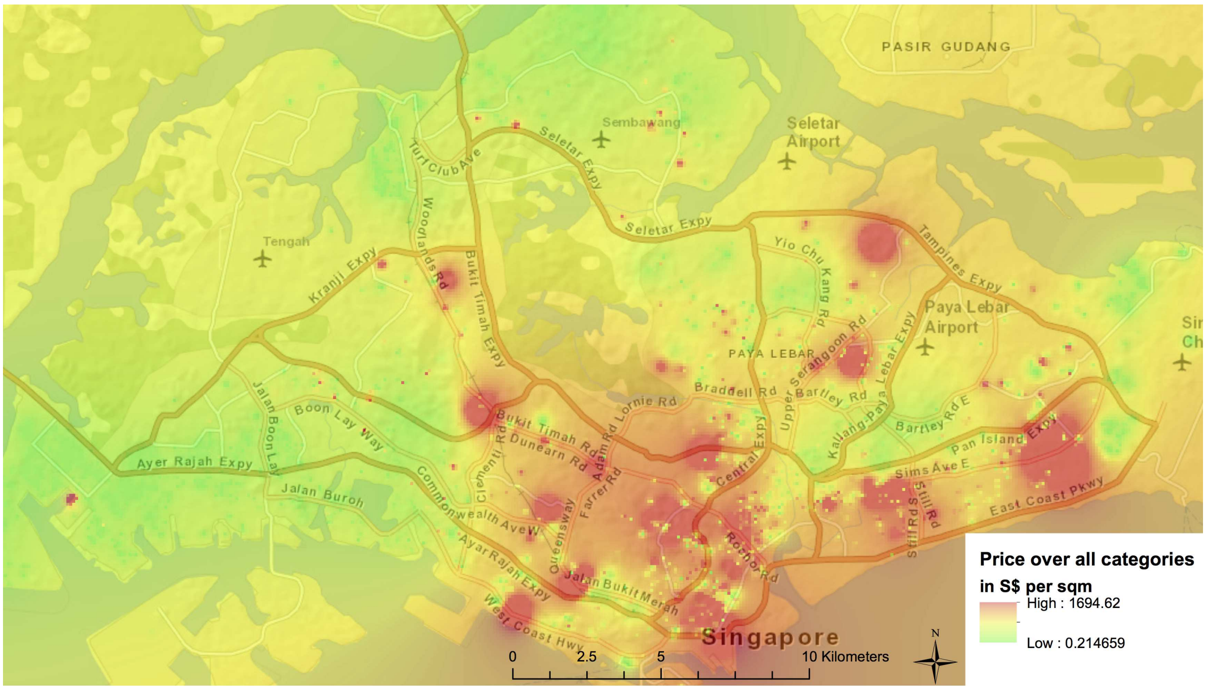


Spatial hedonic regression modelling of commercial and office space prices in Singapore

Luzian Caduff

Supervision:
Prof Dr. K.W. Axhausen, Institute for Transport Planning and Systems, ETH Zurich
M.A.B. van Eggermond, Future Cities Laboratory, Singapore ETH Centre
Dr A. Erath, Future Cities Laboratory, Singapore ETH Centre

Master Thesis
Spatial Development and Infrastructure Systems **July 2013**

(SEC) SINGAPORE-ETH CENTRE 新加坡－ETH 研究中心

(FCL) FUTURE CITIES LABORATORY 未来城市实验室

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

IVT Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Acknowledgements

I would like to thank Professor Kay W. Axhausen for his support, guidance and the valuable comments and suggestions given for my thesis. He gave me the opportunity to write my master thesis at the Future Cities Laboratory in Singapore.

I thank my supervisor Michael van Eggermond for the valuable inputs, the proof reading and the support during the whole thesis. He collected a part of the data and has calculated the accessibilities used in this work. Moreover, I would like to thank Alexander Erath, the module leader of the Future Cities Laboratory in Singapore, for his inputs and recommendations for my work. I also would like to express my gratitude to Anne-Kathrin Bodenbender for proof reading my thesis.

Many thanks go to the team of Module VIII at the Future Cities Laboratory for the support and the great time in Singapore. Further, I would like to thank Nikolaos Theodoratos and Kamila Oglecka, who provided an accommodation during my visit in Singapore and welcomed me warmly.

I am also very thankful to Professor Lino Guzzella and the Erich Degen-Stiftung of ETH Zurich for the generous travel allowance.

Furthermore, I would like to acknowledge the support provided by my family and my friends during my whole studies at ETH and this thesis: Rahel Buss, Anne-Kathrin Bodenbender, Ursin Gaudenz, Beat Sonder and Franziska Tscholl.

Master Thesis

Spatial hedonic regression modelling of commercial and office space prices in Singapore

Luzian Caduff,

Spitalstrasse 25
CH-8952 Schlieren
luziancaduff@gmail.com

July 2013

Abstract

In this thesis a hedonic regression model for the office, retail and industrial space prices in Singapore is developed and estimated. The data used for the regression analysis is collected from a commercial and a governmental online platform. Therefore asking and transaction prices are available for the office, retail and industrial market.
This study uses a hedonic modelling approach for estimating the price determinants. In addition to the linear ordinary least squares approach, spatial models are used. Different spatial modelling approaches are tested and evaluated to find the best performing model. The influence of the public transport access and the location on the property's price is investigated in detail.
The major price determinants are the property's tenure conditions, the access by public transport and the floor area. The office and industrial market show some clustering behaviour. The spatial autocorrelations in the data can be removed completely by using spatial hedonic approaches. Spatial error models (SAR_{err}) performed best and are recommended to use in further studies. The results can be used at the Future Cities Laboratory to model location choice of firms in Singapore by using an agent based transport simulation tool (MATSim).

Keywords
Spatial hedonic regression, office, retail, industrial, transaction prices, asking prices, Singapore

Preferred citation style

Caduff, L. (2013) Spatial hedonic regression modelling of commercial and office space prices in Singapore, *Master Thesis*, Institute for Transport Planning and Systems (IVT), ETH Zurich, Zurich.

Masterarbeit

Spatial hedonic regression modelling of commercial and office space prices in Singapore

Luzian Caduff,

Spitalstrasse 25
CH-8952 Schlieren
luziancaduff@gmail.com

Juli 2013

Zusammenfassung

Diese Masterarbeit befasst sich mit der Regression von hedonischen Preismodellen für kommerzielle Nutzungen in Singapur. Für diese Arbeit stehen Transaktions- und Angebotspreise zur Verfügung, welche von Online-Plattformen bezogen werden können. Die Daten sind unterteilt in Büro, Handel und Industrie.
Für die Bestimmung der Einflussstärken der verschiedenen Faktoren wird ein hedonisches Preismodell erstellt. Einerseits werden einfache lineare Modelle geschätzt, und andererseits verschiedene räumliche Modelle getestet. Die Einflüsse der Arbeitsplatz-Erreichbarkeit, vom Zugang zu den öffentlichen Verkehrsmitteln und von verschiedenen Clustern werden detailliert untersucht und beschrieben. Die wichtigsten Einflussgrössen auf die Immobilienpreise für Büro, Handel und Industrie sind das Pachtverhältnis, der Zugang zu den öffentlichen Verkehrsmitteln und die Grundfläche. Durch die Verwendung von räumlichen Regressionsmethoden konnte die Modellgüte erhöht und die räumliche Korrelation der Daten entfernt werden. Die besten Ergebnisse lieferten die räumlichen Fehlermodelle (SAR_{err}). Die Ergebnisse können am Future Cities Laboratory verwendet um die Standortwahl von Firmen zu modellieren.

Schlüsselwörter

Räumliche hedonische Regression, Büro, Handel, Industrie ,Transaktions- und Angebotspreise, Singapur

Bevorzugter Zitierstil

Caduff, L. (2013) Spatial hedonic regression modelling of commercial and office space prices in Singapore, *Masterarbeit* , Institut für Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

A firm's location choice depends on many factors. A location choice is influenced by different soft and hard location factors. As the firm has chosen a suitable region, the specific location has to be found. For this decision the properties' prices and availability are important to consider. To understand the location choice behaviour better, an understanding of the price components and factors are essential. This thesis concentrates especially on the location choice within a city.

The research module on Mobility and Transportation Planning at the ETH Future Cities Laboratory (2013) (FCL) are implementing a land-use and transportation model for the city of Singapore. The MATSim (2013) model used is an agent-based simulation and works at the building resolution. This means that each single building of the city of Singapore is modelled in the model. Future extensions of MATSim (2013) can even be more detailed. This model shall in near future include the residential and commercial location choice. For this purpose the property's price needs to be calculated for the different scenarios. To gain more understanding on the price determinants a hedonic property price model is estimated. For residential properties a study done by Lehner (2011) at the FCL shows the price determinants for the city of Singapore. No recent studies are available for the commercial property sector of Singapore. This work shows the price determinants for the commercial market.

1.2 Thesis aim

The main aim of this thesis is to answer the question how the price of a commercial property in Singapore is composed. This thesis investigates especially the effects of the public transport access and accessibility on the price, as well as the effect of different locations and clusters in Singapore. According to Lancaster (1966) and Rosen (1974) the price of a property is composed of different characteristics. Lancaster (1966) describes that the price of a good is composed by the different characteristics of the good. Rosen (1984) formulates the first models and approaches to implement this theory.
The properties' characteristics vary between the different countries, for instance of differences in climate, population ethnicities, governmental regulations and more. For example the presence of a heating system in the property is not important at all in tropical countries like Singapore, whereas it is in most European countries. To investigate the preferences of the tenants of commercial properties in Singapore, a hedonic price model will be estimated. The model should

not only answer which characteristics are important in Singapore, but also the relative impact of each parameter. The price models are created for the office, retail and industrial market, where each market has different characteristics and therefore also different price determinants. Such differences will be shown and analysed in this thesis.

1.3 Thesis structure

Literature is reviewed in Section 2 to investigate important price determinants for commercial usages. This section gives also an introduction to the real estate market of Singapore. Section 3 introduces the methodology of the hedonic regression and presents the different model types used in this work. Section 4 describes the available data and the calculations done to create some of them.

Section 5 shows the descriptive statistics and all the results for the different models. This chapter also shows an overview of the different models regressed. The results are discussed and analysed in Section 6. Lastly, the conclusions and recommendations for future research are given in Section 7.

2 Literature review

The literature review gives an overview of possible important variables for the models, which are later introduced for the case study of Singapore. Section 2.1 gives an overview of different hedonic price studies done earlier. The section is subdivided according to the three main commercial classes: office, retail and industrial. The main model types, which will also be used later in this study, are described in chapter 3. Section 2.2 introduces the characteristics of the Singapore property market. Special characteristics are important for using in hedonic regression models.
Section 2.3 summarizes the findings of the literature review and shows the most important price determinants for the commercial market. Section 2.4 proposes hypotheses for price determinants in Singapore, based on the literature and market review.

2.1 Hedonic regression

Hackney *et al.* (2013) distinguish land market models into hedonic and bid auction models. In a hedonic model the property's price is described by its attributes and the locational characteristics. The main weakness of hedonic property price models is its complexity. If variables like accessibility or agglomeration benefits are included, complexity rises, which makes the analysis of the model results more difficult.
Bid auction models maximize the surplus of a consumer. The surplus is the consumer's income minus the rent he has or is willing to pay for the property at a certain location. The rent or price a consumer has to pay for a property is the expected maximum bid across all consumers. Each consumer tries to maximize its utility by varying the location. In this study only hedonic regression models will be used.

Dunse and Jones (1998) describe offices as heterogeneous goods. This means that offices consist out of different attributes, which make each office unit unique on the market. The different attributes are the characteristics of the office. This includes the office itself (with the interior and the quality of the building), the location in the city, the surroundings of the building and other things, depending on the focus of the analysis. With a hedonic regression the value of such attributes can be estimated by using existing data. The result of an hedonic analysis is a function, which describes a dependent variable (normally in the office market the rent or the sale price) with its corresponding attributes.
The same concept of a heterogeneous good can also be applied to retail and industrial properties. Each category has its own different price determinants, but some of the attributes will be found in all three market segments.

2.1.1 Hedonic office price regression

Compared to hedonic regression studies of housing prices, the literature on office prices is relatively scarce. Downs and Slade (1999) ascribe this fact to problems of collecting and to the reliability of data for office markets. Table 1 shows an overview of the studies presented in the next section. Tables 2 and 3 show the attributes analysed in these studies with the expected sign.

Table 1: Overview of hedonic office studies

Citation	Study area	R^2	# obs.	Used model	Type*
Haase (2011)	Zurich	0.62 – 0.67	1010	Linear model	Sqm
Farooq *et al.* (2010)	Toronto	0.44 – 0.47	1302	Linear, box-cox	Sqm
Chalermpong and Wattana (2009)	Bangkok	0.55 – 0.59	85	Spatial (log) linear	
Jennen and Brounen (2009)	Amsterdam	0.53 – 0.56	1465	Log linear	
Ozus (2009)	Istanbul	0.5	94	-	Sqm
Kim (2007)	Seoul	0.54 – 0.82	731	SAR, SEM, SAC	Sqm
Nappi-Choulet *et al.* (2007)	Paris	0.89	2587	Linear/ OLS/WLS	
Enström and Netzell (2007)	Stockholm	0.97	184	Translog model	
Nitsch (2006)	Munich	0.45 – 0.86	46	Linear	Sqm
Tu *et al.* (2004)	Singapore	0.73 – 0.8	2950	BSTAR, 2BSTAR	
Nagai *et al.* (2000)	Tokyo		2497	Box-cox, linear	Sqm
Dunse and Jones (1998)	Glasgow	0.79	477	-	Sqm
Bollinger *et al.* (1998)	Atlanta	0.63	-	-	Sqm

* Sqm: used the price per square meter or feet as dependent variable

The first hedonic office price regressions were undertaken in the United States around the 1980s. Haase (2011) gives an overview of different studies done in the past. A classified listing of the used variables as well as the used dependent variable are included in the overview. A distinction is made between transaction and asking rents. The most dominant variables are according to the overview of Haase (2011) the location, the size and the age of the property, but also spatial variables are increasingly analysed. Tonelli *et al.* (2004) list in their study the literature of hedonic office price models. They show detailed which study includes which determinants and list the used model and equations for these studies. Nappi-Choulet *et al.* (2007) give also a brief overview. In the listing the area, the building age and the distance to the CBD is shown.
Table 1 shows that the R^2 is much higher for studies which consider only the price as dependent variable and the floor area as an independent variable.

Table 2: List of hedonic office rent studies with the significant structural and contract variables and the estimated coefficient sign

	Haase, 2011	Farooq, 2010	Chalermpong, 2009	Jennen, 2009	Ozus, 2009	Kim, 2007	Nappi-Choulet, 2007	Enström, 2007	Nitsch, 2006	Tu, 2004	Dunse, 2002	Nagai, 2000	Dunse, 1998	Bollinger, 1998
Dependent variable*:	T	A		A	A		T	T	A	T	A		A	
Structural variables:														
Building age:	−		−	−	o	+	−			−	−	−	−	−
Total building floor size:	−			+		o						+		+
Number of floors:					−									+
Vacancy rate:	o	−			−									
Elevator:	+		+								o			
Central air conditioning:			−			o			+		+		+	
Indoor parking:	+				o					+		+	+	+
Contract variables:														
Contract size of office:	+/−		−		−	+		+		+	+		+	
Ratio contract/building size:				+		−								
Contract duration:	+/−							+						
Leasehold:										−				

*A = Asking rent, T = Transaction rent; +/−: coefficient sign, o: not significant

Table 3: List of hedonic office rent studies with the significant locational and neighbourhood variables and the estimated coefficient sign

	Haase, 2011	Farooq, 2010	Chalermpong, 2009	Jennen, 2009	Ozus, 2009	Kim, 2007	Nappi-Choulet, 2007	Enström, 2007	Nitsch, 2006	Tu, 2004	Dunse, 2002	Nagai, 2000	Dunse, 1998	Bollinger, 1998
Locational variables:														
Distance to CBD:		−	+		o	−		+	−		−	−	−	
Distance to transit station:	o		−	−		−			−			−		−
Distance to highway:				+	o									+
Proximity to office-cluster:		+	+	+					+					
Distance to lake/sea view:	−				−									
Belonging to commercial area:						+/−								
Neighbourhood variables:														
% Residential in neighbourhood:		−												
% Industrial in neighbourhood:		−												
Micro-location:	+													
Accessibility public transport:	+													
Accessibility private transport:	+													
Accessibility multi mode:					+									
% of pop. with tertiary education:	+													

+/−: coefficient sign, o: not significant

The study of Haase (2011) analyses in a first step the office market in the canton of Zurich with traditional hedonic approaches. In a second step a multilevel modelling approach is used for the hedonic analysis. In the traditional model the location quality variables have the biggest impact on the price. Detailed locational variables as the distance to public transit or the accessibilities are not significant. For the city or Zurich the distance to the lake and the local provision were significant. Further the building age, the rented space and the condition of the building are the main variables for describing the building quality. The condition variables gathered are very basic, because of limited data availability. The real quality of the building is therefore only poorly represented in the model. The vacancy rate is not significant in any model. Through the use of multilevel models the model fit could be improved. The study recommends the use of multilevel modelling approaches, especially when the used data has a hierarchical structure (for example data gathered on the municipality level). The spatial autocorrelation can also be taken into account with the multilevel approach. The use of spatial-lag models should be considered if the required variables are available for a study.

Farooq *et al.* (2010) analyse the office market in the city of Toronto. The study considers an office clustering effect in the analysis. This is reached by introducing the distance to the CBD and by using some dummy variables for different sub-centres. The dependent variables are the asking rent of the office units. As independent significant variables were the quality of the building, effects of business nodes, vacancy rate, the surrounding land usage, vicinity to regional transit stations and subway stations and the vicinity to highway interchanges for suburbs. The effect of vacancy rate on the asking rent is rather small. The authors suggest to use instead of the vacancy rate, the deviation of the current vacancy rate from the vacancy rate of the business node.

Jennen and Brounen (2009) investigate the effects of clustering on office asking rents in the city of Amsterdam. The office clusters are calculated in the first model using a sharp threshold function and in the second model by using a distance decay function. The models include the different cluster types, the building size, the contract size, the age, the distance to the train station and the CBD, a contract building size ratio and different dummy variables for the different sub-centres in Amsterdam. The proximity to train stations generally generates higher rents, but the proximity to highways tends to lower rents. The results show, that the clustering effect generates higher office rents.

Ozus (2009) applied an hedonic office asking rent model for the city of Istanbul. The significant independent variables are: number of floors, vacancy rate, office floor space, architectural quality, banks in the environment, restaurants in the environment, social facilities and accessibility across modes. Sea view is not significant in this study. The author explains this fact with special zoning regulations, which ban high-rise buildings near the sea. Furthermore age, parking availability and distances to the CBD, to sub-centres and to highways are not significant.

The study of Nappi-Choulet *et al.* (2007) creates an hedonic office price model for Paris and its suburbs. The dependent variable is the transaction rent. The study includes locational variables, which describe the different office centres of Paris. Further structural variables and temporal variables are tested. The building age does not have a significant influence on the price.

Rebelo (2011) gives a comprehensive overview of which independent variables are used in a range of studies. Traditionally the main variables which describes office rents are the vicinity to the CBD, nearness to other businesses especially banks and insurances, amenities, accessibility to transport infrastructure, status (prestige) of the location and planning and fiscal tools. Centrality is not such an important factor any more; offices are now situated in suburbs. This evolution leads to poly-centric cities with some suburban office clusters.

2.1.2 Application of hedonic regression in South-East Asia in office price estimation

Tu *et al.* (2004) apply an spatio-temporal approach to the Singapore's office market. The study considers the period of 1992 - 2001. This study introduces a spatio-temporal autoregressive approach (STAR), which allows to investigate spatial dependencies between different properties. The data used for the study is collected from an online real estate transaction database and is obtained from an official authority (Urban Redevelopment Authority, 2013a). The study uses the logarithmic transaction price of the office unit as dependent variable. As independent variables the floor level, age, area and tenure are used. Neighbouring effects (for example view of the office or the distance to the transportation link) are included by using spatio-temporal lags. Different model specifications are considered. The two-order spatio-temporal autoregressive model (2BSTAR) performs best for the office market. The spatial autocorrelation can be reduced quite well with this model. Such STAR models are taken from studies done in the residential market and are now being applied to the office market.

The study of Chalermpong and Wattana (2009) investigates the effect of the access to rail transit on the office rent in the city of Bangkok. The paper uses spatial econometric techniques for regression, which improve the results generated with an OLS estimation. The final model integrates only five independent variables to explain the monthly rent of the office units. Two of these variables measure the accessibility, the belonging to the CBD cluster (Silom area) and the distance to the next metro station. Furthermore the building age, the number of elevators and the air conditioning system are used in the model. The air-conditioning variable defined in this study is opposite compared to other studies, which explains the different direction of the sign. The model shows that the accessibility to the rail transit system only has little impact on the office price. This can be explained through the relative new and small metro system in Bangkok. Most high-rise office buildings in Bangkok were constructed before the metro system. Otherwise, belonging to the Silom area has an high impact on the office price. This means, that a location in an office cluster can be considered as an important factor.

Kim (2007) describes in his study the properties of the office market in Seoul and the influence of transit station locations on office rents. The study investigates the office rent per leased area and the land rent per unit. To reduce spatial autocorrelation of the data, the study uses different regression models. Firstly, a basic hedonic model is regressed using OLS to investigate the differences to the spatial regression models. Further the SAR, SEM and the SAC model are estimated (for details on the different models see Section 3.2). The total floor area of the building and the number of elevators are significant. The age, heating and air conditioning system and the ratio of the contract size to the building size are not significant. The distance from the CBD and the distance to the nearest transit station are significant, but the distance to the nearest sub-centre is not. Furthermore the passenger ridership of the nearest transit station and a location quotient of financial institutions showed to be significant. The identified influence of transit stations on office rents cannot be easily transferred to suburbs. The results can only be used for city centres. The results show that the OLS estimation contains a strong spatial autocorrelation and that spatial regression models should be used for such studies.

