eingegeben.

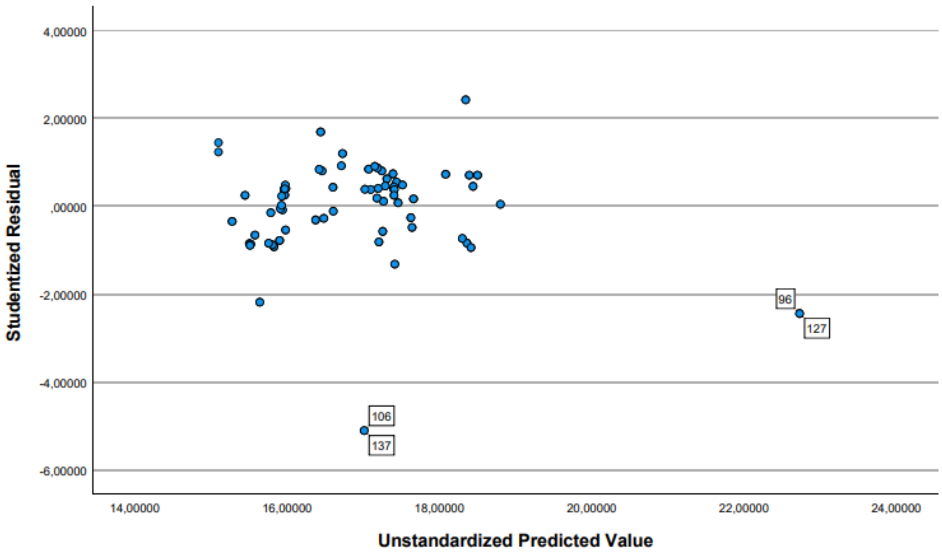
Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,783[a]	,613	,592	,9505522270

a. Einflußvariablen : (Konstante), DCBD, Xm, m2, DM500, DH, DS, DM501

b. Abhängige Variable: LNPP16

Diagramm: Modell 6 1 (2016)



This time the chosen model shows an R-squared value of .613 which shows that the model is effective enough to determine the relationship between the dependent variable (logarithm of property price 2016) and the independent variables (Distance to Central Business District, Property Type Mixed Use, Distance Metro 500m, Distance Hospital, Distance School).

We repeat the Model for the second period in 2018.

Diagram Modell 6 2

Aufgenommene/Entfernte Variablen[a]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	DCBD, Xm, m2, DM500, DH, DS, DM501[b]	.	Einschluß

a. Abhängige Variable: LNPP18

b. Alle gewünschten Variablen wurden eingegeben.

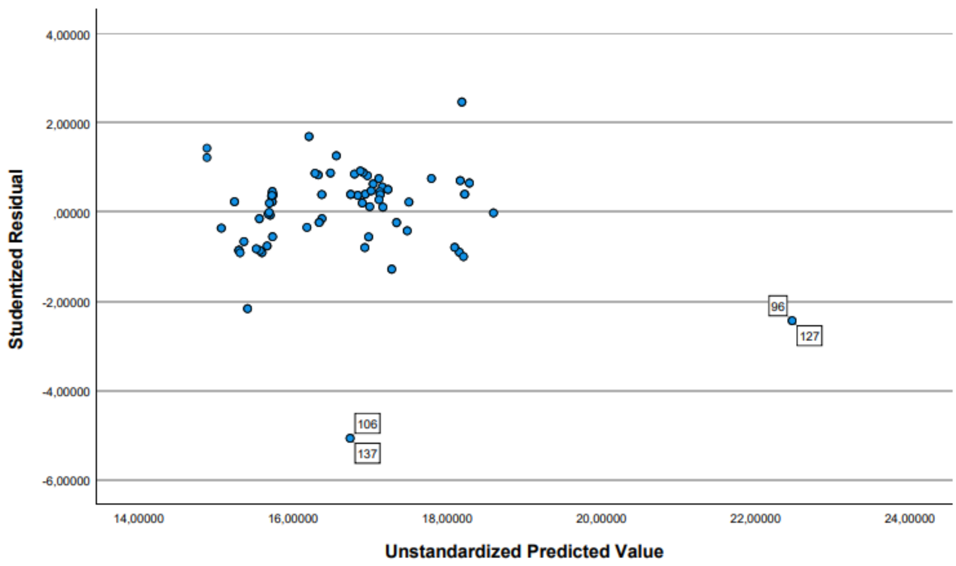
Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,781[a]	,609	,588	.9548442747

a. Einflußvariablen : (Konstante), DCBD, Xm, m2, DM500, DH, DS, DM501

b. Abhängige Variable: LNPP18

Diagramm: Modell 6 2 (2018)



We find a strong explanatory R-square value of .609 for the chosen dependent variable and independent variables for 2018. The difference of R-square to R-square adjusted is 21 which is not far off and will be considered.

Diagram Model 6 3

Regression : Modell 6 3 (2018-2016)

Aufgenommene/Entfernte Variablen[a]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	DCBD, Xm, m2, DM500, DH, DS, DM501[b]	.	Einschluß

a. Abhängige Variable: DLP_18_16

b. Alle gewünschten Variablen wurden eingegeben.

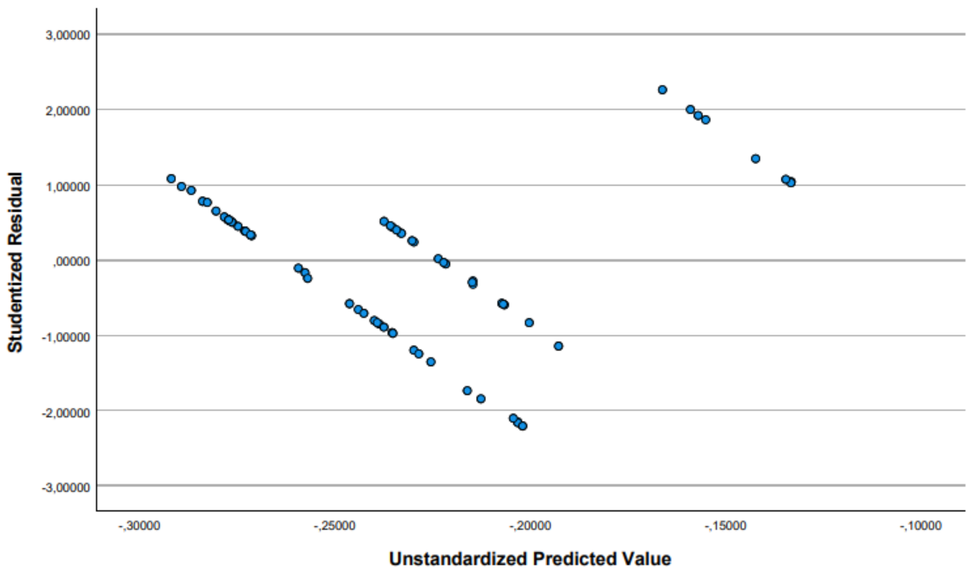
Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,833[a]	,694	,677	,02834

a. Einflußvariablen : (Konstante), DCBD, Xm, m2, DM500, DH, DS, DM501

b. Abhängige Variable: DLP_18_16

Diagramm: Modell 6 3 (2018-2016)



For this model we use the difference in log. prices for 2016 and 2018 as the dependent variable and kept the independent variables from the previous models.

Diagram Model 7

Regression : Modell 7

Aufgenommene/Entfernte Variablen[a]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	m2, DM500, DS, Xm, DH, DM501, DCBD[b]	.	Einschluß

a. Abhängige Variable: LNPP16

b. Alle gewünschten Variablen wurden eingegeben.

Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,783[a]	,613	,592	.9505522270

a. Einflußvariablen : (Konstante), m2, DM500, DS, Xm, DH, DM501, DCBD

b. Abhängige Variable: LNPP16

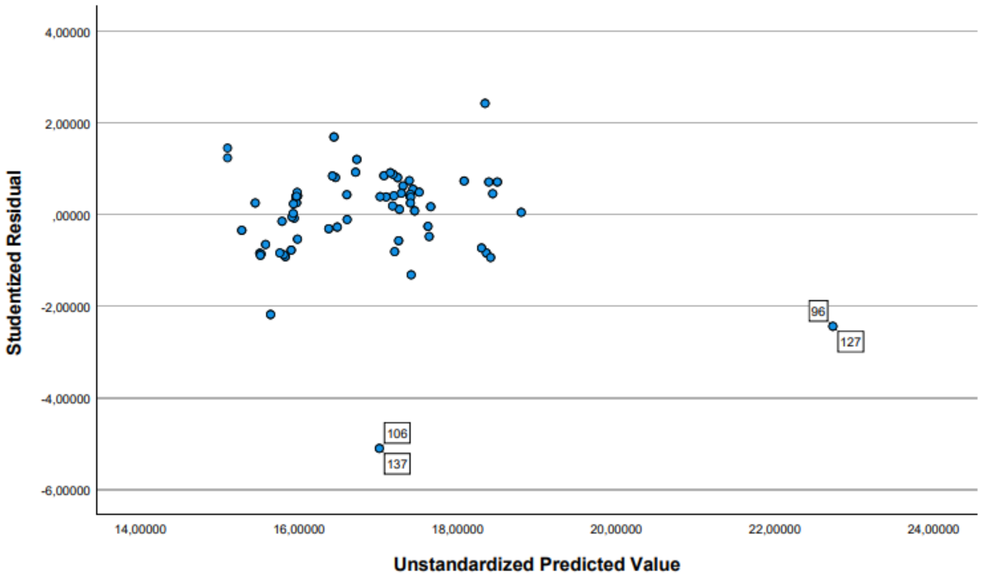


Diagram Model 9

Regression : Modell 9

Aufgenommene/Entfernte Variablen[a]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	m2, DM500, DS, Xm, DH, DM501, DCBD[b]	.	Einschluß

a. Abhängige Variable: DLP_18_16

b. Alle gewünschten Variablen wurden eingegeben.

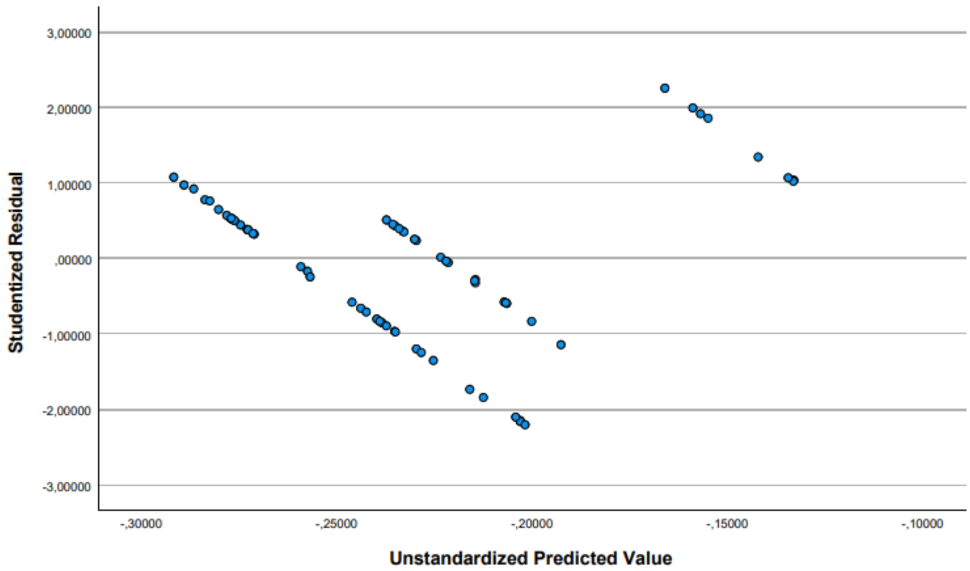
Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,833[a]	,694	,677	,02834