2.1.3 Retail and shopping centre price regression

On the retail side hedonic price regressions are often only done for shopping centres or retail clusters. In shopping centres non-spatial factors are more important as for shops not located in a shopping area. For example, the prestige attraction of an anchor tenant in a shopping centre causes higher rents (Nase, 2013). For this thesis the main focus will be on both, shopping centres and stores situated on the streets. Table 4 shows a summary of retail studies done in the past with the most important attributes and their influence.

Nase (2013) investigates the relationship between retail property price and urban design quality using retail transactions of the city of Belfast. The hedonic price analysis is done by OLS regression. The dependent variable is the unit price per square meter. The significant variables are listed in Table 4. The independent variables are categorized in physical and market characteristics, spatial and time controls and quality design of the shop. The unexpected sign of the age variable can be explained by the centralised characteristic of the properties considered. Further, a shop

Table 4: List of hedonic retail rent studies with the significant variables and the estimated coefficient signs

Study	Nase 2011	Liang 2011	Hui 2007	Francois 2005	Benjamin 1990
Study area	Belfast	Shanghai	Hong Kong	Quebec	Greensboro
# obs.	1300	84	151	939	103
R^2	0.72		0.41	0.48	0.34
Age	+	−	−		
Total bld. floor size	o	o	+	−	
# of floors		+			
Vacancy rate	+		o		
Dist. to CBD	−	−			
Dist. to best mall	+				
Attractivity index	+			+	
Contract duration	+				+
Location dummy		+	o		
Type of shop/mall		+/−		o	
# of shops			+		
National chain	+			−	
International chain	+				

+/−: coefficient sign, o: not significant

being located at a corner is pricier than at other locations. The design quality of the building can be considered as a price-increasing variable.

The study of Liang and Wilhelmsson (2011) estimates a spatial hedonic retail rent model for the city of Shanghai. The dependent variable is the rent of the shops. The study investigates the influence if a shop is located in a shopping centre, in a department store or in a trade market. Further the influences of the different districts in Shanghai are tested. A problem of the study is the limited amount of observations. This leads to multicollinearity effects between the different regions.

A study which considers only the unit rent price in shopping malls is done by Hui *et al.* (2007). The regression analysis considers different malls in the city of Hong Kong, where the mall age, size and number of shops and the mall itself are considered. The dependent variable is the unit rent per square meter. The study concludes, that district centres generate the highest rent levels.

François *et al.* (2005) consider retail unit rents in shopping centres. This study includes a centre attraction index and an economic potential index. The attraction index describes the influence of the mall size and the population living around the mall, weighted by the squared distance. The economic potential index covers the income and potential spending of the customers. For a detailed description of the functions refer to François *et al.* (2005).

2.1.4 Industrial price regression

Studies that cover industrial units are quiet rare and most of them are considered for the United States. Table 5 gives an overview of different studies with the identified significant dependent variables. The work of Jackson (2002) includes a list of previous studies. Dunse and Jones (2005) also list the main previous studies done in the area of industrial hedonic pricing.

Dunse and Jones (2005) develop a hedonic industrial price regression model for the city of Glasgow. The dependent variable is the sales price of the property. The study tests physical and locational variables, including the distance to the city centre and to the next highway entrance. The study also uses locational dummies for testing of existing sub markets. The theory of the declining rental gradient from the CBD and the existence of different sub clusters is investigated in this study. This means that the rental price decreases with increasing distance from the CBD.

Jackson (2002) provides a comprehensive literature overview, with a distinction of studies that include environmental impacts or do not. In his study, the author investigates the influence of the contaminated ground in Southern California. For this independent variables, which indicate, if the land remediation has already occurred or not, are included. Further, several physical and tenant variables are employed.

The study of Dunse *et al.* (2000) for the city of Glasgow shows the impact of different spatial locations for industrial properties. The dependent variable is the asking rent of the properties. Further to the normal regression study, a model where the floor area is divided into 5 size bands and another model by use type are estimated. These two special models do not perform better and do not give very different results compared to the standard model. The main findings of the study are, that the nearness to a highway junction creates a rent premium and regional access points have a high influence on the rent.

Buttimer *et al.* (1997) describe industrial rents in Dallas/Fort Worth. The dependent variable is the real rent per square foot of industrial properties. The study includes physical characteristics,

Table 5: List of hedonic industrial price studies with the significant variables and the estimated coefficient signs

Study	Dunse 2005	Jackson 2002	Dunse 2000	Buttimer 1997	Fehribach 1993
Study area	Glasgow	S. California	Glasgow	Dallas	Dallas
# obs.	429	140	3180	848	170
R^2	0.46	0.81	0.41	0.38	0.98
Age	+/−	−	+/−	−	−
Bld. floor size	−		−	o	+
Bld. surface	+				
Land surface	+				
Ceiling height				−	+
Rail siding				o	o
Dist. to airport					−
Dist. to CBD	−				
Dist. to highway	+		−		
Location dummy	−			+/−	+
Year of sale		+/−			
Type of tenant					−

+/−: coefficient sign, o: not significant

such as age and building size, and locational characteristics. The location is only included by using dummy variables for the different sub markets. A distance variable is not included.

Fehribach *et al.* (1993) investigate the sales price of industrial sites in Dallas/Fort Worth. The model includes physical and locational variables. The building size and age, the amount of office space, the ceiling height, the distance to the airport and a locational dummy affect the price of industrial properties in Dallas.

2.2 The office and commercial market in Singapore

2.2.1 Introduction to the land use regulation

Singapore became independent in the year 1965. The land area at this time was approximately 580 km^2. Through the land reclamation program the state increased the area step by step to 704

km^2 today. As several land reclamations are going on, Singapore's area is still growing. It is assumed that in the next 20 years about 100 km^2 will be reclaimed (Wikipedia, 2013b). The national land-use and planning authority is the Urban Redevelopment Authority (URA). The two main tools at hand for the URA are the Concept Plan and the Master plan. The URA is also responsible for the selling of state land to private investors (Phang, 2000).
The Concept Plan is a long-range planning instrument. The goals are to show the main development directions for land use and the transportation system. It helps the government to plan for the next 40 to 50 years and is revised approximately every 10 years. The Concept Plan is more a large-scale planning tool, which specifies new land reclamation areas and the location of future new towns.
The second main planning tool is the Master Plan. It is the detailed land use plan, which specifies planning guidelines according to the Concept Plan for local areas. It specifies the zoning, the maximum development density and the maximum building height for the individual lots. The Master Plan shows the development over the next 10 to 15 years and is revised every five years (Urban Redevelopment Authority, 2010a). More detailed development constraints are defined in tender conditions, when the state sells land for development.

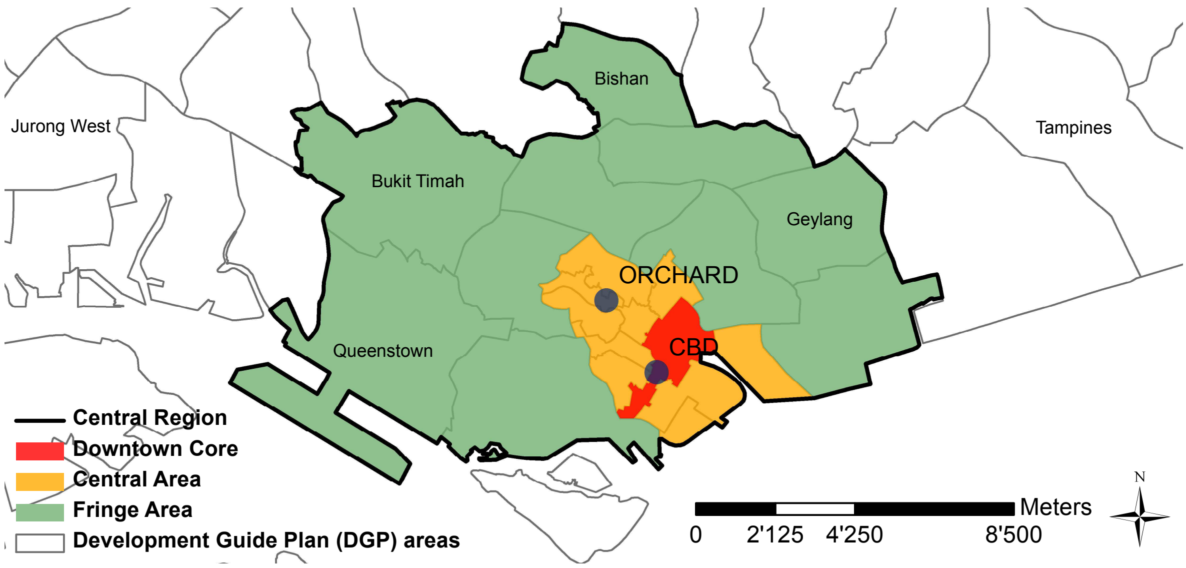
The government of Singapore was searching in the 1960s for a new concept to use land efficiently and to provide public housing to overcome the housing shortage at this time. For this purpose the "new town" planning concept was created. The housing and development board (HDB) and the URA are the main stakeholders in the development of such new towns. One of the main goals of the new towns is decentralisation, especially of housing. This means that away from the city centre new housing was built. To prevent a centralised city, each new town should be self-sufficient regarding the basic social and economic needs of the population. This means, that each new town has to provide the facilities for every day life. Through the self-sufficiency of each new town the population does not need to travel to the congested CBD for such activities. But for working a lot of people commute to the CBD, as the most office spaces are located inside the central area. Usually about 40% of the new town land is used to provide housing. About 20% is used for industry. The new towns are highly hierarchically structured. This ensures the success of the new towns concept (Eng, 1986, Shapiro *et al.*, 2008, Ho, 2007).
The HDB provides some commercial spaces, especially retail spaces in the new towns. The main goal of these commercial units is exactly to provide some everyday goods directly in the new towns (Housing Development Board, 2013). The HDB sells, as does the URA, land to private investors, but the land provided is mainly used for housing. In the housing units some kind of commercial usages can be included (Phang, 2000).

The Jurong Town Corporation (JTC) is the main industrial land developer and owner in Singapore. It provides and manages land for industrial sites and business parks. It is the biggest developer of industrial estates and other related property in Singapore. It also provides land for international firms, that want to begin a business in Singapore (Balamurugan, 2004, JTC Corporation, 2012).

2.2.2 Spatial definitions

Figure 1 shows the most important areas and locations which are used in the following sections. The central region is defined as the union of the central and fringe area. Most statistics for properties in Singapore are subdiveded into theses subgroups.

Figure 1: Location map



Source: Area definitions: Urban Redevelopment Authority (2013a)

2.2.3 Tenure conditions and land sales

The state of Singapore owns four-fifths of the land, which allows the state to regulate and control the development of the city (Phang, 2000). The Planning Act gives it control over the privately owned land. The major stakeholders in selling the state owned land are the URA and the HDB. For developing the state land a competitive bidding is used. The market price of the land is evaluated by using the one-fifth of land in the private sector. In the tender documents the allowed usage is described. This includes for example the plot ratio, building height and layout, time frame for completion and open spaces. The lease of the lot is usually limited to a time frame, normally 99 years for office and retail. For industrial usages the leasehold is normally 60 years. The land parcels which have leases, that are for sale are published on the website of the URA. This bidding framework allows the government to control exactly what is built in which location and in what time frame.

Until the late eighties the URA gave financial and tax incentives to attract more investors and international companies. Nowadays such incentives are not given any more, because the demand for land is high enough (Urban Redevelopment Authority, 1995). Urban Redevelopment Authority (1995) gives an introduction to the land sales program of the URA and contains some examples of development. According to Urban Redevelopment Authority (2010b) 35% of the total commercial space, 39% of the accommodation facility spaces and 29% of the private housing is located on lots, who are leased through the land sales program.

2.2.4 Singapore's office market

Tu *et al.* (2004) present in their hedonic study for the Singapore office market an introduction to the office market. They subdivide the office market in Singapore spatially in five different areas: Downtown core, Orchard Road, Rest of the Central Area, Fringe Area and Outside Central Region. Figure 1 shows a map with these different areas. About 90% of the available office spaces in Singapore are concentrated in the first four areas. Private citizens own about 80% of the office spaces. The contracts for commercial properties are usually only signed for 1 to 3 years. Longer contract periods are rarely signed. After this period an extension of the contract for another 1 to 3 years is generally possible, but with a change of the rent to the recent market conditions.

2.2.5 Shopping centres in Asia and Singapore's retail market

More and more retail activities take place in shopping centres or malls. Wee and Tong (2005) gives an overview of the management of shopping centres in Asia. The main success factors of a shopping centre are:

- Location: The location is the most important attribute. Included in the location is the catchment area. The ideal location is near to public transit facilities with a big potential population in the catchment area.
- Visibility: The shopping centre should be visible to the traffic and to pedestrians. This includes also the design of special building forms and façades.
- Accessibility: The accessibility to private traffic and public transportation are both important. A high amount of passing by traffic is considered as a benefit for the centre.

- Size: Increasing size increases the variety of the stores and attracts a more diverse audience.
- Other factors are space planning and design, anchor tenant, branding, management and marketing.

Ho (2007) gives a short summary of Singapore's retail market. The main shopping area is located near Orchard Road. In the 1960s the URA planned shopping centres for the first time to overcome the problem of land scarcity. As Orchard Road got congested in the 1990s the URA firstly formulated their idea of decentralisation in the Concept Plan. With the decentralisation the new regional centres should have their own suburban shopping centres. According to the development history Ho (2007) distinguishes three retail space categories:

- Primary shopping area: the central shopping area around Orchard Road. The main target group for these shopping centres are tourists, but it also serves the local population.
- Secondary shopping area: outside the primary shopping area, but still inside the central business district.
- Suburban shopping area: all other areas, like the shopping areas in the new towns. Such shopping centres can usually be found near MRT stations and housing settlements in the new towns.

Through the large number of shopping centres the competition is high. This influences the prices for the retail spaces. The competition also causes that some older shopping centres need to be refurbished to be able to compete with the new ones.

2.2.6 Land use zoning classification

The URA defines the land-use very precisely. The Master Plan distinguishes apart from residential special zoning areas for commercial developments, hotel and accommodation facilities and industrial or warehouse developments. The Urban Redevelopment Authority (2011) defines three commercial development groups for the commercial and office market. These usages can also be found on the Master Plan:

- Commercial buildings: These buildings contain only commercial usages, such as offices, shopping centres, etc. They are also allowed to contain mixed commercial usages.
- Mixed commercial and residential buildings. Buildings in this category can contain up to 40% of ground floor area for commercial usage.
- Residential buildings with commercial use on the ground floor. Usually this land use is found in the new towns, especially for providing the residents with everyday goods.

The industrial and warehouse developments are subdivided in three land usage classes. The impact of the usage on the environment is considered in the classification of this usage. The following list shows the different categories in the Master Plan (Urban Redevelopment Authority, 2011):

- Business 1 (B1): For this usage the impact on the environment should not be larger than 50 m around the site.
- Business 2 (B2): This usage has a nuisance buffer greater than 50 m. In this category special industries such as manufacturing of industrial machinery can be found.
- Business Park: This zone is only for non-pollutive industries and businesses. The zone is especially designed for high technology, research and development usages.

The third type of development is intended for hotel and other accommodation facilities. The Master Plan distinguishes only one special "hotel" zone. The other accommodation facilities can also be built in other zones, depending on the purpose of the accommodation. Student hostels for example can be built on sites zoned for educational purposes. Hotels are only allowed in the "hotel" zone, as they have a commercial purpose (Urban Redevelopment Authority, 2011).

2.2.7 Key figures

Table 6 shows the key figures of Singapore's office, retail and industrial market. The available stock is staying constant, except for private office and factory properties, which have been are increasing constantly by about 2% per year over the last years.

Table 6: Key figures of Singapore's office, retail and industrial market

	Office		Retail		Industrial	
	Public	Private	Public	Private	Public	Private
Av. stock ['000 m^2]	1'256	6'014	1'074	2'346	5'052	27'004
Share [%]	17%	83%	31%	69%	16%	84%
Vacancy rate [%]	5%	10.3%	2.6%	6.4%	3%	7.7%
	Central area	Fringe area	Central area	Fringe area	Multi-User	Warehouse
Price [\$/m^2]	11'898	8'592	15'803	12'395	6'352	9'045
Rent [\$/m^2 mo.]	100	44	96	80	23	22

Source: Urban Redevelopment Authority (2013a)

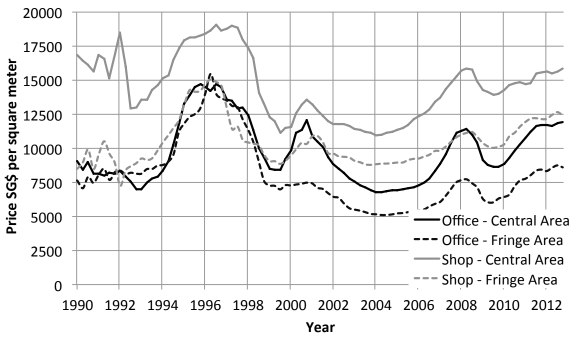
Private investors provide more than 80% of the office space. Over 90% of the available office stocks (public and private) are located inside the central region. This shows the importance of the centrality of the CBD. As the central region is quite large, it is however also possible that inside this area some office clusters exist, for example the CBD. Shop spaces are 70% privately owned and about 70% of the total shop spaces are located inside the central region. The majority of the public owned retail spaces are located outside the central region. This shows that the government tries to decentralize the traffic flows (Urban Redevelopment Authority, 2013a). For a detailed description of the spatial distribution of the data refer to Table 8.

Table 6 shows, that the vacancy rate is smaller for public owned properties. For the private office sector, the vacancy rate is in the fourth quarter of 2012 approximately 10%, it decreased in the last two years by about 5%. One reason for this could be, that public owned spaces are cheaper and so the vacancy rate is smaller, but this assumption cannot be proven. Another reason could be, that private owners and the government have different renting strategies. The public sector is not profit oriented, whereas as the private sector is. Thereby the government is renting their properties for a lower price, as they are not waiting for a potential customer, who could pay more. The private land owners are willing to wait longer for a tenant, if the revenue is higher afterwards.

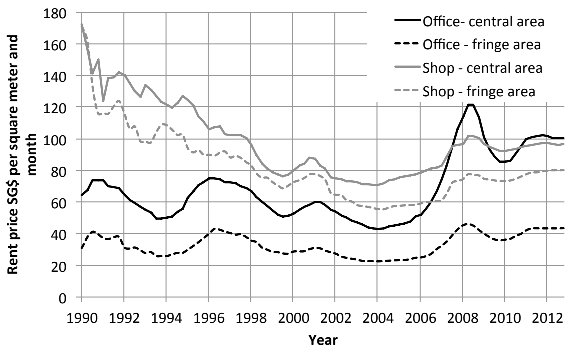
Figure 2 shows the price trend for offices and shops, as well as the vacancy rate for the different sub markets. Figure 2(a) shows the price for buying a property and Figure 2(b) shows the price for renting. The prices in the central area of the city are higher as expected. The price for buying a shop (per square meter) is in 2012 nearly 4000 $ higher than the price for a square meter of office space. In the rent prices this difference does not exist at all. The prices in the central area are nearly the same for shops and offices. Obvious is the influence of the global financial markets. The global financial crisis begun in 2008 and prices also fell in Singapore (Hui *et al.*, 2013). Other influences on the property market are according to Hui *et al.* (2013) the Asian financial crisis, which begun mid 1997, a recession in the city of Singapore in 2001 and the rise of luxury property prices, which also affected the office market. Figure 2(c) shows the vacancy rate. From the data source it is not clear if the vacancy rate is valid for the rental and the sale market or only for the sale market. It can be seen, that the vacancy rate is higher during the periods when the prices are lower (compare Figures 2(a) and 2(b)). During a low price period the vacancy rate is higher as more spaces are available and the sellers cannot charge a high price.

Figure 2: Median prices for office and shop properties

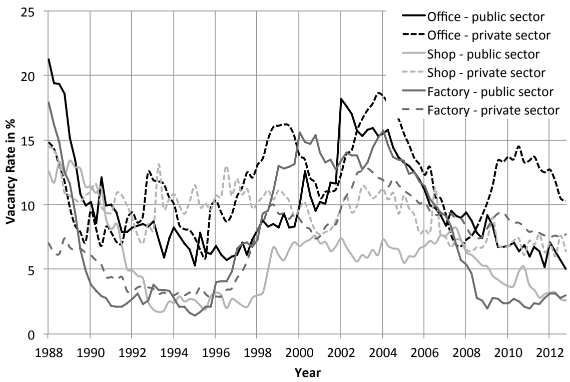
(a) Property price



(b) Rental price



(c) Vacancy rate



Source: Urban Redevelopment Authority (2013a)

2.3 Overview of the major price determinant variables

The literature review shows, that the most important variables are floor size, building age and location of the property. The location can be subdivided in different variables. For the office sector the distance to the central business district is important. The accessibility to traffic infrastructure like MRT stations, bus stations and highway junctions is also important. The availability of air-conditioning, centralised or unit-wise is an important factor especially for tropical countries. Contractual variables are not often considered, because of the lack of data. For industrial usages, the distance to the CBD is not that important. The distance to highway junctions or to the next port are more important for industrial usages.

Different special variables are mostly used for the different types of commercial usages. For office usages variables describing the neighbourhood can be considered. This can be the percentage of residential or industrial land use in the neighbourhood, as well as an affiliation to a cluster location. For retail the passenger flows are more important. The higher the passenger flows (the traffic is sometimes also considered, including the visibility and the attractiveness of the retail's property façade) the better for the retail location. This means that especially property locations at corners or in front of a shopping mall can be considered as premium location. For the industrial land use sometimes a differentiation between single user and multi user is made.

Table 7 gives a summary of the main price determinants for the commercial sector, based on the literature reviewed above.

Table 7: Major price determinant variables according to past studies for commercial properties

Variable / Characteristic	Expected sign and strength		
	Office	Retail	Industrial
Structural variables:			
Floor area	++	++	++
Building age	−−	−−	−
Building size	+		
Air-conditioning	+		
Multi/Single user factory			−
Premium location (front, corner, shopping mall)		++	
Passenger & traffic flows		++	
Contractual variables:			
Contract / Tenure duration	+	+	+
Locational variables:			
Distance to CBD	−−	−	
Distance to nearest MRT station	−	−	−
Distance to nearest bus station	−	−	−
Distance to next highway junction	+	−	−
Neighbourhood characteristics			
% of residential in the neighbourhood	−	+	
% of commercial in the neighbourhood	+	+	−
Proximity to cluster	++		

2.4 Hypotheses

According to the literature review and the presentation of the property market in Singapore, hypotheses are set up. This reveals potential effects on the properties in Singapore, as the market is quite different compared with other countries. The compact size of the state makes it also unique.

2.4.1 Structural and contract characteristics

The structural characteristics of a building are always important factors that influence the price. In a price per square-meter model the price should be lower for a bigger unit, as it can be assumed that with increasing floor area some discounts are given on the price. This effect can also be modelled with the ratio of the floor area to the plot area. It is assumed, that with increasing percentage of the sold/rented area of a property the price per square meter of the rented area sinks. The better the facilities (for example availability of air-conditioning) of the building and unit are, the higher the price for the unit should be.
The floor level is especially for the retail sector an important factor. Shops at the ground floor should be pricier than shops in higher levels. An explanation for this is, that more people will pass a shop near the entrances of the building, which are assumed to be usually at the ground floor. For office units the floor level is not that important. Only small prestige effect of the level number could be found. For industrial purposes the level should not play a big role.
In addition to the usual structural characteristics of a property also the design of the building could have an effect, especially for the retail sector. Good examples for this are the new shopping centres at Orchard Road, where a huge effort is put on the building architecture and the façade structure. This increases the recognition effect. But all such efforts do not help a lot if the visibility of the building is small. Therefore the visibility to passengers and probably traffic flows should be maximised to obtain higher property prices. Heritage buildings could have a similar effect. Such properties could generate a higher price then normal ones.

The tenure is a quite important determinant. The longer the tenure conditions are valid, the higher the price for the property should be. The highest prices should be paid for freehold properties. Property prices vary over time (see Fig. 2). This means that the contract date should have an influence on the price, depending on the actual market situation.

2.4.2 Location characteristics

Besides the structural characteristics, the location of an unit should have a big effect on the price. Accessibility of the property plays an important role for the location choice. Through a good location of the unit the firm becomes more attractive to employees. Thus the public and private transport accessibility should play a role in the price determination. The accessibility can be measured in very different ways. This study will test different location parameters and show which one can measure the differences between the locations best.
Further it is assumed, that companies cluster to profit from clustering benefits. For retail units the location plays an even bigger role then for office or for industrial properties. Through a good

location of the shop, more people can be attracted. This effect can vary for different retail types. For example a furniture outlet is usually not located at premium locations.
The distance to the airport is assumed, to not play a big role for Singapore, as it is located at one side of the city and in the surroundings no high density of commercial usage can be found. For other cities, where international firms are located next to the airport this effect might be higher.
In the literature no studies can be found analysing the effect that parking spaces have on the property prices. A reason for this could be, that usually it is not known how many parking lots are available at which place. For bigger companies the higher management decides where to create a new office. This means, that parking could have an influence on the property's price.

2.4.3 Neighbourhood characteristics

The neighbourhood of a property has an influence on the price. Firstly, the land use in the neighbourhood describes what other usages are around an unit. For example, for offices it could be the best if the commercial land use in the neighbourhood is high. This can also indicate a kind of cluster.
Similar to the average land use in the neighbourhood, the average income and the percentage of the population with a tertiary educational degree can influence the price. These two factors indicate, that high qualified labour is available just close to the property's location. Such variables are mainly interesting for office usages and not for industrial or retail. For retail usages the population density or the working people in the catchment area of the shop or restaurant could be an important factor. The higher the population or labour density, the more potential customers are inside the catchment area of a shop or restaurant.
The vacancy rate should have an effect on the price, as Figure 2(c) shows the price is correlated to the vacancy rate. It is assumed, that if the vacancy rate is small, which means that only a few properties are for sale, the prices tend to be higher.

3 Methodology

The consumer buys a good according to its different attributes. The total price of the good can be seen as the sum of the prices of the attributes. Rosen (1984) postulated the first hedonic model for durable goods, such as property.
Firstly, Section 3.1 gives an overview of the most common regression method, the linear regression. Section 3.2 shows the concept of spatial regression methods and Section 3.3 gives an overview of the different test statistics, which are available for regression models.

3.1 Linear regression

The simplest form is to assume a linear relationship between the dependent and the independent variables. An other possibility is to use the natural logarithm to transform the variables. A linear model is called log-log model, when the dependent and the independent variables are transformed into a logarithmic variable.
A linear model (without a transformation) estimates the price value for the attributes, a logarithmic model estimates their elasticities (Dunse and Jones, 1998). Elasticities show the percentage change of the independent variable, if the dependent variable changes by 1%. The simplest way to estimate hedonic property prices is to use a multiple linear regression model. The most common way to fit a linear model is to use the ordinary least squares method (Weisberg, 2005). The linear regression model takes the form

$$P = \beta X + \varepsilon \tag{1}$$

where P is the dependent variable, β is a vector of the parameters of the model (parameters for the independent variables), X is a matrix whit the observations according to the independent variables and ε is a vector with the errors.
In this study the dependent variable is the property price. The study will be done for transactional and asking prices, as they are both available for the city of Singapore. A detailed description of the used dependent variable can be found in Section 4.2.

3.1.1 Standardized coefficients

To be able to compare different parameters with different variances, standardized coefficients are calculated for the linear regression models. This coefficients are only calculated for continuous variables and are calculated by multiplying the normal regression coefficient by the standard deviation of the independent variable and dividing it by the standard deviation of the dependent variable (Stahel, 2013). The estimated parameter describes the strength of this variable and can be compared to the other ones. If the value of the independent variable is increased by one standard deviation, the explanatory variable increases (or decreases) by the standard deviation times the standardized coefficient.

3.2 Spatial hedonic regression

By using spatial data in a linear regression the results are often spatially autocorrelated. To approach these problems in the regression, spatial regression models can be used. There are different approaches to accommodate the spatial autocorrelation. Two major subgroups of spatial regression models are distinguished, global and local models. Global models do not allows the parameters to vary over space. Local models allow the estimated parameters to vary depending on their location. By using dummy variables of different locations a very basic local model can be established (Löchl, 2010). Simultaneous autoregressive models (Section 3.2.2) are global methods and geographically weighted regression (Section 3.2.3) is a local method.

3.2.1 Calculation of spatial neighbours and weights

The spatial regression methods use a neighbouring matrix (W), which describes the neighbourhood of each location. Each element w_{ij} of this matrix describes the relationship between the locations i and j. It is also possible to describe such a matrix on a non-spatial background.
The most common type for calculating the neighbours is to use a k nearest neighbours algorithm. Such an algorithm searches for a location the k nearest other locations.
A problem which happens with the data in high density environments is, that many listings are located at the same location. This means, if there are more listings at the same building as the algorithm searches for neighbours, some actual neighbours are missing in the matrix. For example, if we look at a building with 20 units and we search for the next 12 neighbours, we will miss 8 neighbours in the same location and there will be no units in the neighbours list of other locations. To prevent this problem, the neighbours matrix can be made symmetric. This means if i is a neighbour of j, j also needs to be a neighbour of i. The missing neighbours are just added to the matrix, what makes the matrix symmetric.

The second possibility to find neighbours is to define a critical distance. If other locations are located within the critical distance they are neighbours. This can be expressed as follow

$$\begin{aligned} w_{ij} &= 1 \text{ if } d_{ij} < d_{cr} \\ w_{ij} &= 0 \text{ otherwise} \end{aligned} \tag{2}$$

where w_{ij} is the weight for the observation i and neighbour j. d_{ij} is the distance between locations i and j.
After the neighbours have been found, they need to be weighted according to their supposed influence on the location. The simplest way to do this, is to give all neighbours the same weight. This means, that the distance does not plays a role. The second method is to weight the neighbours according to the distance between the original location and the neighbour. This method is also called inverse distance weighting. An example of a weighting function that can be used is shown in Figure 3.

In this study the package “spdep” of the software R (R Development Core Team, 2013) has been used for calculation of the k nearest neighbours. To calculate the nearest neighbours by distance a function had to be implemented.

3.2.2 Simultaneous autoregressive models (SAR)

The spatial simultaneous autoregressive approach assumes that each observation in the regression is dependent of its neighbours. Kissling and Carl (2008) give a good overview of the SAR models. This is done by using the spatial relationships of the observations.
Three different modelling approaches are distinguished in the simultaneous autoregressive modelling. This differentiation is based on where the spatial autocorrelation is assumed to occur.

The spatial error model (SAR_{err}) expects the spatial autocorrelation to happen in the error term. This occurs when a spatially distributed variable could not be included in the model. The model form is

$$P = \beta X + \lambda W u + \varepsilon \tag{3}$$

where P is the dependent variable, β a vector with the coefficients for the independent variables and X a matrix with the independent variables. This part is the same as for a linear regression model. The SAR_{err} model is completed with the term $\lambda W u$, where W represents the previously introduced weights, u the spatially dependent error term and λ the spatial autoregression coefficient.

The second distinguished model is the SAR lagged model (SAR_{lag}). Here it is assumed, that the autocorrelation only occurs in the dependent variable. The SAR_{lag} model form is

$$P = \rho WP + \beta X + \varepsilon \tag{4}$$

where ρ is the autoregression coefficient. As it can be seen in equation 4, the weights are applied to the dependent variable P. When it is assumed, that the spatial autocorrelation occurs at both places, in the dependent variable and the independent variables, the SAR mixed model is used (SAR_{mix}). This model is also called spatial Durbin model (SAR_{durbin}). The assumption for this model type is, that the independent variables are spatially dependent and that the dependent variable is spatially autocorrelated. The form is

$$P = \rho WP + \beta X + WX\gamma + \varepsilon \tag{5}$$

where γ is the spatial autoregressive coefficient of the independent variables. Further information on SAR models can be found in the following sources: Anselin (1988, 2002) and Fortin and Dale (2005).

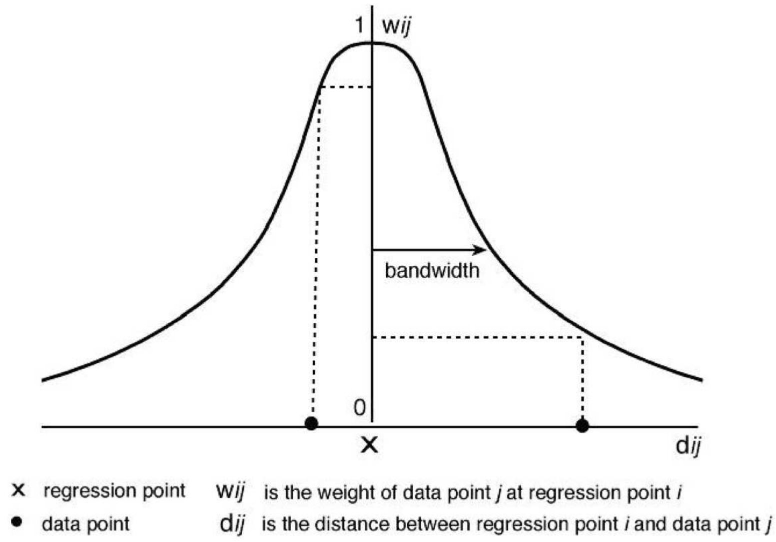
3.2.3 Geographically weighted regression

Geographically weighted regression (GWR) is another approach to handle spatial autocorrelated data. The GWR approach calculates for each location a linear regression and allows therefore the variation of the regression coefficients over space. Fotheringham *et al.* (2002) introduced such GWR models. The model form can be written as

$$P_i = \beta_0(x_i, y_i) + \sum_k \beta_k(x_i, y_i)x_{i,k} * \varepsilon_i \tag{6}$$

where P_i is the ith observation, β_0 is the intercept, $\beta_k(x_i, y_i)$ is the coefficient for the kth parameter and the ith observation at location x_i, y_i and ε_i is the error for each observation. As for each observation individual coefficients are regressed, more unknown than observed variables are present. The model assumes, that the coefficients however are not distributed randomly. Thus in the estimation neighbouring observations are used, similar as in the SAR models. The observations at each point are smoothed according to its neighbours. The standard function for smoothing is to use a Gaussian distribution. Observations that are further away than a critical distance are not included any more for the smoothing. This critical distance is also called bandwidth and is a result of an optimisation process done before the actual regression. If the bandwidth is chosen too small, the variance of the local estimators will be quite big. On the other side, with a too big bandwidth a large bias can be found in the local estimators. Figure 3 shows a possible function to use for calculating the weights.

Figure 3: Typical weighting function for calculating GWR weights



Source: Fotheringham *et al.* (2002)

For estimating the coefficients β_i the following problem has to be solved:

$$\beta_i = (X^T W_i X)^{-1}(X^T W_i P) \tag{7}$$

where W is the spatial weighting matrix and W_i represents a row of the matrix.

3.2.4 Other spatial regression methods

In addition to the presented models, other ones that are only rarely applied in current regression studies can be found. The following list gives a short overview of possible other models, which are able to address spatial autocorrelation (Fotheringham *et al.,* 2002):

- The spatial expansion method: This method allows the parameter estimates to vary locally by using a function for defining each parameter. The parameters can be expressed by a function of the location, that includes the spatial context in the model.
- Spatially adaptive filtering: Similar as the previous method, the spatially adaptive filtering allows also the coefficients to vary locally by using adaptive filters and smoothing (Gorr and Olligschlaeger, 1994).

- Multilevel modelling: This method separates the effects of property's characteristics and its locational characteristics. The locational characteristics can also be aggregated. Such a model is good if some data is available only on an aggregated level such as municipality. An application of it can be found in Haase (2011).

3.3 Regression analysis tests

This section gives a brief overview of which problems can occur with the regression data and which tests can be applied to the data to find such problems. The overview is not conclusively.

3.3.1 Heteroscedasticity

The term heteroscedasticity in statistics is used to describe a dataset, which contains a variance in the error term. This means that the errors are not normal distributed and that the error term has a different variance. In regression analysis normally the assumption of normal distributed error terms is supposed. When a dataset is heterosceadastic a part of an explanatory variable is missing in the model (Rönz and Förster, 1992).
The used statistical software (R Development Core Team, 2013) provides different tools for testing if a model has heteroscedasticity. Two methods for finding heteroscedasticity are the Breusch-Pagan test and the Goldfeld-Quandt test. For detailed explanation of the tests refer to Rönz and Förster (1992) and Breusch and Pagan (1979).

3.3.2 Spatial autocorrelation and heterogeneity

Spatial data is usually always spatially autocorrelated. Tobler (1970, 236) stated in his first law of geography: "everything is related to everything else, but near things are more related than distant things". This means, that near things tend to be spatially autocorrelated. Spatial autocorrelation and spatial dependence can be regarded as the same. The spatial dependence has to be regarded in a modelling processes (Rosen, 1984).

On the other side, spatial heterogeneity plays a big role in modelling property prices. Spatial heterogeneity means, that the space is not uniform. In every country and city regions with higher and some with lower amount of office and commercial spaces. Usually the distance to the central business district is an important factor for the attractiveness of a business district. Such effects need to be regarded in a regression, if possible (Rosen, 1984). If they are ignored the results are averaged for all areas, what means, that for example an area with cheap office prices is overestimated.

The statistical software R (R Development Core Team, 2013) provides several tests for finding spatial autocorrelation. The most commonly used one is the Moran's I test for spatial autocorrelation. It checks with a spatial weights matrix (see Section 3.2.1) if the used dataset is spatially autocorrelated. The Durbin Watson test checks if the disturbances are autocorrelated. An other possibility, which allows to do a distinction between spatial lag and error models, is to use the Lagrange multiplier test statistics (Anselin, 2003). These test statistics will be reported for the analysis of the regression models.

3.3.3 Akaike information criterion (AIC)

The Akaike information criterion (AIC) is useful when creating models. The lower the AIC is, the better the model fit. The main usage is to compare different model specifications. This can be models with different variables selection or also the regression method. The AIC cannot be used to compare different models with different data against each other. For this other tests have to be used (Burnham and Anderson, 2004). The AIC can be calculated as follows:

$$AIC = 2k - 2ln(L) \tag{8}$$

k is the number of parameters in the model and L is the Likelihood of the estimated model.

3.3.4 Pseudo R-squared

To compare the different model results of the spatial with the linear model, a pseudo R^2 is calculated. For the calculated models the McKelvey (McKelvey and Zavoina, 1975) pseudo R^2 worked best. It can be formulated as follows:

$$R^2_{McKelvey} = \frac{\widehat{Var}(\widehat{P^*})}{\widehat{Var}(\widehat{P^*}) + Var(\epsilon)} \tag{9}$$

where ϵ is the error term and $\widehat{P^*}$ the predicted price of the model. Further to the pseudo R^2, the AIC (see 3.3.3) is suitable for comparing the different models. The model and variable selection is done by comparing the different values for the AIC.

4 Data analysis and description

This chapter describes the collection and calculation of the available data. Chapter 4.1 gives a short spatial overview of the available data. Section 4.2 describes and analyses the transaction and asking prices for the units. Section 4.3 describes the available structural variables, Section 4.4 the locational variables and Section 4.5 the neighbourhood variables.

4.1 Spatial distribution of the available data

Table 8 shows prices per area in Singapore. These areas and corresponding definitions can be found in Figure 1 (page 14). For the industrial listings a differentiation between central and fringe area is not done, because the amount of the industrial listings in the central area is marginal. The biggest difference between the available stock and the listings can be found for the industrial asking transactions. The available stock data is gathered from the state's real estate information platform (Urban Redevelopment Authority, 2013a). The amount in the central region of the actual stock is about 20%, but approximately 50% of the listings are inside the central region. This means, that more properties are listed on the platform near the city centre than in the suburbs. A reason for this big difference could be, that the industrial properties in the western part of the island are only sold by the Jurong Town Corporation. If the retail and office data is compared, the biggest differences can be found between the fringe area and outside the central region. Here the same effect as for the industrial listings can be seen.

Table 8: Spatial distribution of the data compared with actual stock in Q1 2013

			Central region		Outside	Total number
			Central area	Fringe area		
Stock		Office	67.5%	24.6%	7.9%	
		Retail	38.9%	27.4%	33.7%	
		Industrial	20.4%		79.6%	
Listings	Trans.	Office	61.4%	37.8%	0.8%	4'765
		Retail	46.3%	40.3%	13.5%	9'481
		Industrial	36.2%		63.8%	18'889
	Asking	Office	68.7%	26.1%	5.2%	7'299
		Retail	40.5%	37.3%	22.3%	5'924
		Industrial	48.6%		51.4%	5'507

Source: Urban Redevelopment Authority (2013a), Allproperty Media Pte Ltd (2013)

4.2 Transaction and asking prices

For the regression analysis asking and transaction prices are available. Transaction prices describe real prices for units. Asking prices are usually higher than the transaction ones. The asking prices are the prices offered by the property owners.
The transaction prices are gathered from the REALIS (real estate information system) of Singapore's government (Urban Redevelopment Authority, 2013a). The data can be downloaded from this information system. The dataset contains all the transactions for industrial, retail and office property sales for the period between 1995 and 2013. Beyond the transaction price, the dataset contains the unit size, the contract date, the property type, the tenure conditions and the postal address with postal code and zone. Through the postal address and the postal code the location of the building can be found. The transactions in REALIS are unit fine, what means that the exact unit is known including the floor level and the unit on this floor. The information has to be extracted from the address, where the floor level is usually indicated. The transaction data is subdivided by the URA into sales of land and sales of strata properties. The URA gives no detailed description what the difference is between these two categories. According to Wikipedia (2013c) a strata title is the ownership of an apartment or of a property part. There are always some shared spaces in the building, like staircases, which are not owned by a individual person.

Table 9: All available listings of CommercialGuru and REALIS (since 2011)

		Strata		Land		Total	
REALIS		#	%	#	%	#	%
	Office	1,733	15.2%	10	0.9%	1,743	13.9%
	Retail	2,169	19.0%	586	50.9%	2,755	21.9%
	Industrial	7,528	65.9%	555	48.2%	8,083	64.2%
	Total	11,430		1,151		12,581	
		Sale		Rent		Total	
Commercial Guru		#	%	#	%	#	%
	Office	1,216	22.3%	4,613	43.7%	5,829	36.4%
	Retail	1,915	35.1%	3,348	31.7%	5,263	32.9%
	Industrial	2,327	42.6%	2,590	24.5%	4,917	30.7%
	Total	5,458		10,551		16,009	

The asking prices are gathered from the online platform of the Allproperty Media Pte Ltd (2013) (www.commercialguru.com.sg), which is one of the main platforms for selling and renting commercial properties. The data are collected by using a robot, which captures all data available on the platform. This data is collected by Van Eggermond (2013). The gathered attributes depend on what is listed online. The agents decide which attributes of a property to mention. They do not select from a pre-defined set of possibilities. This leads to very different coverage of the attributes.

Table 9 shows the number of the units that are available. The numbers shown are exclusive of wrong and duplicate entries (see Section 4.2.5). For the regression analysis not the whole dataset is used. This is done to prevent that past influences and events on the property and financial market are distorting the results and because built environment can change over time. As described in Section 2.2.7 and shown by Figure 2 the market has been stabilizing since 2011. These effects can also be seen in the available data if the average price over the years is plotted. For this reason, the analysis will only contain the transactions since January 2011.
The observed share of different markets (office, retail and industrial) differs in the asking and transaction data. The biggest difference can be found for the industrial properties. The transaction data contains up to 65% of industrial property sales. The asking data just contains approximately 43%, when only considering asking (sales) prices. One reason for this could be that the revenue of office and retail properties are higher and therefore more interesting for agents to sell.
It can also be seen, that the land sales (in the transaction data) are actually only for retail and industrial properties. In the office category only a few are sold over the observation period. The reason for this is not understood.