a. Einflußvariablen : (Konstante), m2, DM500, DS, Xm, DH, DM501, DCBD

b. Abhängige Variable: DLP_18_16

Diagramm: Modell 9



1.4.3.1 Limitations and possible bias

The hedonic pricing model and its function is commonly used in monocentric cities. In the case of Qatar, it is correct to use the approach as the CBD is the centre of the city. As the city evolves, triggered by the 2022 World Cup and its Vision 2030 the city will transform into a multi-centre city with subcentres. This must be included in future models.

As building prices seem to have a measurable impact, they should be included in models on property prices. This work is limited to data obtained during the phase of announcement and soft launch. It would be desirable to extend the work with data collected during the year of opening and after consolidation i.e., five years' time.

There is also the possible bias with regards to the blockade-imposed June 5th on Qatar by Saudi Arabia, the United Arab Emirates, Bahrain, and Egypt among other Countries. This might have had an impact on business and/or trade of the land plots. The blockade is declared over on 6th January 2021. During the time of blockade there was limited trade between these countries and might have influenced property values.

Limitations due to model data. We have used panel data obtained at two points in time. Mathematical: Linear covariates perform better in a regression than logged covariates. They yield similar conclusions as quadratic covariates in a model.

1.4.4 Omitted variable bias

How to control for potential omitted variable bias? In a study in the United States by Sirmans et al. he found a total of 360 independent variables. The variables are classified into eight categories: construction and structure, house features, house amenities, environmental, neighbourhood and location, public services, marketing, occupancy, financial and fiscal issues. (Sirmans, Macpherson, & Nietz, 2005)

To control for inconsistencies or possible biased estimates of the effect of infrastructure on property values researchers used accessibility to employment centres (Billings, 2011) and accessibility to shopping centres. This study uses the variable distance to CBD as it is home to the highest density of shopping and employment.

1.4.4.1 Unobserved heterogeneity

Unobserved heterogeneity, hence, a biased linear regression model, arises whenever the explanatory variables do not account for the full amount of individual heterogeneity in the conditional mean of the dependent variable (Winkelmann, 2008).

Bias testing in three steps:

I	"The usual" written as	A	B →
II	"The reversed" written as	A	B ←
III	"Simultaneity" written as	A	B ⟺

1.4.5 Checking for Reversed Causality

With checking for reversed causality, we discern if the cause stated in the hypothesis is the cause, or the effect.

Hypothesis statement I: Proximity to new metro stations has a (positive/negative) impact on land values in Doha, Qatar.

A = (proximity) metro station causes B = high land values in Doha

B = high land value causes A = metro stations

A causes B and B causes A = metro stations cause high land values <u>and</u> high land values cause metro stations

Higher (positive) land values account for the construction of new metro lines (metro stations).

The alignment of the metro had to serve the Qatar vision 2030. The Qatar national vision 2030 aims to direct Qatar towards a balance between developmental needs and the protection of its natural environment, whether land, sea, or air. It provides the roadmap towards Qatar becoming an advanced society capable of sustainable development with the goal of providing high standard of living for all citizens by the year 2030. (General Secretariat for Development Planning, 2008)

The alignment process is driven by the vision. It must connect without disturbing the city's fabric. This results in the rational to use mayor roads and their land area on a subsurface level. This minimizes disruption on existing buildings where possible.

City planning and mobility studies are used to create hubs and transportation nodes. Last technical boundaries such as Qatar's special nature of the ground contributed to the now built alignment. Last, the awarded 2022 World Cup raised the need to connect the relevant soccer stadiums.

We conclude that we can foreclose the possibility of reversed causality.