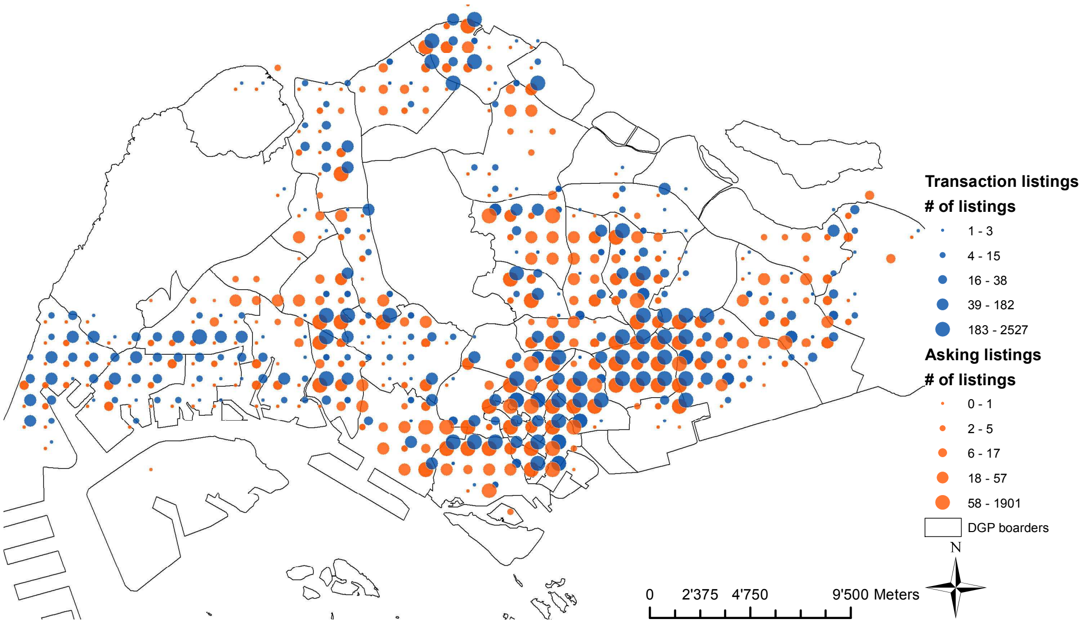
Figure 4 shows the distribution of the transaction and asking data across the island. The absolute number of transactions cannot be compared (see Section 4.2.2). Most properties are sold or offered around the city centre. This is similar for the asking and transaction data. It can be seen, that on the online platform of "Commercial Guru" much fewer listings are published for the industrial estates in the west of the island. This could be a consequence of the smaller amount of asking transactions for industrial properties. More "Commercial Guru" properties are published for the region around "Ang Mo Kio" and "Changi".

4.2.1 Price determinants

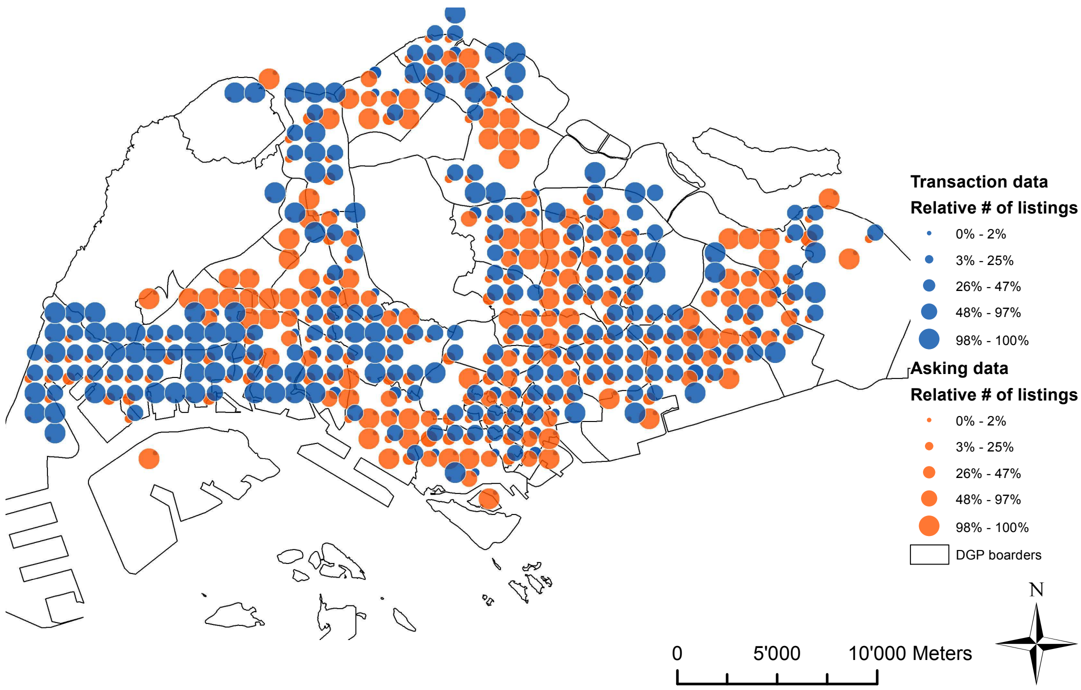
As the gathered data shows, the property price variations are high. The highest price listed on the REALIS platform is 700 mil. S$ for an industrial estate signed in 2012. The highest price per square meter is paid for an office land property next to the Raffles Place MRT station. The price is nearly 230 Mio. S$ and the price per square meter nearly 300'000 S$. The highest price for a retail location is paid at the Lucky Plaza in Orchard Road: 150'000 S$ per square meter.

Figure 4: Distribution of transaction and asking data

(a) Absolute distribution of the listings



(b) Relative distribution of the listings



The highest prices in the asking data are similar to the transaction ones. Although there is a tendency that retail locations generate higher prices than the other categories.

Table 10 shows the statistics for the price per square meter. It can be seen, that the asking sale prices vary more. The minimum and the maximum of the asking prices per square meter are lower/higher than the transaction prices. The asking prices also tend to be higher in general than the transaction ones. This fact corresponds to the expectations, that asking prices are higher and can be seen by comparing the mean or the median of the prices together. In Table 10 the differences between the asking and transaction are also listed. The retail and office price of the transaction data is about 15 % lower than the asking prices. The difference in the industrial category is much higher than for the other ones: about 40%. The industrial property prices listed on the online platform are too high.

Table 10: Main statistics of asking and transaction prices in 1000 S$ per m^2

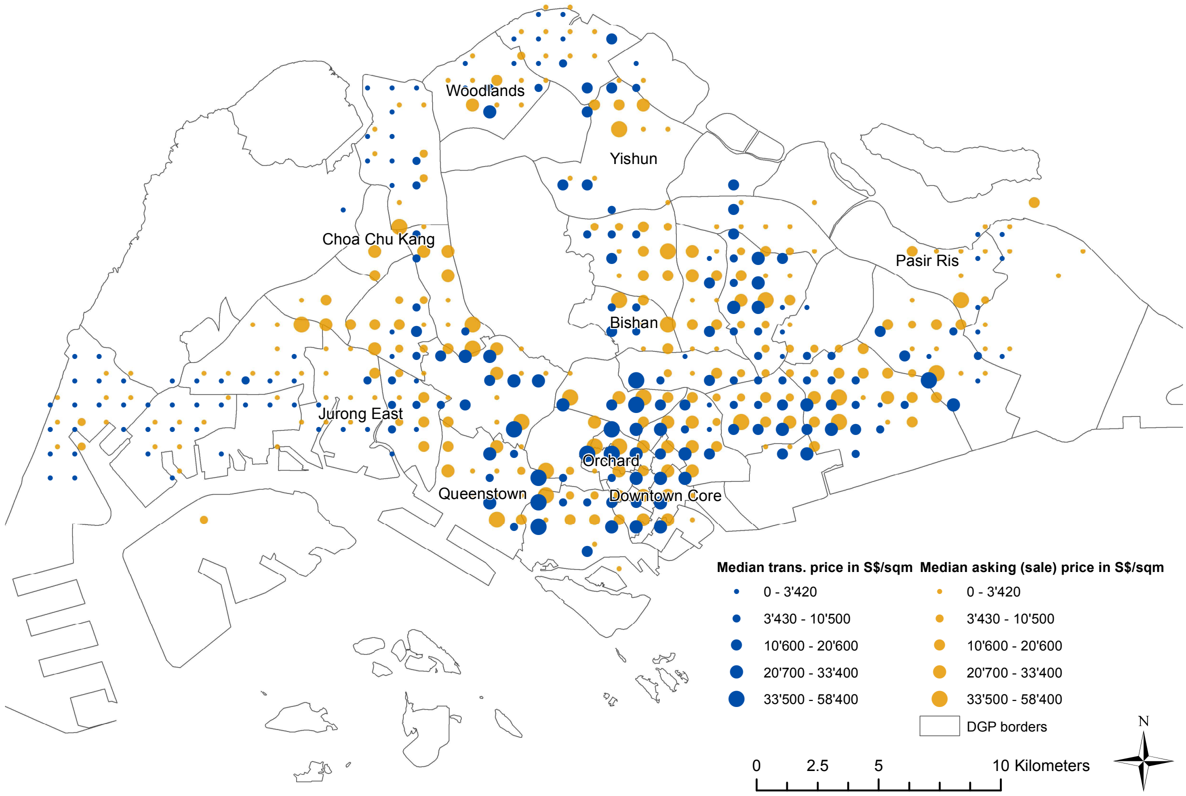
	Asking - Sale			Transaction - Strata			Difference		
	Retail	Office	Ind.	Retail	Office	Ind.	Retail	Office	Ind.
Mean	39.7	22.9	6.8	34.6	19.8	4.8	14.7%	15.3%	40.4%
Median	35.1	21.3	6.4	30.5	18.8	4.4	15.0%	13.5%	44.1%
Min	0.2	0.4	0.8	5.0	4.0	0.7	-4.8	-3.6	0.1
Max	217.4	80.7	26.9	150.5	57.9	22.6	66.9	22.8	4.3
Std.	26.9	9.6	3.0	18.4	6.3	2.1			

Figure 5 shows the spatial distribution of the median price per square meter for the transaction and asking data. The data are aggregated over a square kilometre grid. Noticeable is the difference in the outer regions. The transaction prices seems to decrease with increasing distance from the city centre, excluding some sub-centres, where the prices are again higher. This decrease is less strong for the asking data. The prices in the outer regions are also quite high, compared to the city centre and the overall price is higher for asking listings, as seen before. A reason for this could be that the agents of the asking data try to get similar prices for the properties as in the city centre, but the actual transaction price will be much lower. There are some regions, where only asking listings are available, but no transaction data. This can already be seen in Figure 4. As all sales are listed in the transaction data, it is not known why there are regions with only asking listings. A reason could be that in such places new buildings will be constructed and the units are already on the market. Figure 5 also shows that the industrial estates, especially in the north and west of the island, are relatively lower priced compared to the other categories.

Figures 6 to 8 show the distribution of the square-meter price for each category (office, retail and industrial). The office units (Figure 6) are highly concentrated in the city centre. There are only a few listings at the suburbs. The price distribution is similar between the asking and transaction data.

Figure 7 shows the distribution of the retail prices. In the city centre no difference is observed between transaction and asking prices. In the outer districts more transaction retail units are

Figure 5: Spatial distribution of transaction and asking price per square meter (all categories)



sold and the prices are still quite high compared to the city centre. A clear price decrease with increasing distance from the city centre cannot be seen. This fact shows, that the retail spaces in the new towns are mainly not sold over a commercial platform and that the prices are similar to those in the city centre.

A similar figure is shown in Figure 8 for the industrial properties. The closer the properties are to the city centre, the higher the prices, but some sub centres with increased prices can also be found. For example, such a sub-centre can be seen at Jurong East or at the most western part of the island at Tuas. A lot of the industrial properties are also not listed on a commercial platform. It is assumed that these units are sold over a separate platform, for example directly over the biggest provider for industrial land, the Jurong Town Corporation.

Figure 9 shows the distribution of the asking price per square meter of the asking data. The price for the different usage categories differs greatly while the prices for retail properties are much higher than for office or for industrial. This difference can be seen in the rental and sales prices. The minimum price per square meter is about 220 S$ for the sale and about 10 S$ for the rental observations. The retail category also has a lot of outliers, compared to the other two categories. These high prices are mainly observed in the shopping malls around on Orchard Road and the

Figure 6: Spatial distribution of transaction and asking price per square meter (office)

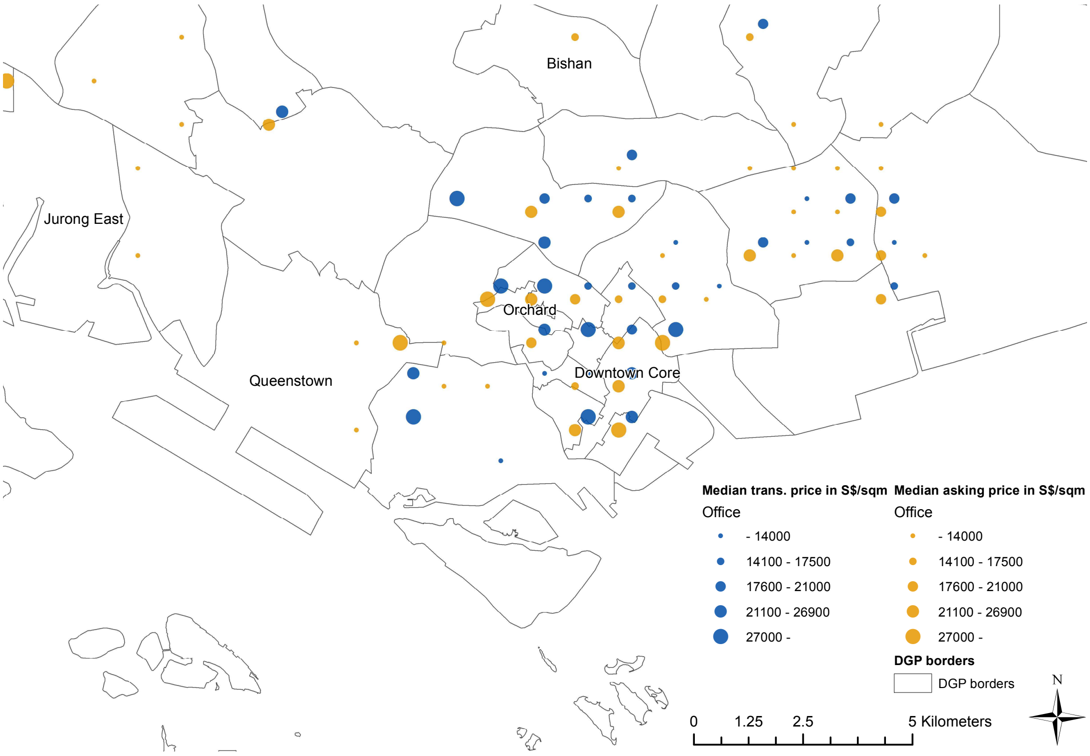


Figure 7: Spatial distribution of transaction and asking price per square meter (retail)

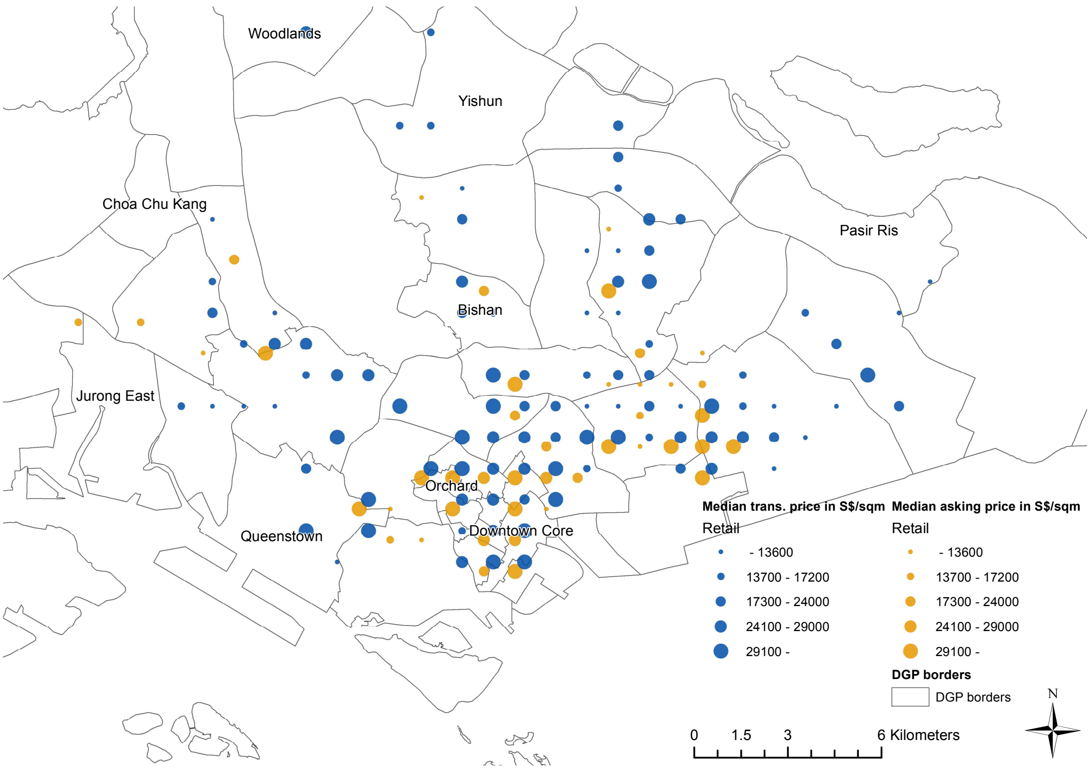


Figure 8: Spatial distribution of transaction and asking price per square meter (industrial)

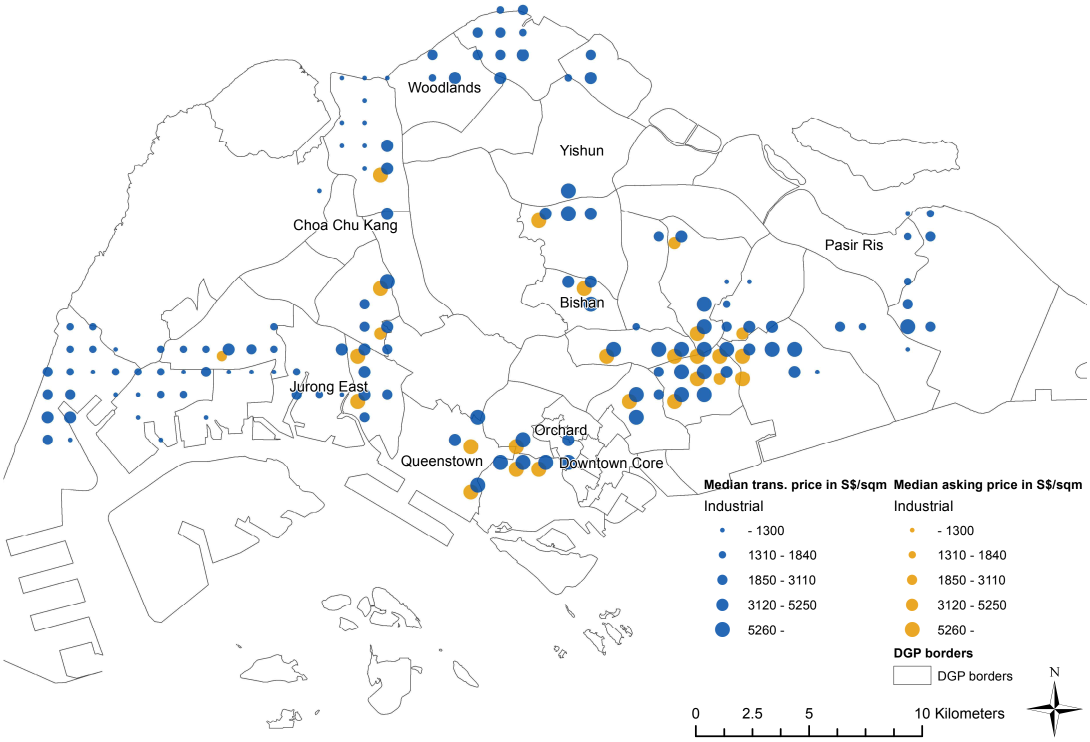


Figure 9: Boxplots of main property types for the asking data

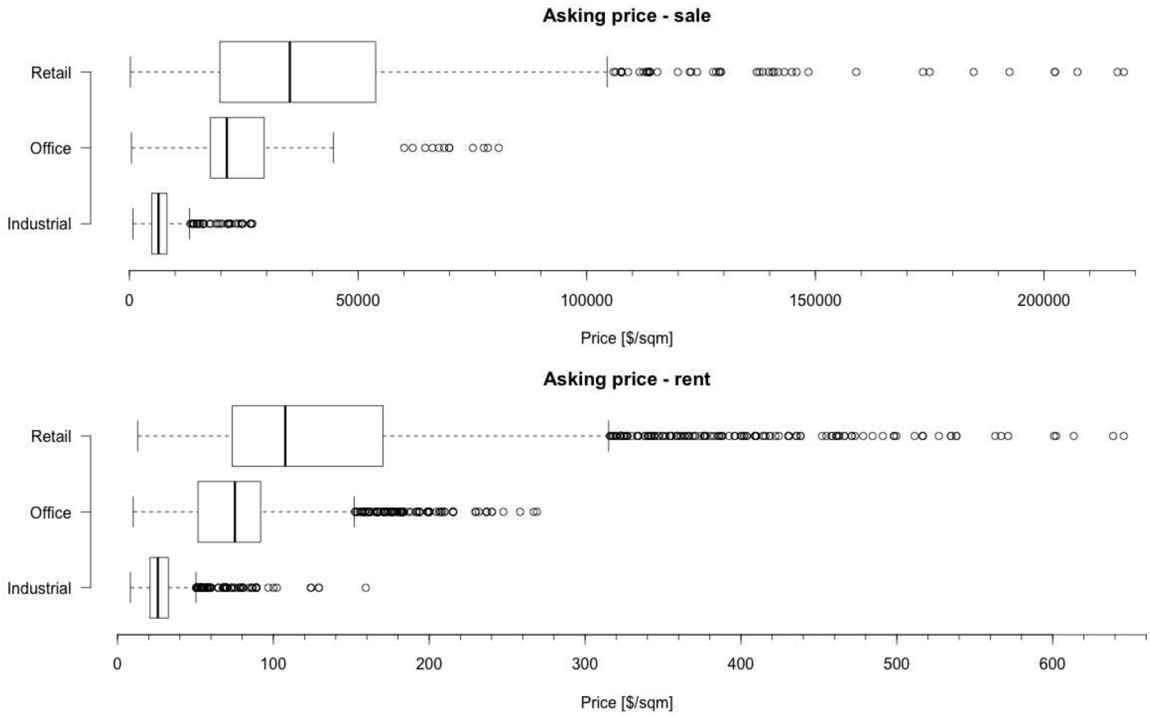
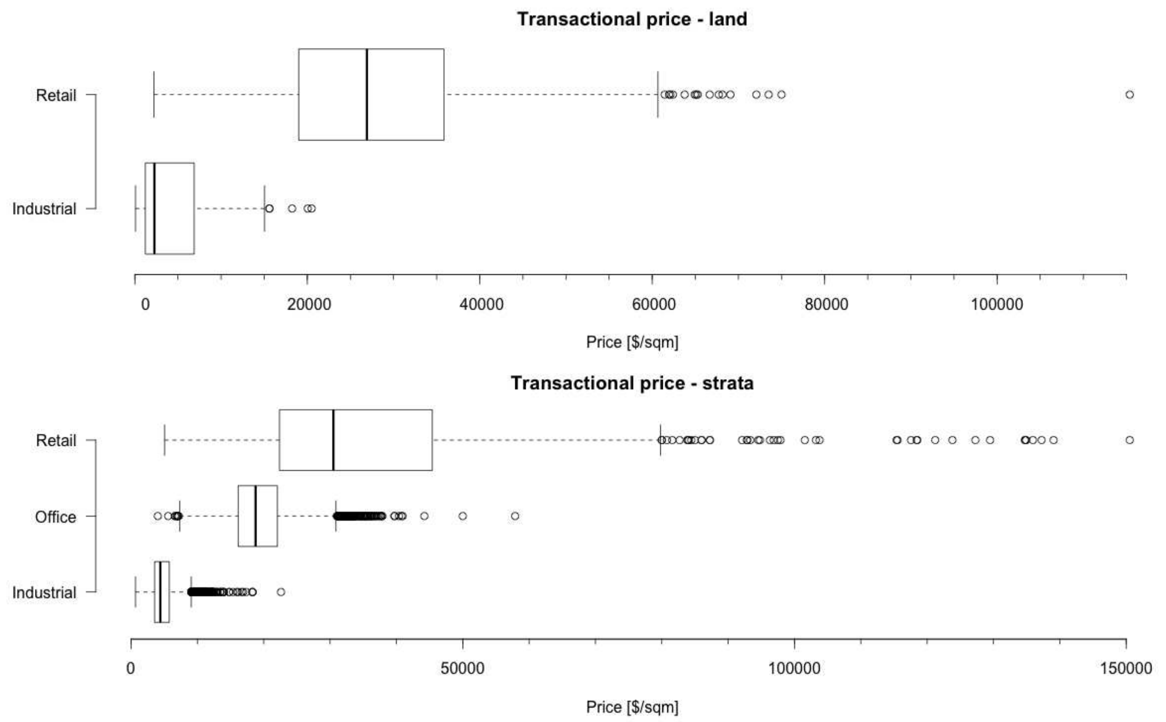


Figure 10: Box plots of main property types for the transaction data



city centre. The category of retail also contains all transactions of food stalls and restaurants. The price for a food stall in a hawker centre is assumed to be low compared to prices for shops. This leads to a very high spread of the data.

Figure 10 shows the box plots for the transaction data. The office category of the land data is excluded from the graphics, because of the small number of observations. It can be seen, that the prices of industrial properties are not widely spread.

4.2.2 Market segments

The analysis of the available data shows, that different market segments exist. The asking data is subdivided into rental and sale. These two segments are further subdivided by the type of usage: office, retail and industrial. The transaction data is subdivided into strata and land transactions. Same as for the asking data, the transaction data is further subdivided by the usage type. For the transaction data the regression analysis has to show whether separate models are necessary. If not this data can be regressed within one model.

In Singapore land is owned either by the public or by a private entities. The owner of the property is only available for a few asking data listings. The transaction data does not contain information about if the government or a private person is owning the land. For the asking dataset, the information is only available for about 40 % of the entries. Out of these, just 3 % of the properties are owned by the government. Due to this circumstance of data availability no differentiation is done by owner type in the regression analysis.

4.2.3 Price-rent ratio

The price rent ratio shows the price difference for the rent and sales data. The ratio can be compared across different countries and cities. The price-rent ratio is defined as follows (Wikipedia, 2013a):

$$\text{Price-rent ratio} = \frac{\text{property price}}{\text{property rent} \cdot 12} \tag{10}$$

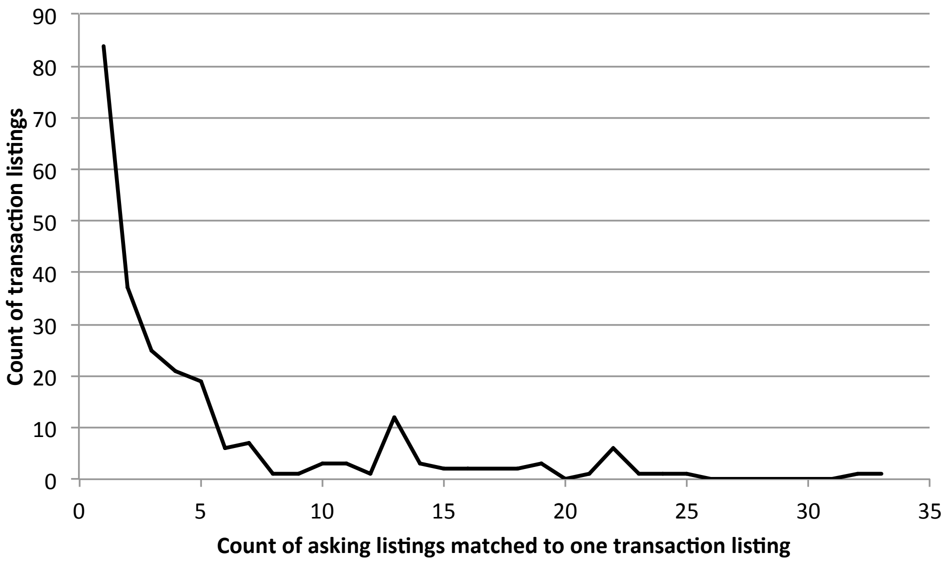
The ratio is usually used to indicate the onset of a property bubble. A property bubble is likely to occur, when the price-rent ratio increases rapidly. This index has to be calculated over a longer period of time. For Singapore some values for housing could be found in the Global Property Guide (2012). A value for the commercial sector could not be found. The price-rent ratio in 2012 for housing in Singapore is about 34 according to (Global Property Guide, 2012). For the available data an average ratio is calculated. The ratio is about 20 for the asking data over all categories together. Office and retail properties have a similar ratio of 22, industrial properties have a ratio of 17.

4.2.4 Transaction and asking data matching

As transaction and asking data are both available for Singapore, it would be interesting to match the data to each other. This would allow to show the difference in prices between the offered price listed on the online platform and the real price.
There are several problems with data matching. First the contract date (real contract signing date) and the listed date on the commercial online platform cannot be compared. The contract is signed later than the property is listed online. A second problem is, that the asking listings are not unit fine. But as the majority of the buildings in Singapore are skyscrapers, the information about the unit is quite important for matching the data. For one location multiple listings can be listed, what makes it very difficult to match.

The results of the data matching trial are briefly shown here. The asking data is only available for the period between 13-01-2013 and 11-02-2013. This means that the transaction listings need to be newer than 13-01-2013. Approximately 600 contracts are signed after this date. The matching is done by using the location and size of the properties. The biggest problem in the

Figure 11: Comparison of the amount of asking transactions that could be matched to one transaction unit



matching is, that the listings cannot be joined precisely. A transaction listing matches to multiple asking listings. As it is assumed that multiple agents publish a property multiple times a distinct matching is not possible. It could also be possible that in a building multiple identical units are listed in the asking data. Due to this fact it is not possible to make a statement about the matching quality. Further, one asking listing can also match to multiple transaction listings. This leads to the same problem as described before.

For about 250 of the transaction listings at least one asking listing could be assigned. Figure 11 shows how many asking listings could be matched to one transaction listing. For example, 33 (most right of the figure) asking listings could only be matched to one transaction listing, 84 (most left of the figure) transaction listings have exactly one asking listing that could be assigned (one to one connection).

The results of the data matching are not satisfactory, because of the problems described before. To use the matched data for a regression analysis, further effort should be undertaken to improve the matching quality.

4.2.5 Data issues

The transaction data is considered to be correct, despite the fact that there are some entries with a very high transaction price. This issue can be reduced by dividing the dataset into land and strata sales.

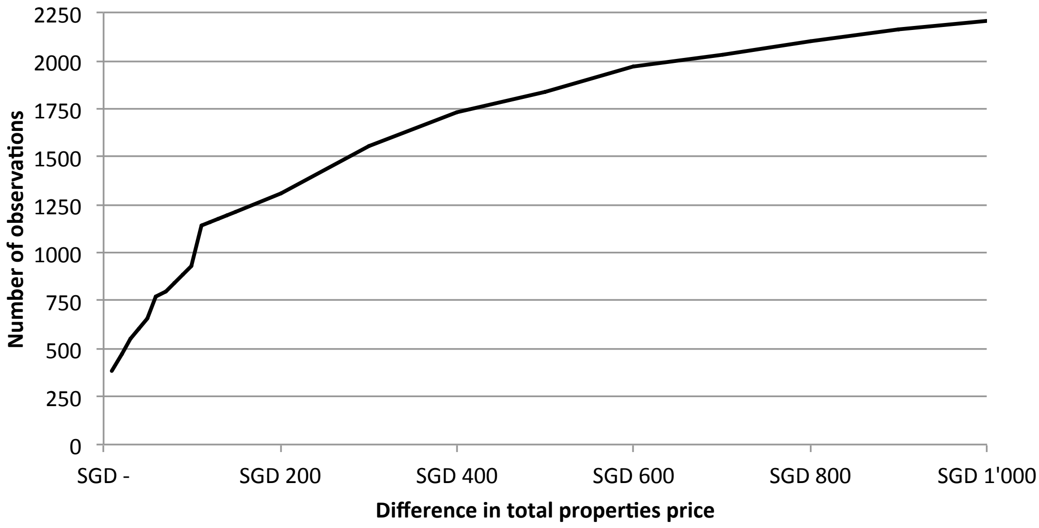
The asking price data are more prone to errors as the entries on the website are not monitored. Approximately 31'000 asking entries are available for this study. The first issue is that some listings are published more than once. A possible explanation for this is, that multiple agents put the same property online, which occurs in multiple entries. Some of the duplicates can be considered as real ones. These entries have exactly the same attributes and are excluded from the dataset. A little more than 1700 entries can be considered real duplicates. These entries are excluded from the dataset.
Furthermore a lot of entries have certain attributes which are the same or similar. If the address, the price and the floor area are identical, another 2000 duplicates can be identified. In a last step for identifying duplicates, listings with similar prices are considered. For this the difference in the price is calculated. For example some entries can be found, where the price differs only by a few dollars, and the rest of the attributes listed are the same. Figure 12 shows the number of duplicates with the according maximum difference in price. For these entries it is not known if there are really multiple similar properties for sale or rent, or if different agents put them online with varying attributes. Because of this reason these listings are not excluded from the dataset. Van Eggermond (2013) describes some problems of the geocoding of the entries. Because of missing properties addresses or postal codes, a geolocation cannot be found for all listings. For the regression analysis only the geocoded entries are used.

4.3 Structure and contract variables

4.3.1 Floor area

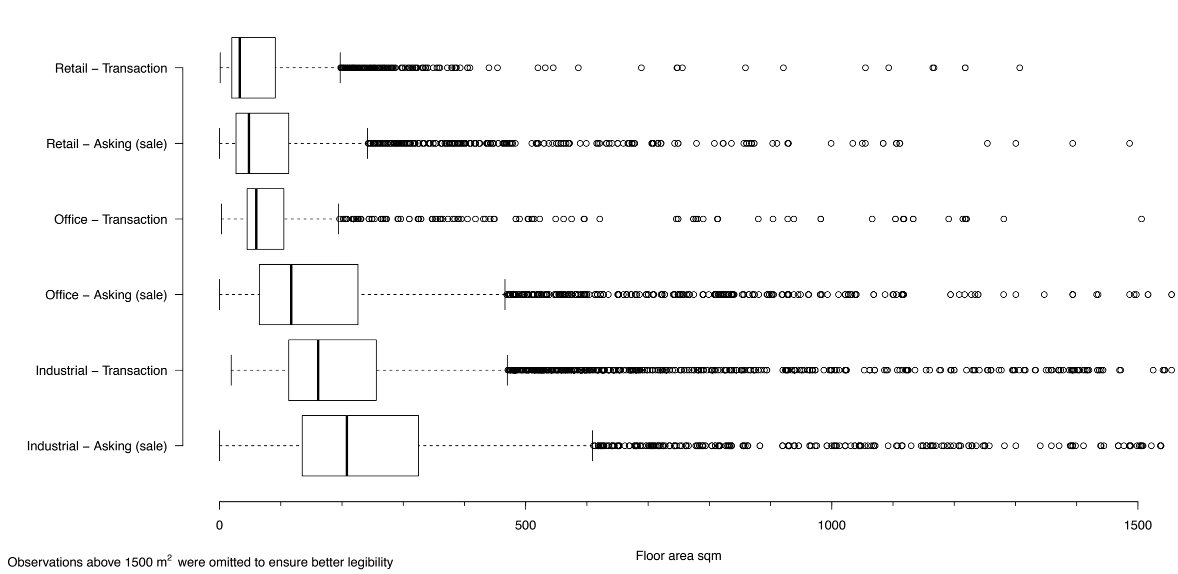
The floor area of the properties varies across the different sources and usage types. An industrial usage generally needs more space than an office or retail usage. Figure 13 shows the box plots for the floor area of the sales data. The retail's floor area range differs the least. A reason could be that a standard restaurant or little shop needs usually approximately the same space. This reduces the variance in the floor are for the retail usage. Office and industrial usages use a wider range in sizes, dependent on the actual main activity and size of the company. Clearly it is the industrial usage that needs most land or floor area for producing goods. This fact can be seen well in Figure 13.

Figure 12: Number of duplicates by varying price range



The mean value of the floor space is not similar when comparing transaction and asking data. The interquartile range of the transaction data is smaller than the one of the asking data. The biggest difference between the transaction and asking data is found for the office usages. The size of the transaction offices are similar to the retail units. The size of the office asking data is more similar to the industrial usages. A reason for this could not be found.

Figure 13: Box plot of the floor area for the different data sources and main usage type



4.3.2 Property features

Different structural features of the properties are available. For the transaction data only the information about the tenure is available (see Section 4.3.4).
For the asking dataset a lot more structural information is available. The agents usually list what facilities are available for a building. These are extracted into the database, where they can be used to create variables. The following variables are prepared to be included in the regression analysis: availability of air conditioning, central air conditioning, cargo lifts, heavy current connection and parking, extra high ceilings, high floor load capacity, high lift capacity (for cargo) and the number of passenger lifts. Table 11 shows which variables are available for how many listings. The listed features show a special characteristic of the buildings and are used in the regression analysis as dummy variables, only showing if an unit has this facility or not. The feature "high floor load capacity" in Table 11 shows for which listings the floor is suitable to store heavy items on it. Most of the listed facilities are only used for industrial facilities, as they do not make much sense for a retail or office unit. A full overview of the available features and facilities in the data can be found in Van Eggermond (2013).

Table 11: Available facility information for the asking data

Feature	Available	Not available	% available
Air conditioning	2,775	13,234	21.0%
Central air conditioning	328	15,681	2.1%
Cargo lift	399	15,610	2.6%
Extra ceiling height	1,349	14,660	9.2%
Heavy current connection	908	15,101	6.0%
High floor load capacity	1,036	14,973	6.9%
Cargo lift capacity	200	15,809	1.3%
Parking availability	395	15,614	2.5%
Number of passenger lifts	462	15,547	3.0%

4.3.3 Lot area and building footprint

Further to the variables described, the lot area and the footprint of the building are available. Some previous studies used the ratio of the rented or sold floor area to the whole building floor area as an independent variable. The total floor area of the building is not available for Singapore. But the ratio between the floor area and the footprint, as well as the ratio to the lot area can be calculated. It has to be shown if theses variables can explain some of the variance of the property prices.

4.3.4 Tenure conditions

The transaction data contains good information about the tenure. About 25% of the transaction properties are sold with freehold. The other 75% are based on leasehold. Freehold properties are privately owned. Leasehold is owned by the government and a leasehold is given for a certain amount of years to use the land according to the tender conditions. More information about the duration and the leasehold starting date are provided for the leasehold properties. This means, that for the transaction data all information about the tenure conditions are available. The median of the leasehold duration is 60 years.
The asking data does not contain such reliable information. About 30% of the rent and sale data are based on leasehold and 30% of the sale data and 10% of the rent data are freehold. For the rent data, it is not clear if the indicated freehold is based on the building and if freehold is relevant for the rental market. One fourth of the asking data contains information on the leasehold duration. The median is 99 years. To improve the data quality and quantity of the tenure conditions for the asking sales data, the available information about the properties in the transaction data is transferred to the asking listings. If a building is sold in the transaction data as freehold, the same should hold for the asking data.

4.4 Locational variables

To describe the unit's location different approaches are used. The simplest one is to calculate the straight line distance to important points. Different distances are included to describe the accessibility by public and private transport. Although Euclidean distances do not describe the distance that has to be travelled from a person or a car to reach a location very well. The real walking distance to a MRT station can be much longer than the straight line distance, as often obstacles like overhead bridges have to be overcome. A better approach is to use a time based distance, which includes the actual walking or travel time to a location (see Section 4.4.2).

As seen in the literature overview, the distance to the central business district is an important determinant, especially for office usage. The median distance for the transaction office data is 1.9 km, the one for the asking data 1.4 km. The value for the overall dataset is approximately 6.5 km. Only the asking rent listings differs, with only 2.9 km. One reason for the lower value in the rent category could be, that most listings are for office properties and these are likely to

be closer to the CBD. The higher distance for all listings shows, that the office usage is highly clustered in the area around the CBD.
For the retail properties the distance to the main shopping area “Orchard Road” could be a important variable. The median distance for the transaction retail data is 3.2 km, 3.6 km for the asking sale data and 3.0 km for the asking rent data.
To calculate the straight-line distance to the CBD the reference point is taken at the centroid of the CBD area. The same is done for the reference point at Orchard Road. The area definitions are taken from the Development Guide Plan (Urban Redevelopment Authority, 2013b).

4.4.1 Public & private transport access

The access to public transport is measured in the analysis by using the distances to the nearest MRT and bus stations. The distances are always calculated as straight-line distances, because a footpath network is not yet available for the city of Singapore. This approach is also just a simple representation of the accessibility. For a better implementation a real accessibility index is used and compared (see Section 4.4.2).
The access to the bus stops is measured in different ways. First just the simple straight-line distance is calculated. This measure does not include any kind of attractivity. Thus the number of bus lines which are serving a bus stop are counted. This is done for the five nearest stations, but a bus line is just counted once. This means that the bus line is unique in the attractivity index. Further the number of bus stops and lines are counted in a perimeter of 500 m around the unit. The different variables which describes the accessibility of the bus stops are tested against each other in the regression analysis to find the one which works the best.
For the transaction data the median distance to a MRT station is about 750 m, the distance to the nearest bus stop is approximately 120 m. For asking data the values are 100 m for the bus stops. The distance to the nearest MRT station differs between the rental and sales listings. For the rental listings the median distance is 430 m and for the sales listings 640 m. This difference in the distance is probably due to rental objects being located in the city centre and less in the outer areas.

The private transport accessibility is measured by using the straight-line distance to the nearest highway entrance, as well as the distance to the nearest car park. For the industrial land use the distance to the nearest truck parking is calculated. Further the amount of parking in the neighbourhood of 500 m is calculated. The parking amount is taken from a dataset of the Urban Redevelopment Authority (2013b). This dataset does not contain parking lots in private buildings. That means, that the exact amount of parking cannot be reproduced by this variable. It just measures differences in the parking availability between the different regions. The regression analysis will show, if this simplified parking availability index has a significant influence on the price of the properties.

4.4.2 Accessibility

Apart from the mentioned methods an accessibility index is used to describe the units' location. The accessibility index is only calculated for the public transport. The index for the private transport is not yet available. The accessibility describes for each postal code (building) in the state of Singapore the potential opportunities that can be reached. The accessibility index is calculated for working and residential opportunities. These indexes are calculated by Michael van Eggermond from the Future Cities Lab (FCL) and can be formualted as follows (Ortúzar and Willumsen, 2011):

$$P_i = \sum_{j=1}^{n} O_j exp(-\lambda t_{ij}) \tag{11}$$

where P_i is the potential number of activities accessible to building i, O_j the opportunities in building j, λ the distance decay factor and t_{ij} the travel time from i to j. The travel times between the buildings are calculated using the agent based transport simulation MATSim. For each building the five nearest public transport stops are selected in a radius of maximum 500 m. From these five stops the travel times to the destination are calculated and the fastest route is chosen. The accessibility calculation does not include yet the access time to the transit stop. For this purpose a pedestrian network is generated and later used for calculating detailed access times.

4.4.3 Poly- vs. mono-centric city analysis

Figure 16 shows different cuts through the data for analysing price differences at different locations. Sometimes the assumption of a mono-centric city is made according to economic theory. For example the variable which measures the distance from the city centre is a typical mono-centric city structure indicator. As this theory does not hold for a lot of cities, this section shows the price distribution over different "cuts" through the city. The different cuts which are analysed can be seen in Figure 14. According to these lines the listings which are closer than 1 km to the line are analysed and plotted at the distance that they are away from the reference point.