1.4.5.1 Checking for Simultaneity

With the argument above we discern that high land prices account for the construction of a Metro station.

1.4.6 Choice of functional form for Hedonic Price functions

Another bias can arise from the chosen functional form for the HP function.

Hedonic pricing models use 6 functional forms: semi log, linear, quadratic, double-log, quadratic box-cox, and linear box-cox (UKEssays, 2018). In the work of Cropper et al. (Cropper, Deck, & McConnell, 1988) they highlight that linear and quadratic Box-Cox forms produce lowest mean percentage errors: however, when some attributes are unobserved or are replaced by proxies, linear and linear Box-Cox functions perform best.

Also, simultaneity bias occurs because of the non-linearity of the hedonic function. It arises in implicit markets for individual characteristics/attributes because of the nonlinearity in the market (hedonic) price function for housing. The standard method for dealing with is to apply the two-stage least squares method.

The semi-log model is used to deal with non-linearities in the hedonic pricing model as the land and/or real estate used in our model cannot be restocked in the market. The semi-log structural model permits the value of a given characteristic to vary proportionately with the value of another characteristic.

We use a log model to account for land being a non-linear utility with spatially delineated amenities.

1.4.7 Plausibility Check of the Regression Coefficients

Some researchers argue that Rosen´s global model does not account for spatially varying relationships such as transport accessibility and land value. They argue to expand the model by using Geospatial data in a geographically weight regression (GWR).

Impact of New Metro Stations on Surrounding Land Values in Doha, Qatar

We contribute by presenting for the first-time correlations between land value and variables used in HP function. To our knowledge this is the first analysis using a unique dataset showing the impact on land value related to the new metro in Doha.

We contribute by establishing a correlation between land value and different catchment zones from new metro stations in Doha, Qatar.

We can establish a correlation between stable land values within the 500m to 1.000m catchment zone and the new Metro.

We show a correlation between land value and maturity of metro project i.e., during the chosen time segments of construction.

1.5 Comparison of the Empirical Findings

We compare the empirical findings according to the three types of land use (commercial, residential, and mixed), second according to the three distance bands and last the two-time segments.

First: The analysed data shows a drop in land values during both periods under observation.

Second: The drop in land values is most significant for residential and commercial land plots.

Third: Mixed use land plots are more resilient.

Fourth: The plots within the 500m to 1.000m distance to Metro show a significant lower drop in land value.

Fifth: Distance to CBD shows no significant correlation to land values.

Sixth: Distance to education and medical facilities is significantly correlated to land value

1.6 Conclusion

The presented research examines the impact of the new Metro system with its new Metro stations in Doha, Qatar on property values. The property values are divided into commercial, residential, or mixed use. We use highly detailed GIS data and geographical information on a hedonic pricing function to potentially explain the variations in land values. The research is based on detailed information on the Doha Metro “Red Line” development and is analysed using a set of three distance buffers as well as two distinct project milestones “metro announcement” and “metro soft launch”. We use the ordinary-least-square panel regression model where each point in time subsumes appr. 2.000 data points per station and additional externalities on land values over a fixed period in 2016 and 2018 respectively. The information on land values was granted by the Ministry of Municipality in Doha and must be treated confidential.

The findings indicate that plot land values within the first 0-500m distance to the new Metro stations decreased between 2016 and 2018. We can see a stronger negative trend for residential, and commercial plots.

To evaluate the above, we must contextualize the findings on Real Estate Values within the 2016- and 2018-time trend. Therefore, we include the Real Estate Price Index for the State of Qatar published by the Qatar Central Bank (QCB), based on the data issued by the Ministry of Justice, for the relevant period 2016 to 2018.