Figure 15 shows the cut though the city centre. The reference point for this line is taken near the Harborfront MRT station. This graphic is especially interesting for the office sector, as

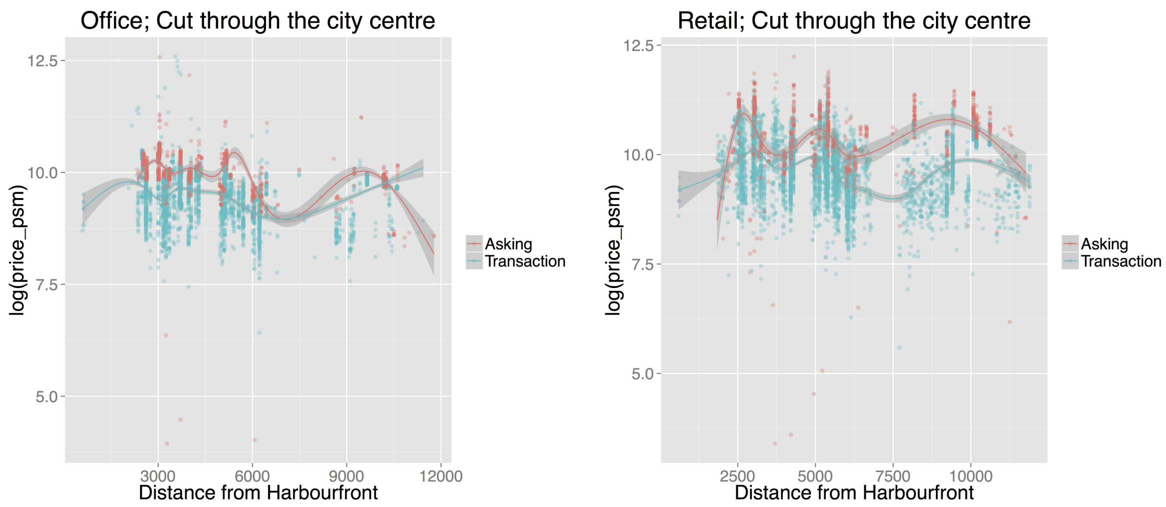
Figure 14: Overview of analysed lines

Source: Background: Esri Inc. (2013)

the majority of the office properties are located within this area. The figure shows, that some fluctuations in the price per square meter can be found. It can also be seen, that the asking prices are throughout higher than the transaction prices. This fact has already been shown before.

Figure 16 shows the results for the different cuts. For these lines the industrial properties are also analysed. The reference point for these lines is at the city centre near the Marina Bay MRT station. Obvious is that the price for the asking units is nearly always higher than the transaction price. The variance of the asking price is also higher. In the transaction price some variations can be found, but the overall price is quite similar over the analysed distances. This means, that the assumption, that the price will decrease with increasing distance form the city centre is probably not valid. The figure clearly shows, that there are some areas with very few listings and some with a lot of listings. This reflects in some kind the land use regulations of the government, as it is not allowed at every part to build commercial properties. This fact can be seen quite good for the retail listings in direction east. It shows another cluster at approximately 5 km from the city centre, after the Marina Bay, but the prices per square meter at this location (around Katong) are only slightly smaller than in the CBD.

Figure 15: Price per square meter analysis for the city centre

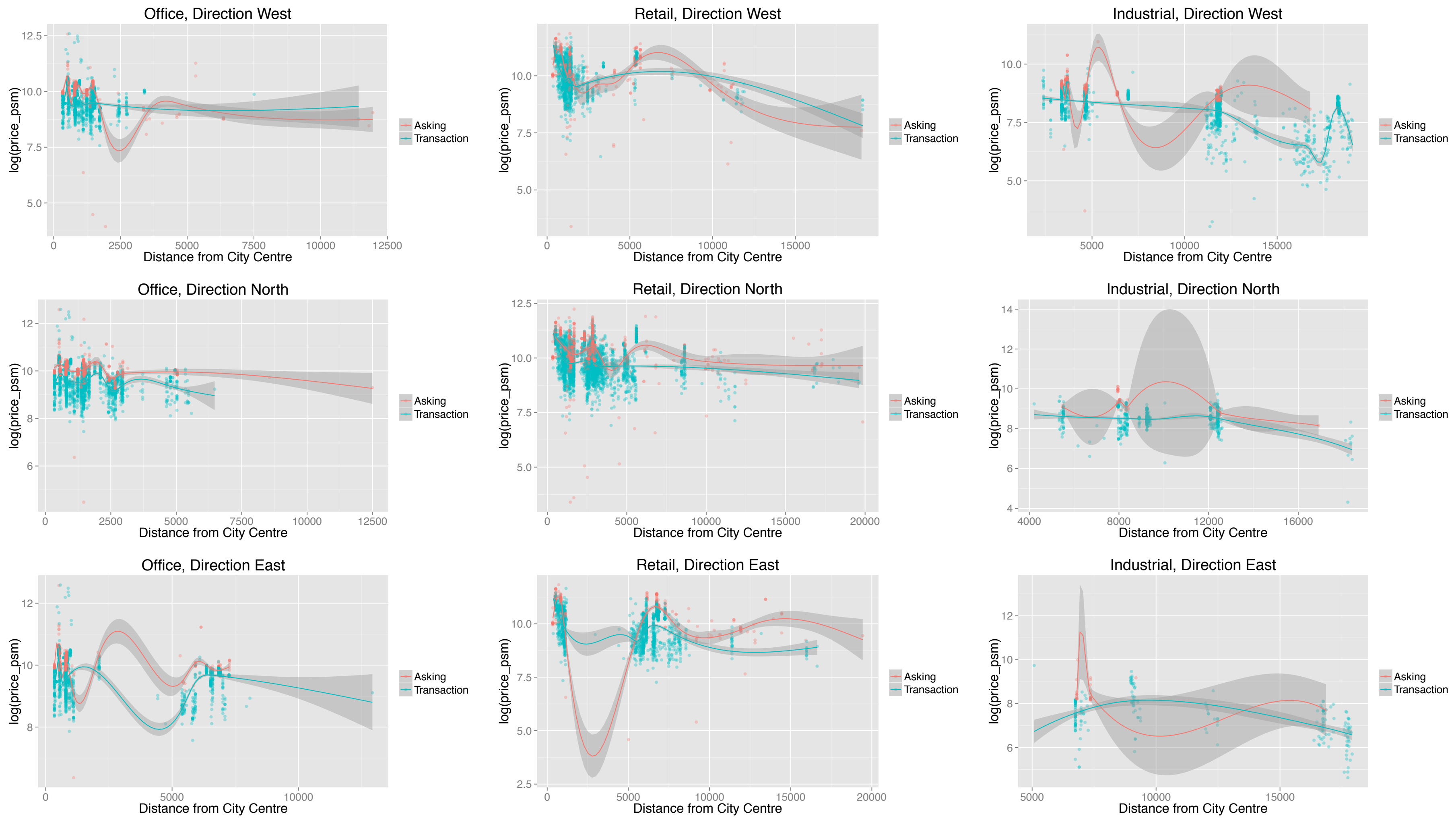


4.5 Neighbourhood variables

The neighbourhood of a property has an influence on the price. The neighbourhood describes the influence of different characteristics in the surrounding of the property. The location describes where the property is located more on a higher level.
The percentage of the different land uses in the neighbourhood is calculated. The land use information is taken from the Master Plan and aggregated in the following four categories: residential, industry, commercial and green space. Green space aggregates all parks, sports areas, recreational and open spaces and water body areas. For the calculation all areas 500 m around the property are considered. The open space indicator has to be analysed with caution, as not all green spaces are accessible for the public.

Previous studies have used the vacancy rate of the building or of the neighbourhood as an independent variable. The influence however is controversial. For Singapore the vacancy rates are available for the different usage types on the level of the Sub Development Guide Plan (sub DGB) area, except for the industrial category, where the data is only available on the DGP level. The data is subdivided into a public and private rate. These values are downloaded from the REALIS platform and matched to the listings. The data was retrieved for the last quarter of 2012.
The median private vacancy rate for the office listings is approximately 10 %, for retail the value is 6 % and for the industrial usage about 8 %. The values seem to be more or less randomly distributed over the planning areas. The computed average vacancy rate in the neighbourhood of the listings is similar to the actual ones (see Section 2.2.7).

Figure 16: Price per square meter analysis for the cardinal directions

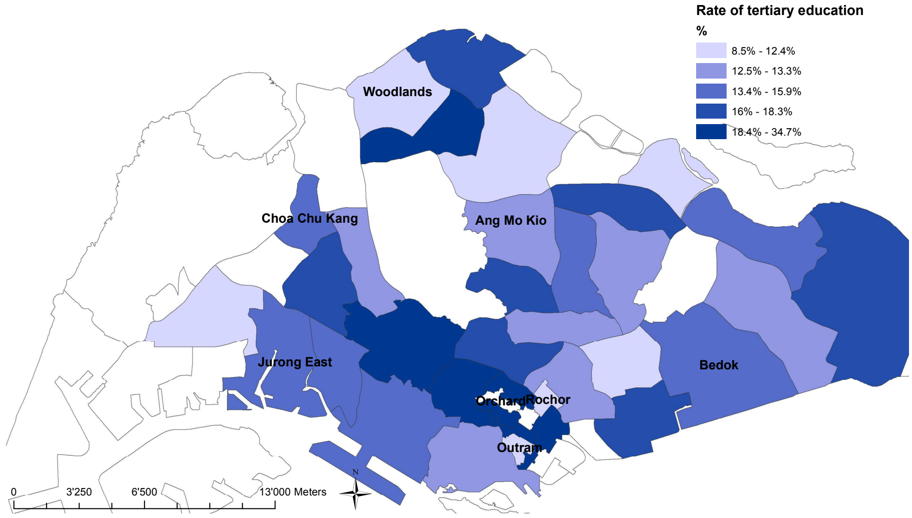


4.5.1 Tertiary education and income

To test the influence of the education level and the income of the population in the neighbourhood of the property, these variables are included. The data is only available on the Development Guide Plan (DGP) level. The data is taken from the department of statistics Singapore (Department of statistics Singapore, 2013).

Figure 17: Spatial distribution of the tertiary education and average income

(a) Percentage of population with a tertiary education degree



(b) Average income based on DGP

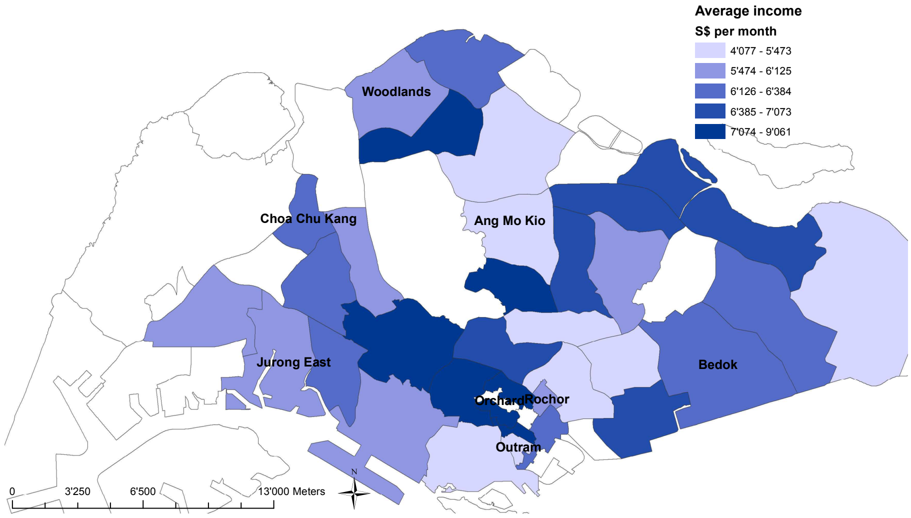


Figure 17(a) shows the share of population with a tertiary education degree. The highest rates can be found in the area of "Mandai", "Bukit Timah" and around the city centre. Figure 17(b) shows the average income per person in the different areas. The absolute value of the average income has to be interpreted with caution. The data on the income of the population is only available in categories. For example the data says how many people earn more than 10'000 S$. To calculate the average, the categories had to be split up in an absolute value (mean of each category) and normalized by the count of population living in the area. The map shows that the distribution of the income is similar to the one of the tertiary education.

4.5.2 Cluster indicators

It is assumed that clusters are especially important for office units, but also other land use types cluster together at certain locations. The influence of such clusters is investigated by using two different approaches. The first and simplest one is to check whether an unit is inside a defined area or not. The areas for this are the CBD and Orchard Road planning area defined by the Urban Redevelopment Authority (2013b). The boarders of the areas can be seen in Figure 19. The CBD area is used for the office properties and the Orchard Road area for the retail properties, as the main shopping area in Singapore is located around Orchard Road. These two areas are administrative boundaries for planning purposes and do not indicate cluster locations very well. Because of this a second method is used. All the available listings are used to find areas where the density of listings is higher than in other regions. The algorithm (Stack Exchange Inc., 2013) creates an area around all listings which are not further away as a defined critical distance. With this distance the cluster sizes can be controlled. The cluster areas are calculated for each main usage type. The resulting cluster areas are shown in Figure 19. The critical distance used for the office clusters is 250 m, for retail 150 m and for industrial 250 m. To describe the importance of each cluster, the percentage of the total amount of listings is calculated for each cluster area and included as variable.

Figure 18 shows the box plot of the price by the office category. The price is higher if the office is located inside the CBD or inside the office cluster area. For the properties outside of theses areas, the average price per square meter for the asking and transaction data is nearly the same. This is especially surprising, because if the price is considered over the whole island there is a huge gap between the average prices (see Table 10). For the other usage types (retail and industry) the results of the clustering are not so clear, but there are still clustering effects visible.

Figure 18: Box plots for office data and clustering effects

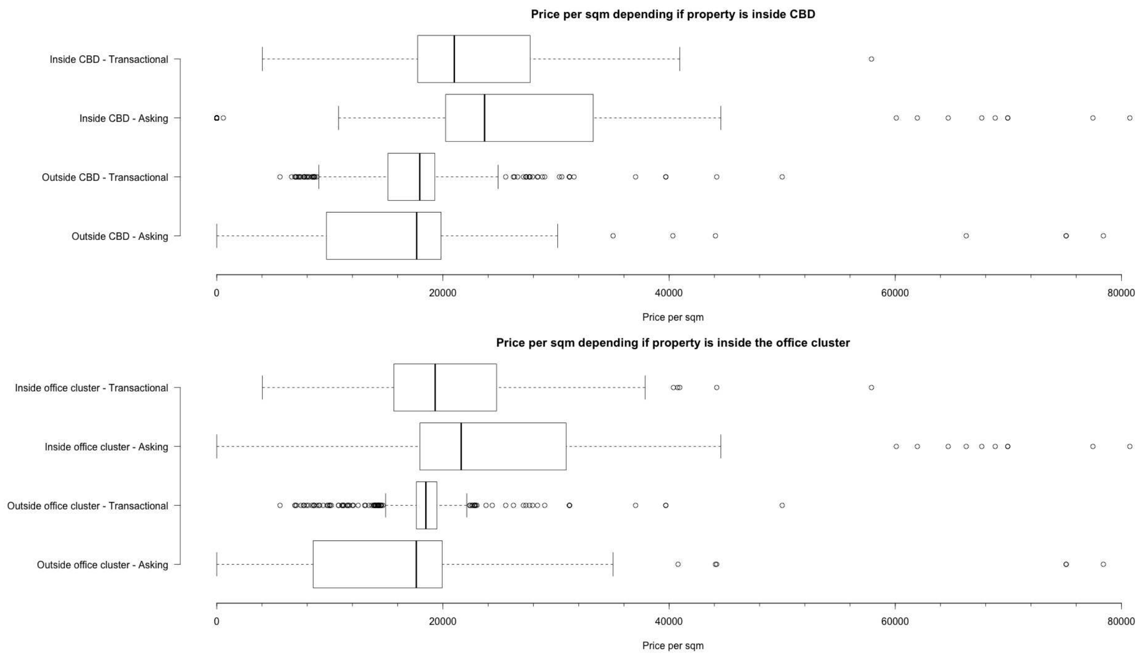
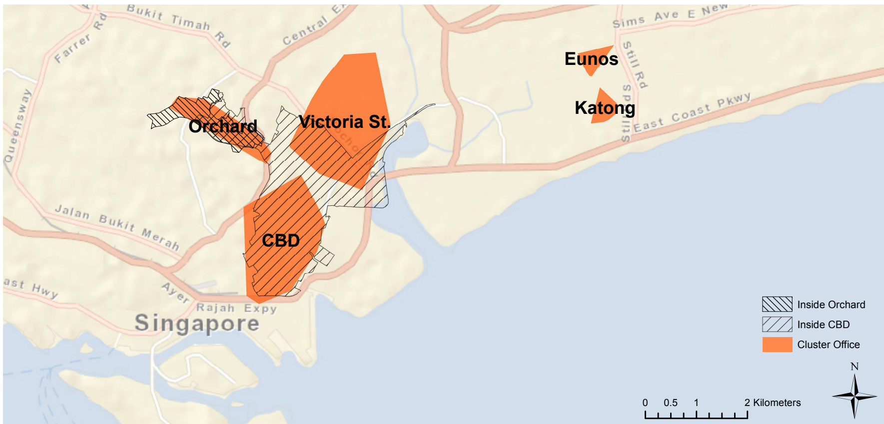
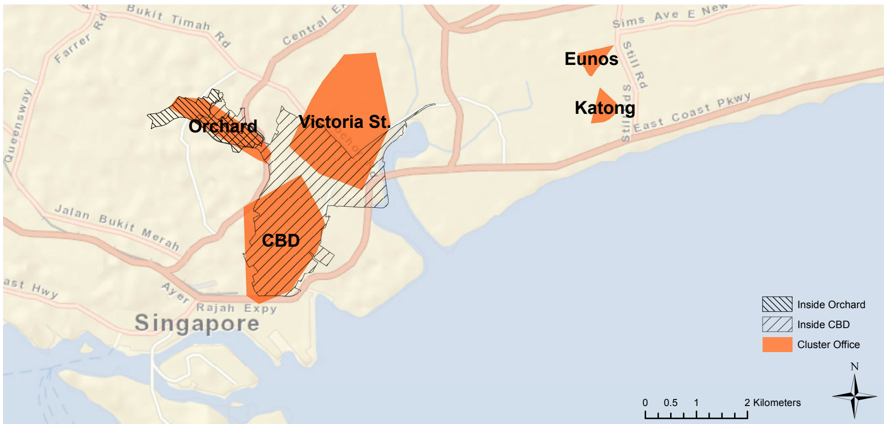


Figure 19: Cluster definition

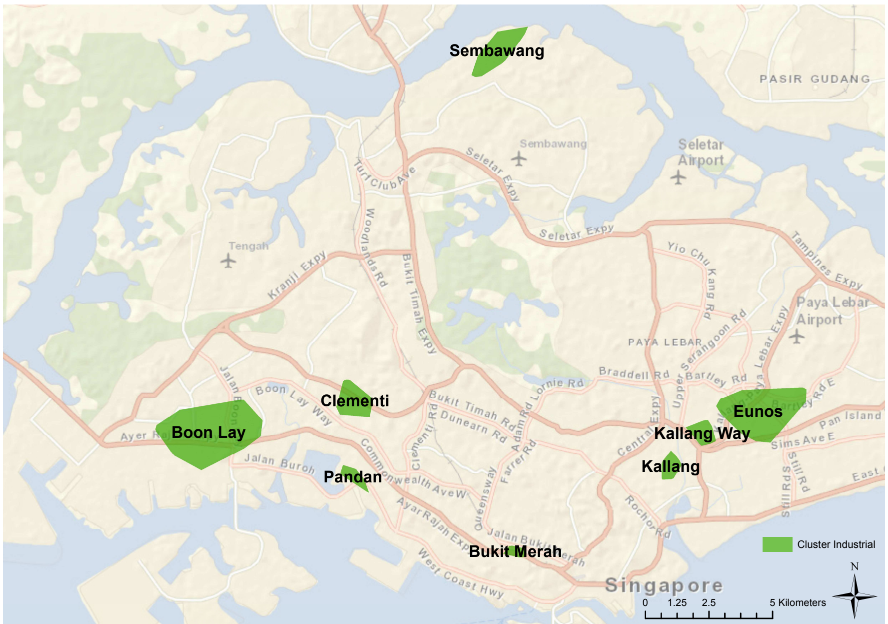
(a) Office



(b) Retail



(c) Industrial



Source: Background: Esri Inc. (2013)

5 Model estimations

This chapter describes the different models and the results. Firstly, Section 5.1 describes the variables used in the models. The following section gives an overview of the models which are presented in this study (Section 5.2). Section 5.3 gives the descriptive statistics for the data and Sections 5.4 to Section 5.7 present the results.

5.1 Variables definition

Table 12 gives an overview of the different variables used in the following models. The definitions of the variables are the same over all models and subgroups, but not all variables are available for all models. This is noted in the last two columns of the Table 12. Some variables are only intended to be used in a specific market. For example the variable "in_shopping_mall" is only used in the retail models. Table 12 also shows the expected sign and the expected strength of the variables. For specific effects, different definitions of the variables are used. This allows to test which specification works best for which data type. For example for the access to the bus stop five different variable definitions are used and tested against each other to find the best working one. The data and sources used are described in Chapter 4 in more detail.