We acknowledge a decline from 292.2 in 2016 down to 246.5 by the end of 2018; a decline of 45,7 points. This drop in real estate and land prices may be connected to international developments such as the Brexit and US elections (the Qatar Rial is backed by the US Dollar) as well as national developments i.e., a sharp fall in oil prices in the early 2016 resulting in fiscal and current account deficits. At the same time the return on investments declines from 16.15% in 2015 to 14.49% in 2016. The inflationary situation sustained in 2018 due to geo-political tensions. The GDP in 2016 of 3.1% is stable on a low level and is 2018 at 3.2%. (Qatar Central Bank, Financial Stability Review 2018, 2018) (Qatar Central Bank, Financial Stability Review, 2016)

The above negative time trend for land values in Doha during the time of observation might confound with the regression estimates and must not be ruled out. With the introduction of multiple distances and two time periods we aim to control for time trends. To consolidate regression estimates the presented panel data could be extended using further longitudinal data.

The presented regression finds no strong correlation between land value and distance to CDB. This is in line with findings by Berawi (Berawi, Miraj, & Saroji, 2020) and Anderson (Anderson & West, 2006) for example.

The introduced HPM is based on the logic of a monocentric city and the concentric Burgess Model. The Burges Urban Land Use Model assumes that the greater the distance from the CBD, the smaller the economic rent value. (Rodrique, 2020). Incorporating the future developments in Doha the model should be augmented to the model of a polycentric city using transportation lines and sectors as proposed by Homer Hoyt. (Beauregard, 2007)

Looking at our model we find that distance to education and medical facilities has a higher influence on site value. This might be due to Qatar's specific census. Expatriates being a majority over Qatari citizens seem to value travel-time to education over travel time to work.

This might also explain why direct or walking distance to metro stations (0m - 500m) is not valued. School runs are mostly done by car and not by public transport. Hence walking distance to metro stations proximity is more a convenience highly valued in the second distance band 500m -1.000m from metro stations.

It shows that hedonic pricing models need to have a stronger socio-economic view specifically tailored to its unique cultural setting as intra-metropolitan influences on land values vary accordingly.

For further research on the topic of hedonic pricing models we would encourage to include other site-specific census-track influences on intra-metropolitan commercial, residential, and mixed-use units.

A key finding of the presented regression is that the assumption that spatially centralized workplaces and therefore accessibility to the CBD is a major determinant for location specific land values is not correlated in the case of Doha, Qatar.

The findings show that land values overall remain more stable in the second distance band to Metro stations i.e., 500m – 1.000m. This is in line with findings in Jakarta and Dubai and other studies suggesting negative correlation due to Metro proximity and land uplift.

Being in the second distance band might, in the case of Doha, include the commodity but exclude the negative impacts. We assume this to coincide with findings by Higgis and Kanaroglou that proximity to rail can affect land value in one positive and two negative ways. "Being located very close to a transit station or along a transit line can have negative impacts, including noise and air pollution from trains, and increased local vehicle traffic from transit passengers with parked cars. These nuisance effects may reduce property values." (Higgis & Kanaroglou, 2018)

Taken the results the study asserts that that mixed-use land plots seem more resilient to value decay than commercially and residentially utilized land plots.

1.7 Further research

Moving forward from our results we would suggest future studies on the impact of the metro in Doha and to adopt a polycentric approach using multiple nodes as suggested by the Polycentric Research Group or Momer Hoyt (E.Heikkila, Gordon, J.I., Peiser, & Richardson, 1989)

In addition, the data from the neighbouring new metro projects in Oman and Saudi-Arabia will provide additional data to understand trends for value uplift due to proximity to new metro systems in middle eastern countries.

As the results show no uplift on land values in walking distance to metro stations but positive effects on land values in the second and third distance bands, we add to research how to monetize this uplift and recoup construction costs.

The following section introduces value capture as a strategy to monetize land value to finance public infrastructure. The various strategies are introduced as well as the relevant stakeholder using data from Singapore, Hong Kong, and Japan. Concluding we present a land value strategy for Qatar in the form of a case study applying data from Qatar Rail on Transit Oriented Developments in the chapter “Financing transit systems through value capture using TOD´s”.