Table 12: Definition and description of the variables

Name in models	Description	Units	Sign	T	A
Dependent variables					
price	Price of the unit	S$		x	x
price_psm	Price per square meter	S$/$m^2$		x	x
Structural variables					
floor_area_sqm	Floor area	m^2	++	x	x
floor_plot_ratio	Floor plot ratio	-		x	x
contract_build_area_ratio	Contract to building area ratio	-	-	x	x
floor_number_fact	Floor number	-	+	x	
condition	Condition (factor: 1:fitted / 0:bare)	D	+		x
gpr_num	Ground plot ratio out of Master Plan	-	+	x	x
aircon_available	Air-conditioning available	D	+		x
aircon_central	Air-conditioning centrally managed	D	+		x
floor_height	Floor height	D	+		x
cargo_lift	Cargo lift available	D	+		x
lift_capacity	Lift capacity in terms of weight	D	+		x
electricity_supply	High voltage electricity supply	D	+		x

#: Quantity; D: Dummy variable; -: Dimensionless

T: Available for transaction data; A: Available for asking data

Name in models	Description	Units	Sign	T	A
floor_loading	Floor loading	D	+		x
parking_available	Parking available	D	+		x
Locational variables					
distance_bus	Distance to next bus stop	m	-	x	x
distance_mrt	Distance to next MRT station	m	-	x	x
distance_cbd	Distance to the CBD	m	- -	x	x
distance_airport	Distance to the airport	m	-	x	x
distance_hgw	Distance to next highway ramp	m	-	x	x
distance_car_parking	Distance to next car parking	m	-	x	x
distance_truck_parking	Distance to next truck parking	m	-	x	x
dist_per_bus	Sum of distance to next 5 bus stations / number of bus lines	m	-	x	x
num_bus_station_500	Number of bus stations in 500 m	#	+	x	x
num_bus_line_500	Number of bus lines in 500 m	#	+	x	x
bus_lines_fact	Number of bus lines at next 5 st.	#	+	x	x
bus_number_unique	Number of unique bus lines at next 5 stations	#	+	x	x
in_cbd	Inside CBD	D	++	x	x
in_orchard	Inside Orchard	D	+	x	x
in_cluster_landuse	In cluster of land use type	D	++	x	x
cluster_size_landuse	Cluster size of land use type in % of total amount of transactions	%	++	x	x
cluster_landuse	The different clusters used as a factor variable	D	+/-	x	x
in_shopping_mall	In shopping centre	D	++	x	x
connected_to_mrt	Shopping mall connected to MRT station	D	+	x	x
accessibility_work	Hansen based accessibility for work		++	x	x
accessibility_residential	Hansen based accessibility for res.		++	x	x
Neighbourhoood variables					
average_income	Average income	D	+	x	x
perc_tertiary_edu	% Tertiary education	D	+	x	x
perc_open	% Open space in 500m	%	+	x	x
perc_res	% Residential land use in 500m	%	-	x	x
perc_ind	% Industrial land use in 500m	%	-	x	x
perc_com	% Commercial land use in 500m	%	++	x	x
land_use_vac_rate_public	Vacancy rate public	%	-	x	x
land_use_vac_rate_private	Vacancy rate private	%	-	x	x
total_lots	Total parking lots	#	+	x	x
car_lots	Car parking lots	#	+	x	x
trailer_lots	Trailer parking lots	#	+	x	x

#: Quantity; D: Dummy variable; -: Dimensionless

T: Available for transaction data; A: Available for asking data

Name in models	Description	Units	Sign	T	A
work_density_500	Working population in 500m	#	+	x	x
totpax_500	Population in 500 m	#	+	x	x
Contractual variables					
tenure_type	Leasehold	D	+	x	x
tenure_type	Freehold	D	++	x	x
tenure_type	Tenure count years	#	+	x	x
tenure_ongoing	Duration of tenure from year 2013	Years	+	x	
contract_date	Contract date	Date	+/-	x	
top_year	Year of Temporary Occupation Period: Date when the first owners can move in to the unit	Year	+		x

#: Quantity; D: Dummy variable; -: Dimensionless
T: Available for transaction data; A: Available for asking data

5.2 Model overview

In this study different models are estimated. The models can be categorized into three different types: the best performing or full models, the comparison models and the one-variable models. The following list describes the models in more detail. The data for the models is always the same and is presented in Section 5.3. Combined datasets are not used in any model.

- **Best performing models:** This model type is estimated for each dataset. The specification of the model depends on which variables work best for each dataset. That makes the results difficult to compare (see Section 5.4).
- **Comparison Models:** These models allow for comparison of asking and transaction data. To do so, for each market (office, retail and industrial) a model is estimated containing the same variables. This allows to directly compare the strength of each included variable in the different markets (see Section 5.5).
- **One-variable models:** The one-variable models show the isolated influence of only one important determinant on the price.
  - **Cluster/location analysis model:** To compare the influence of different locations across the island, defined areas with higher density of listings are used to examine if they have an influence on the price (see Section 5.6).
  - **Access/Accessibility analysis model:** The accessibility of a property plays an important role for the price. To show which access or accessibility determinant performs best, an analysis is done in Section 5.7.

5.3 Descriptive statistics

This section gives an introduction to the data used. The data is first of all subdivided into transaction and asking data and into the three land use categories: office, retail and industrial. Further is the transaction data subdivided into strata and land listings and the asking data into sale and rent (see Section 4.2.2 for details). Table 13 gives an overview of the descriptive statistics for all types of data. This table only includes the price per square meter and two major variables. A full descriptive statistic for each data type can be found in Appendix A. Table 13 shows in the right column the table number for the full descriptive statistic.

Table 13: Overview of the descriptive statistics

			Number of listings	Median price psm [S$ /sqm]	St.Dev. price psm [S$ / sqm]	Median floor area [sqm]	Median accessibility_work	Table with full desc. statistics
Transaction	Office	Strata	1271	18117.78	7051.31	65.00	131430.80	Tab. 34
	Retail	Strata	2756	30478.18	20080.99	28.00	108510.99	Tab. 37
		Land	670	25784.62	14016.72	142.00	97050.07	Tab. 38
	Industrial	Strata	9292	4832.95	2420.93	181.00	60754.21	Tab. 41
		Land	1545	1372.78	3224.21	4805.00	17889.24	Tab. 42
Asking	Office	Sale	1576	20499.39	11377.76	108.00	124416.52	Tab. 35
		Rent	7333	75.35	39.01	110.46	118847.06	Tab. 36
	Retail	Sale	3049	39559.34	29796.66	55.00	113930.73	Tab. 39
		Rent	5860	114.16	93.75	55.74	109511.80	Tab. 40
	Industrial	Sale	5402	6458.35	3076.39	216.88	60008.57	Tab. 43
		Rent	7635	22.64	11.16	260.87	55946.07	Tab. 44

Table 13 shows that the median of the asking price is higher than the transaction price. Furthermore the standard deviation of the price per square meter is also higher for the asking data. The highest prices are paid for the retail properties. The median of the floor area is the highest in the industrial sector, followed by the office sector and the retail sector. The asking listings tend to have higher floor areas compared to the appropriate categories of the transaction data. The working accessibility is the highest for the office sector, followed by the retail and the industrial sector. This makes sense, as the office properties are mainly located in the CBD, where the accessibility is highest. The retail properties are spread over the whole island, where the accessibility is not that high. The industrial land units have a low accessibility. A reason for this could be, that in these new areas the data for the accessibility calculations is not up to date or that these land plots have a low accessibility in reality.

5.4 Results of the best models

First the results of the best performing models are reported. The models are created by including all significant variables, but with exclusion of correlated variables. All variables which have a correlation coefficient higher or smaller than 0.4/-0.4 are excluded. The results tables (Tables 15 to 25) include the coefficients for the linear model and the three different spatial models (SAR_{err}, SAR_{mix} and SAR_{lag}). Furthermore the standardized coefficients are (see Section 3.1.1) calculated and reported for the linear model.

Table 14: Overview of the regression results for the best models

			Dependent var.	Num. of variables *	Linear model Adj. R^2	Linear model Moran's I	SAR_{err} Pseudo R^2	SAR_{mix} Pseudo R^2	Weights matrix*
Office	Transaction	Strata	log(price_psm)	9	0.63	0.46	0.79	0.80	12 NND
	Asking	Sale	log(price_psm)	6	0.56	0.83	0.92	0.98	6 NND
		Rent	log(price_psm)	7	0.59	0.74	0.86	0.87	6 NND
Retail	Transaction	Strata	log(price_psm)	10	0.62	0.55	0.80	0.82	8 NND
		Land	log(price_psm)	5	0.54	0.65	0.76	0.78	8 NND
	Asking	Sale	price_psm	10	0.52	0.38	0.61	0.62	16 NND
		Rent	log(price_psm)	7	0.42	0.55	0.66	0.66	16 NND
Industrial	Transaction	Strata	log(price_psm)	9	0.79	0.65	0.94	0.94	16 NND
		Land	price_psm	7	0.80	0.79	0.95	0.96	16 NND
	Asking	Sale	log(price_psm)	8	0.72	0.62	0.87	0.87	12 NND
		Rent	log(price_psm)	6	0.53	0.59	0.75	0.71	12 NND

*Factors counts as 1 variable; *k* NND: *k* nearest neighbours weighted by distance

Table 14 gives an overview of the results. All models except the "Retail asking sale" and the "Industrial transaction land" are estimated by using the logarithmic price per square meter. Those

two models perform better without the logarithm. The model fit varies between 0.42 and 0.80 in the linear models and between 0.66 and 0.98 in the spatial models. The SAR_{mix} specifications works best for most models, but the differences between the SAR_{mix} and the SAR_{err} are small. Models with a high spatial autocorrelation in the linear model (high value of the Moran's I test statistic) tend to have a higher difference between the linear and the spatial models in the R^2. This can be seen by comparing the "Office asking sale" model with the "Retail asking sale model". The R^2 for the office model improves by nearly 0.3, the retail model R^2 with a low spatial autocorrelation improves only by 0.09. The high autocorrelation in the office model is assumed to happen as office property is less spatially dispersed than retail property.
All spatial models are estimated using a k nearest neighbours matrix, which is weighted by the distance. Other methods for calculating the weights did not work as well. The nearest neighbours by distance algorithm generated similar results.

Model fit increases with spatial models. The SAR_{lag} model generally does not remove all spatial autocorrelation. This fact can be seen over all estimated models. For the SAR_{lag} models, the spatial autocorrelation is assumed to happen in the dependent variable. This means, that for the used data the autocorrelation is not happening mainly in the dependent variable. The SAR_{err} model performs a little worse than the SAR_{mix} model. The parameters of the SAR_{err} are comparable to the coefficients in the linear OLS model. This is because the autocorrelation is only removed in the error term and not in the explanatory variables. The coefficients SAR_{mix} are sometimes very different to the linear ones. The sign of the coefficient can change the direction. This fact can be explained by the method, since the SAR_{mix} model reduces the spatial autocorrelation in the dependent and independent variables.
Furthermore some variables are not any more significant in the spatial models. This happens especially in the SAR_{mix} models, as there the spatial autocorrelation is also removed in the independent variables. This is not the case for the SAR_{lag} and SAR_{err} models. Overall the SAR_{err} models are preferred because of the easier interpretation of the results and because of the simpler model specifications.

5.4.1 Transaction office model

Table 15 shows the results of the transaction office model. This model includes only strata entries, as in the transaction office data no land entries are available (as for retail and industrial). The transaction office model includes nine variables. The following variables are not significant: perc_open, perc_res, distance_airport and floor_plot_ratio. Further some variables are highly correlated to each other, especially the in_cbd and the distance_cbd. In this model the distance_mrt is correlated to bus_number_unique. This happens only in that model. The most important variable are the tenure conditions (tenure_ongoing) and the access (bus_number_unique).

In the spatial models the floor number shows higher significance. The results show, that the prices at the lower floors are the highest. It is also visible, that the higher the office is located in the building, the higher the rent is, but the effects are rather small. The SAR_{mix} model did not perform very well for this data. The variable cluster_office had to be excluded because otherwise the estimation did not work. Because of these estimation problems, the SAR_{err} model is assumed to be the best performing.

Table 15: Regression results for the office transaction model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	10.01		***	10.31	***	2.52	***	3.35	***
log(floor_area_sqm)	-0.06	-0.15	***	-0.10	***	-0.11	***	-0.05	***
office_vac_rate_public	0.00	0.12	***	0.00	*	-0.04	*	0.00	***
log(distance_hgw)	-0.14	-0.20	***	-0.13	**	0.20	*	-0.06	***
car_lots	0.00	-0.40	***	0.00	***	0.00		0.00	***
log(tenure_ongoing)	0.09	0.60	***	0.07	***	0.01		0.03	***
contract_date									
2011 Q1	0.00			0.00		0.00		0.00	
2011 Q2	0.01			0.05	*	0.03		0.02	
2011 Q3	0.11		***	0.09	***	0.09	***	0.08	***
2011 Q4	0.06		*	0.08	***	0.05	*	0.06	**
2012 Q1	0.15		***	0.13	***	0.12	***	0.12	***
2012 Q2	0.21		***	0.14	***	0.11	***	0.13	***
2012 Q3	0.22		***	0.19	***	0.17	***	0.16	***
2012 Q4	0.20		***	0.21	***	0.19	***	0.16	***
2013 Q1	0.23		***	0.23	***	0.21	***	0.20	***
floor_number									
< 3	0.00			0.00		0.00		0.00	
3 - 5	-0.06		**	-0.17	***	-0.17	***	-0.08	***
6 - 10	-0.02			-0.17	***	-0.18	***	-0.06	***
11 - 15	0.08		**	-0.14	***	-0.16	***	-0.03	
> 16	0.08		**	-0.12	***	-0.12	***	-0.01	
cluster_office									
Outside	0.00			0.00		-		0.00	
Katong	0.21		***	0.27	**	-		0.06	*
Orchard	0.38		***	0.42	**	-		0.16	***
Eunos	0.13		***	0.16	.	-		0.07	**
CBD	0.48		***	0.52	***	-		0.20	***
Victoria St.	0.32		***	0.34	***	-		0.16	***
bus_number_unique									
[0,22)	0.00			0.00		0.00		0.00	
[22,33)	0.17		***	0.20	***	-0.01		0.07	***
[33,60)	0.10		***	0.09		-0.66	*	0.04	.
[60,81)	0.23		***	0.26	**	-0.42		0.06	*
Adj. R^2/PseudoR^2	0.63			0.79		0.80		0.77	
AIC	-176.05			-763.89		-787.51		-670.93	
Moran's I	0.46		***	0.03	**	0.02	.	0.09	***
Lambda / rho				0.76	***	0.72	***	0.68	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.2 Transaction (strata) retail model

Table 16 shows the results of the transaction (strata) retail model. For this model the parking description variables and most neighbourhood variables (perc_com, perc_res, perc_tertiary_edu, average_income) are not significant. The distance_cbd and distance_hgw are correlated.
The floor area of the retail property is the most important variable. Further the tenure and the access are important, a leasehold of 999 years has a positive effect on the price compared to freehold. For retail the floor is more important than for the other categories. The ground floor is the priciest floor. The price for a property located inside a shopping mall is higher than if it is located outside.
The vacancy rate is only significant in the linear model and there the influence is very small. Interesting is that the percentage of the open land use has a negative effect on the price. This means, the more open land use, the lower the price. The influence of this variable is however small.
The SAR_{err} model performs the best. The parameters of the SAR_{mix} model are hard to interpret, especially for the variable cluster_retail. Because of this the SAR_{err} model is the best specification for the transaction (strata) retail model.

Table 16: Regression results for the transaction (strata) retail model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	11.49		***	11.31	***	3.52	***	4.29	***
log(floor_area_sqm)	-0.22	-0.29	***	-0.14	***	-0.15	***	-0.15	***
in_shopping_mall	0.68	0.26	***	1.00	***	0.08		0.40	***
log(distance_mrt)	-0.15	-0.22	***	-0.18	***	-0.45	***	-0.06	***
bus_number_unique	0.01	0.28	***	0.01	***	0.01	**	0.00	***
perc_open	-0.74	-0.12	***	-0.76	**	0.22		-0.07	
retail_vac_rate_public	-0.01	-0.06	***	0.00		0.00		0.00	
contract_date									
2011 Q1	0.00			0.00		0.00		0.00	
2011 Q2	-0.04			0.11	***	0.12	***	0.07	**
2011 Q3	0.04			0.10	***	0.12	***	0.06	*
2011 Q4	0.09		*	0.26	***	0.27	***	0.19	***
2012 Q1	0.13		***	0.29	***	0.25	***	0.18	***
2012 Q2	0.22		***	0.32	***	0.31	***	0.23	***
2012 Q3	0.35		***	0.31	***	0.32	***	0.26	***
2012 Q4	0.31		***	0.35	***	0.38	***	0.32	***
2013 Q1	0.28		***	0.47	***	0.45	***	0.33	***
floor_number									
not available	0.00			0.00		0.00		0.00	
basement	0.18		***	0.12	*	0.00		-0.13	***
ground floor	0.45		***	0.31	***	0.22	***	0.10	***
2	0.23		***	0.06		-0.06	***	-0.10	**
> 2	0.09		.	-0.14	**	-0.25	***	-0.23	***
cluster_retail									
Outside	0.00			0.00		0.00		0.00	
CBD	-0.20		***	-0.11		-87.83	***	-0.03	
Orchard	0.27		***	-0.20	*	271.17	***	0.08	**
Katong	-0.30		***	-0.20	**	0.01		-0.15	***
Geylang Rd.	-0.21		***	-0.01		-1.01		-0.05	
Rochor	-0.09		***	0.19	**	0.07		0.08	***
tenure_type									
Freehold	0.00			0.00		0.00		0.00	
99 Years	-0.33		***	-0.28	***	-0.22	**	-0.09	***
999 Years	0.18		***	0.20	***	-0.10		0.10	***
Adj. R^2/PseudoR^2	0.62			0.80		0.82		0.78	
AIC	1816.94			168.63		-79.16		388.16	
Moran's I	0.55		***	-0.07		-0.03		0.02	**
Lambda / rho				0.76	***	0.70	***	0.65	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.3 Transaction (land) retail model

Table 17 shows the results for the transaction (land) retail model. This model is similar to the strata model shown previously. The model fit is approximately the same. For the land data fewer variables are available. A leasehold of 999 years is assumed to be equal to a freehold property. The properties located inside the CBD are much more expensive than those located at other locations. This effect is not that big for the strata model. The floor area of the property is not as important as in the strata model.

Table 17: Regression results for the transaction (land) retail model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	10.62			9.85	***	4.64	***	5.00	***
log(floor_area_sqm)	-0.21	-0.19	***	-0.06	*	-0.04		-0.11	***
bus_number_unique	0.01	0.14	***	0.01	***	0.01	*	0.00	***
contract_date									
2011 Q1	0.00			0.00		0.00		0.00	
2011 Q2	0.13		**	0.15	***	0.15	***	0.10	**
2011 Q3	0.10		*	0.11	**	0.10	*	0.07	.
2011 Q4	0.27		***	0.21	***	0.21	***	0.19	***
2012 Q1	0.22		***	0.23	***	0.23	***	0.19	***
2012 Q2	0.35		***	0.34	***	0.34	***	0.28	***
2012 Q3	0.31		***	0.33	***	0.31	***	0.27	***
2012 Q4	0.37		***	0.44	***	0.41	***	0.34	***
2013 Q1	0.51		***	0.47	***	0.47	***	0.45	***
cluster_retail									
Outside	0.00			0.00		0.00		0.00	
CBD	0.78		***	0.76	***	0.47	*	0.35	***
Katong	0.05			0.07		-0.40		0.02	
Geylang Rd.	-0.05			-0.09		-0.03		-0.05	
Rochor	0.23		***	0.25	***	-0.02		0.11	***
tenure_type									
Freehold	0.00			0.00		0.00		0.00	
99 Years	-0.27		***	-0.30	***	-0.39	***	-0.16	***
999 Years	0.00			0.00		-0.06		-0.01	
Adj. R^2/PseudoR^2	0.54			0.76		0.78		0.74	
AIC	481.50			160.35		153.81		201.04	
Moran's I	0.65		***	0.06	*	0.06	*	0.13	***
Lambda / rho				0.61	***	0.58	***	0.52	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.4 Transaction (strata) industrial model

Table 18 shows the results for the transaction (strata) industrial model. The distance_hgw, distance_airport, perc_ind, work_density, average_income and the vacancy rates are not significant. Correlated are the distance_cbd and the distance_car_park. For this model the floor area as well as the access have a high influence on the price. The cluster variable shows, that the industrial estates on the western side of the island are much cheaper than those at other locations. The dist_truck_parking has an unexpected sign. The coefficient states, that if the distance increases the price is increasing as well. A possible explanation for this unexpected sign is that the variable probably measures another unknown influence. Another problem with the distance to the next truck parking is that only the public truck parks are included in the calculations. This does not describe the distance to a truck parking very well.
The influence of the gpr_num is positive, as expected. The strength of the parameter is rather small.

The model fit of this model is quite high, especially the spatial models increase the fit to 0.95. The spatial autocorrelation in the data can be reduced with the SAR_{err} and the SAR_{mix} model. As the model fits of these two models are approximately the same, the SAR_{err} model is concluded to be the best because of easier interpretable parameters. The sign of some parameters for the SAR_{mix} model are unexpected for some variables.

Table 18: Regression results for the transaction (strata) industrial model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	9.59		***	9.36	***	1.26	***	2.38	***
log(floor_area_sqm)	-0.25	-0.30	***	-0.11	***	-0.10	***	-0.09	***
gpr_num	0.14	0.09	***	0.12	*	-0.86	***	0.02	***
bus_number_unique	0.02	0.27	***	0.02	***	0.04	***	0.00	***
log(distance_mrt)	-0.11	-0.15	***	-0.17	***	-0.01	***	-0.02	***
log(dist_truck_parking)	0.08	0.13	***	0.08	***	-1.00	***	0.00	*
contract_date									
2011 Q1	0.00			0.00		0.00		0.00	
2011 Q2	0.14		***	0.10	***	0.11	***	0.11	***
2011 Q3	0.18		***	0.16	***	0.18	***	0.16	***
2011 Q4	0.29		***	0.21	***	0.21	***	0.20	***
2012 Q1	0.26		***	0.24	***	0.24	***	0.22	***
2012 Q2	0.30		***	0.28	***	0.28	***	0.25	***
2012 Q3	0.32		***	0.31	***	0.31	***	0.28	***
2012 Q4	0.43		***	0.34	***	0.35	***	0.31	***
2013 Q1	0.45		***	0.37	***	0.38	***	0.34	***
floor_number									
<= 1	0.00			0.00		0.00		0.00	
2 - 3	-0.28		***	-0.22	***	-0.23	***	-0.20	***
4 - 5	-0.29		***	-0.23	***	-0.24	***	-0.20	***
6 - 8	-0.26		***	-0.24	***	-0.24	***	-0.20	***
> 8	-0.19		***	-0.22	***	-0.23	***	-0.18	***
cluster_industrial									
Outside	0.00			0.00		0.00		0.00	
Bukit Merah	-0.13		***	0.30	*	-0.01		-0.02	.
Kallang	0.07		*	0.45	***	-0.55	*	0.00	
Boon Lay	-0.68		***	-1.44	**	4.06	***	-0.33	***
Pandan	-0.19		***	0.30	**	-0.01		-0.04	**
Eunos	0.06		***	0.03		0.01	***	0.00	
Sembawang	0.02			0.02		-0.01		-0.07	***
Kallang Way	-0.05		***	-0.03		-11453.59	***	-0.05	***
Clementi	0.07		***	0.00		2364.13	***	0.01	.
tenure_type									
Freehold	0.00			0.00		0.00		0.00	
30 Years	-0.77		***	-0.96	***	-3.63	***	-0.22	***
60 Years	-0.39		***	-0.35	***	-2.18	***	-0.10	***
99 Years	-0.05		.	-0.42	***	-2.26	***	-0.01	
999 Years	0.18		***	-0.23		-9.40	***	0.05	***
Adj. R^2/PseudoR^2	0.79			0.94		0.94		0.93	
AIC	-672.14			-11366.06		-11744.34		-10418.52	
Moran's I	0.65		***	-0.03		-0.02		0.10	***
Lambda / rho				0.92	***	0.88	***	0.79	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.5 Transaction (land) industrial model

Table 19 shows the results for the transaction (land) industrial model. This model uses the price per square meter as dependent variable. The performance of the model is comparable to the strata model. The leasehold of 999 years is not similar to the freehold properties in this model. All spatial models cannot remove the spatial autocorrelation of this data. The Moran's I statistics is for all models significantly different form 0.

Table 19: Regression results for the industrial transaction land model

Dependent: price_psm

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	15838.58			13701.34	***	3632.51	***	5537.93	***
log(floor_area_sqm)	-783.38	-0.30	***	-782.40	***	-604.30	***	-294.69	***
log(distance_mrt)	-364.43	-0.08	***	-529.29	***	-6971.85	***	-127.60	***
log(dist_truck_parking)	129.65	0.05	***	453.81	***	786.24	***	50.46	***
bus_number_unique	94.02	0.19	***	107.84	***	-2.10		33.54	***
contract_date									
2011 Q1	0.00			0.00		0.00		0.00	
2011 Q2	-116.97			-122.57		155.53		-46.33	
2011 Q3	-35.43			99.22		279.59	.	49.12	
2011 Q4	221.53			968.25	***	1454.98	***	179.13	.
2012 Q1	171.35			359.55	*	877.08	***	138.11	
2012 Q2	671.41		***	1042.71	***	1490.36	***	349.83	***
2012 Q3	892.58		***	1525.80	***	1952.88	***	517.21	***
2012 Q4	956.65		***	1871.12	***	2364.56	***	589.47	***
2013 Q1	1415.64		***	1599.14	***	1838.64	***	703.44	***
cluster_industrial									
Outside	0.00			0.00		0.00		0.00	
Bukit Merah	476.73			-5813.09	***	-15603.21	***	5.41	
Kallang	-2709.92		***	-2481.98	***	-176.76		-1128.08	***
Boon Lay	-827.65		***	-1141.14	***	18262.82	***	-317.43	***
Pandan	209.19			544.13		37009.18	**	246.19	
Eunos	2201.88		***	1541.29	***	-6513.43	***	602.02	***
Sembawang	-291.68			177.33		5892.95	***	-142.81	
Kallang Way	718.51		*	628.94		20090.12	***	205.83	
Clementi	434.59			-1112.90	.	-27028.74	***	111.97	
tenure_type									
Freehold	0.00			0.00		0.00		0.00	
30 Years	-6285.55		***	-4784.24	***	-3697.48	***	-2129.37	***
60 Years	-5668.33		***	-4404.00	***	-3475.90	***	-1866.00	***
99 Years	-6051.63		***	-5425.80	***	-4739.85	***	-2079.76	***
999 Years	-5201.11		***	-3595.84	***	-3291.14	***	-1665.50	***
Adj. R^2/PseudoR^2	0.80			0.95		0.96		0.93	
AIC	26853.18			25552.43		25174.40		25866.63	
Moran's I	0.79		***	0.16	***	0.15	***	0.23	***
Lambda / rho				0.76	***	0.75	***	0.64	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.6 Asking (sale) office model

Table 20 shows the results for the asking (sale) office model. The following variables are not significant: car_lots, aircon_available, in_cbd, gpr_num as well as the vacancy rates. The access and the floor area of the property are the most important determinants for this sector. The tenure is not compared to freehold, but to the 30/60 years class. It can be seen, that a leasehold of 999 years has a similar effect on the price as freehold. The tertiary education in the neighbourhood has a small effect on the price. The price does not increases with an increasing tertiary education percentage. Reasons for this could be, that in the neighbourhood of the central business district more people live in old buildings than at remoter areas. Further the TOP year variable is included in this model, but the variable is only significant at a level of 5%.
The SAR_{err} and SAR_{mix} show similar model fits and the spatial autocorrelations are removed by both models. The SAR_{err} is considered as the best model, concerning the very few significant variables in the SAR_{mix} model.

Table 20: Regression results for the asking (sale) office model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	10.34			10.98	***	0.48	***	2.35	***
log(floor_area_sqm)	-0.12	-0.18	***	-0.11	***	0.39	***	0.30	***
log(distance_mrt)	-0.19	-0.27	***	-0.15	**	-0.01		-0.05	***
bus_number_unique									
[0,21)	0.00			0.00		0.00		0.00	
[21,33)	0.33		***	-0.21	***	-0.13	***	0.06	***
[33,50)	0.33		***	0.23	**	0.04		0.07	***
[50,87)	0.36		***	0.14		-0.06		0.08	***
tenure_type									
30/60 Years	0.00			0.00		0.00		0.00	
Not available	0.87		***	0.30	**	0.01		0.27	***
99/999 Years	0.84		***	0.42	***	0.07	**	0.25	***
Freehold	0.92		***	0.35	***	-0.02		0.28	***
perc_tertiary_edu									
[0, 0.124)	0.00			0.00		0.00		0.00	
Not available	0.09		*	-0.74	***	-0.64	***	0.06	***
[0.124,0.262)	0.21		***	0.05		-0.05		0.03	***
[0.262, 0.311)	0.14		*	-0.28	.	-0.56	***	0.07	***
top_year									
Not available	0.00			0.00		0.00		0.00	
-2013	-0.11		*	0.06	**	0.03	**	-0.02	
2013 - 2020	0.10		*	-0.03		-0.02		0.05	***
Adj. R^2/PseudoR^2	0.56			0.92		0.98		0.90	
AIC	1270.70			-1141.78		-3838.58		-2005.90	
Moran's I	0.83		***	-0.01		0.01		0.57	***
Lambda / rho				0.91	***	0.89	***	0.40	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.7 Asking (rent) office model

The results for the asking (rent) office model are shown in Table 21. The most important variables are the distance to the next MRT station and the floor area. The variable describing the bus access is correlated to the distance to the MRT. Not significant are the car_lots, the vacancy rate and the floor_plot_ratio. The results for this model show that a centralized air-conditioning has a negative effect on the price, but the effect is rather small. In the spatial models this variable is not significant any more. It is assumed that this variable does not work well because of the lack of data. The perc_tertiary_edu has a significant effect on the price in this model. Properties located in the CBD or in the Orchard area generate higher prices.
The SAR_{err} and the SAR_{mix} model have a similar model fit. The cluster_office variable is only significant in the SAR_{err} model where it shows meaningful results. The SAR_{err} is considered to be the best model.

Table 21: Regression results for the asking (rent) office model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	6.78			6.54	***	0.50	***	0.78	***
log(floor_area_sqm)	-0.10	-0.22	***	-0.13	***	-0.06	***	-0.03	***
contract_build_area_ratio	-0.01	-0.10	***	-0.01	***	0.00	***	0.00	***
aircon_central:True	-0.09	-0.03	***	0.01		0.00		-0.02	**
log(distance_mrt)	-0.22	-0.37	***	-0.24	***	-0.19	***	-0.03	***
log(distance_hgw)	-0.12	-0.16	***	-0.05	**	0.01		-0.01	***
cluster_office									
Outside	0.00			0.00		0.00		0.00	
Katong	-0.07		.	0.07		0.07		-0.02	.
Orchard	0.25		***	0.18	**	-0.01		0.01	
Eunos	-0.04			0.18		0.26		0.00	
CBD	0.30		***	0.34	***	0.05		0.03	***
Victoria St.	0.17		***	0.15	***	-0.12	**	0.01	***
perc_tertiary_edu									
Not available	0.00			0.00		0.00		0.00	
[0.085,0.133)	-0.38		***	-0.30	***	0.13	***	-0.05	***
[0.133,0.262)	-0.12		***	-0.13	.	0.02		-0.02	***
[0.262,0.311]	0.12		***	0.26	***	0.54	***	0.00	
Adj. R^2/PseudoR^2	0.59			0.86		0.87		0.84	
AIC	4626.39			-2482.27		-15002.90		-13998.64	
Moran's I	0.74		***	-0.04		-0.03		0.04	***
Lambda / rho				0.81	***	0.81	***	0.77	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.8 Asking (sale) retail model

Table 22 shows the results for the asking (sale) retail model. The model works best with the price per square meter as dependent variable. The distance of the MRT station and the floor area are the most important variables and have as expected a negative effect on the price. The tenure_type and if the property is located inside a shopping mall are also important. The leasehold of 999 years is not as similar to freehold as in other models. The TOP year also shows that if the building will be finished after 2013 the price per square meter is higher. If the property is already fitted (furnished), the price will be lower. In the retail sector the shops and restaurants are usually fitted according to the brand style. This could explain why the effect of this variable is negative, since the buyer cannot design and fit the property according to his idea. Another reason could be, that the fitted items need to be removed, which causes more costs. The cluster_retail variable shows that a property located in the Orchard area has a higher price. However, this works only in the linear model. The results in the SAR_{mix} model are very different and in the SAR_{err} the Orchard cluster is not significant.

The parking variables and the aircon_available variables are not significant for this model. It is assumed, that the data quality in terms of the availability of air-conditioning is not very good, which could explain why this variable is not working.

Noticeable in this model is, that the spatial autocorrelation in the linear model is with 0.38 relatively small compared to the values in the other models. This also leads to a smaller increase in the model fit of the spatial models. The floor_plot_ratio and the vacancy rate are not significant any more in the spatial models. A reason for this could be, that the effect in the linear model is only very small. The SAR_{err} and the SAR_{mix} model perform nearly equally. In the SAR_{err} more of the included variables are significant. Because of this the SAR_{err} model is considered as the best specification for this market.

Table 22: Regression results for the asking (sale) retail model

Dependent: price_psm

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	115443.46			112015.78	***	49169.47	***	73833.74	***
log(floor_area_sqm)	-7833.47	-0.29	***	-9138.50	***	-10034.35	***	-6734.81	***
floor_plot_ratio	-3781.55	-0.06	***	-1171.62		2010.18	*	-321.21	
retail_vac_rate_public	-270.58	-0.05	***	-127.27		164.72		-75.63	
log(distance_mrt)	-8724.04	-0.22	***	-6992.02	***	-5065.64	***	-4396.70	***
in_shopping_mall:True	24109.90	0.16	***	28182.37	***	-30242.96		13424.03	***
condition:True	-3487.47	-0.05	***	-2947.77	***	-3223.38	***	-2529.64	***
bus_number_unique									
[0,12)	0.00			0.00		0.00		0.00	
[12,24)	-4597.18		***	-7078.01	***	3993.72		-5318.93	***
[24,33)	9557.19		***	2401.03		2261.79		1370.24	
[33,81)	12758.38		***	5735.61	*	4386.63		2612.30	*
cluster_retail									
Outside	0.00			0.00		0.00		0.00	
CBD	-9633.58		***	-7674.29	**	-15309.72		-5435.07	***
Orchard	28418.69		***	-4287.93		-478001531.44	***	-846.91	
Katong	-12267.97		***	-10628.67	***	-79396.83		-7831.16	***
Geylang Rd.	-18003.41		***	-17892.25	***	34927.78	.	-12534.25	***
Rochor	-6271.72		***	-5432.59	*	-27185.84		-3665.20	**
tenure_type									
Not available	0.00			0.00		0.00		0.00	
99 Years	1361.92		***	2563.35		2185.48		1629.84	
999 Years	10691.09		***	16794.01	***	12281.22		7350.49	***
Freehold	19112.30		***	19222.01	***	14005.14	***	10107.84	***
top_year									
Not available	0.00			0.00		0.00		0.00	
-2013	-3041.71		*	-2660.32	*	-2714.31	*	-1343.43	
2013 - 2020	6731.46		***	6204.05	**	5860.98	.	2298.26	
Adj. R^2/PseudoR^2	0.52			0.61		0.62		0.60	
AIC	69285.42			68704.26		68682.22		68778.08	
Moran's I	0.38		***	0.00		0.00		0.04	***
Lambda / rho				0.50	***	0.48	***	0.45	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.9 Asking (rent) retail model

The results for the asking (rent) retail model can be found in Table 23. The main determinant for this model is the floor area and if the property is located inside a shopping centre. Similar to the previous model, the Orchard cluster shows a higher price compared to the listings outside the investigated areas. The private vacancy rate has a positive effect on the dependent variable. The following variables are not significant: all parking variables, in_orchard, aircon_available and totpax_500.

As in the previous sale model, the condition has a negative effect on the price. The reason for this negative effect is unknown. The SAR_{err} model performs best for this data and specifications.

Table 23: Regression results for the asking (rent) retail model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	6.37			6.84	***	0.89	***	1.18	***
log(floor_area_sqm)	-0.27	-0.41	***	-0.25	***	-0.11	***	-0.08	***
in_shopping_mall:True	0.39	0.22	***	0.36	***	0.15	**	0.04	***
retail_vac_rate_private	0.01	0.11	***	0.01	***	0.00	.	0.00	***
condition:True	-0.10	-0.07	***	-0.07	***	-0.03	***	-0.03	***
log(distance_mrt)	-0.12	-0.17	***	-0.21	***	-0.10	***	-0.02	***
bus_number_unique	0.01	0.14	***	0.01	***	0.00	**	0.00	***
cluster_retail									
Outside	0.00			0.00		0.00		0.00	
CBD	-0.15		***	-0.27	***	-0.17	*	-0.03	**
Orchard	0.19		***	0.54	***	-0.51	*	0.08	***
Katong	-0.19		***	-0.20	*	0.01		-0.06	***
Geylang Rd.	-0.49		***	-0.43	***	0.04		-0.07	***
Rochor	-0.21		***	-0.28	***	0.06		-0.04	***
Adj. R^2/PseudoR^2	0.42			0.66		0.66		0.63	
AIC	8209.79			5353.39		-4466.12		-4032.23	
Moran's I	0.55		***	-0.01		-0.01		0.06	***
Lambda / rho				0.70	***	0.70	***	0.63	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.10 Asking (sale) industrial model

Table 24 shows the results for the asking (sale) industrial model. The main determinants for this model are the floor area and the distance to the MRT station. Further the number of unique bus lines is another important variable. The tenure_type variable shows, that the shorter the leasehold, the lower the price. A leasehold of 999 years has a similar effect as the category freehold. In the spatial models the difference between the category freehold and 999 years is quite big. But the category 999 years is not any more significant in these models, as only few properties are sold with a leasehold of 999 years in the industrial sector. The Pandan cluster shows a positive effect on the price, in the linear model as well as in the spatial models.
The model fit is very high for this model. The pseudo R^2 is 0.87. As the model fit for all spatial models is the same, the SAR_{err} is considered the best one. In this model the SAR_{lag} model does not show any spatial autocorrelation. This is remarkable, since in the other models presented, the spatial autocorrelation could never be completely removed in the SAR_{lag} model.

Table 24: Regression results for the asking (sale) industrial model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	9.73		***	10.30	***	1.02	***	1.13	***
log(floor_area_sqm)	-0.16	-0.27	***	-0.09	***	-0.04	***	-0.02	***
log(gpr_num)	0.23	0.04	***	-0.55	***	-0.77	***	-0.01	
log(ind_vac_rate_private)	0.54	0.18	***	0.26	**	-0.40	***	0.03	*
total_lots	0.00	0.04	***	0.00		0.00		0.00	***
log(distance_mrt)	-0.13	-0.19	***	-0.21	***	-0.20		-0.02	***
bus_number_unique									
[0,9)	0.00			0.00		0.00		0.00	
[9,17)	0.06		***	-0.10	**	-0.06	*	-0.01	**
[17,23)	0.11		***	-0.03		-0.33	***	0.00	
[23,36)	0.23		***	0.05		-0.19	***	0.01	*
cluster_industrial									
Outside	0.00			0.00		0.00		0.00	
Bukit Merah	0.16		***	0.14		-32076.38	***	0.01	
Kallang	0.25		***	0.34	**	0.02		0.02	
Boon Lay	0.12		**	-0.04		0.00		0.03	*
Pandan	0.27		***	0.32	**	15563.88	***	0.03	***
Eunos	0.01		.	0.00		0.23	***	0.00	
Sembawang	-0.28		***	-0.21		-121.19	***	-0.02	.
Kallang Way	-0.11		***	-0.07		-1.73		-0.01	*
Clementi	0.07		***	-0.10	*	5749.48	***	-0.01	.
tenure_type									
Not available	0.00			0.00		0.00		0.00	
30 Years	-0.29		***	-0.33	***	-0.17	***	-0.04	***
60 Years	-0.11		***	0.16	***	0.01		0.00	
99 Years	-0.07			0.19	*	0.04		0.00	
999 Years	0.30		***	0.07		-0.13	*	0.04	**
Freehold	0.36		***	0.66	***	0.19	***	0.05	***
Adj. R^2/PseudoR^2	0.72			0.87		0.87		0.87	
AIC	11.60			-3748.33		-12962.57		-12745.19	
Moran's I	0.62		***	-0.03		-0.03		-0.01	
Lambda / rho				0.80	***	0.78	***	0.76	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.4.11 Asking (rent) industrial model

Table 25 shows the results for the asking (rent) industrial model. The most important variable for this model is the floor area. Interesting is that for this sector, the number of bus lines in a distance of 500 meters works better than all other bus access measurement variables. This is the only model, where this variable works the best. For this model the work place density is significant and shows a positive effect. The variable is however only significant in the linear model, but has a rather big influence, comparable with the access variables. As already seen in other industrial models, the Boon Lay cluster has the lowest prices. In the SAR_{err} model the Boon Lay and the Sembawang cluster have a large negative effect on the price. This can be explained by the location of these clusters.

The spatial autocorrelation can be removed with the SAR_{err} and the SAR_{mix} model. The results for the SAR_{mix} model are harder to interpret. Because of this reason and because of the similar model fit, the SAR_{err} is considered the best model.

Table 25: Regression results for the asking (rent) industrial model

Dependent: log(price_psm)

	Linear OLS			SAR_{err}		SAR_{mix}		SAR_{lag}	
Variable	**Coef**	**StCoef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**	**Coef**	**Sign**
Intercept	4.23		***	4.27	***	0.49	***	0.77	***
log(floor_area_sqm)	-0.14	-0.47	***	-0.16	***	-0.07	***	-0.04	***
work_density_500	0.00	0.11	***	0.00	.	0.00		0.00	***
log(distance_mrt)	-0.06	-0.10	***	-0.05	**	-0.17	***	-0.01	***
num_bus_line_500	0.00	0.12	***	0.00	***	0.00		0.00	*
perc_tertiary_edu									
[0, 0.112)	0.00			0.00		0.00		0.00	
[0.112,0.130)	0.10		***	0.04		-3695.94		0.00	
[0.130,0.162)	0.19		***	0.15	**	-3696.10		0.02	***
[0.162,0.311]	-0.05		*	-0.08		-3695.91		-0.01	**
cluster_industrial									
Outside	0.00			0.00		0.00		0.00	
Bukit Merah	0.03			0.01		2600.85	***	-0.01	
Kallang	0.10		***	0.02		-0.09	*	0.01	
Boon Lay	-0.19		***	-0.31	***	-5884.46	***	-0.06	***
Pandan	-0.18		***	-0.07		-0.01		-0.04	***
Eunos	-0.09		***	-0.05	.	-0.04		-0.02	***
Sembawang	0.26		***	-0.23	***	0.04	.	-0.02	*
Kallang Way	-0.12		***	-0.10	**	6.52		-0.02	***
Clementi	0.04		*	0.07	.	101.88	***	-0.01	*
Adj. R^2/PseudoR^2	0.53			0.75		0.75		0.71	
AIC	258.37			-4090.36		-16860.49		-15955.12	
Moran's I	0.59		***	-0.03		-0.03		0.06	***
Lambda / rho				0.73	***	0.73	***	0.64	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.5 Results of the comparison models

This section shows the results for the comparison models. These are estimated with the same variables over the asking sale and transaction data, so that the results can be compared with each other. Further to the linear models the SAR_{err} models are reported for each data type. Because of the experience from the full models presented in Section 5.4, only the SAR_{err} model results are reported. All models presented in this section have the logarithmic price per square meter as dependent variable.

5.5.1 Office comparison model

The results of the office comparison model are shown in Table 26. The sign for the parameter of the distance to the next MRT station has an unexpected direction in the transaction data. This is caused by a high correlation between the number of bus lines and the distance to the next MRT station. This value should therefore not be compared with the other ones. The values for the tenure are not very similar between the different models. A problem with this variable could be, that the data availability is very different. In the transaction data no office property is sold with a leasehold of 30 or 60 years, but in the asking data such entries can be found. The influence of the number of bus lines is much more higher for the transaction data.

In general the asking model performs better than the transaction model. The spatial autocorrelation is not reduced completely in the transaction SAR_{err} model.

Table 26: Regression results for the comparison model office

Dependent:	**Transaction**					**Asking Sale**				
log(price_psm)	**LM**			**SAR$_{err}$**		**LM**			**SAR$_{err}$**	
	Coef.	StCoef	Sign.	Coef.	Sign.	Coef.	StCoef	Sign.	Coef.	Sign.
Intercept	9.05		***	9.13	***	10.67		***	11.06	***
log(floor_area_sqm)	-0.01	-0.03		-0.10	***	-0.11	-0.17	***	-0.11	***
log(distance_mrt)	0.08	0.19	***	0.12	**	-0.08	-0.12	***	-0.13	**
bus_number_unique	0.01	0.44	***	0.01	***	0.00	0.17	***	0.00	
tenure_type										
Freehold	0.00			0.00		0.00			0.00	
Not available	-			-		-0.22		*	-0.07	
30/60 Years	-			-		-0.87		***	-0.60	***
99 Years	-0.25		***	-0.19	***	-0.16		***	-0.13	*
999 Years	0.00			-0.09		0.12		*	0.29	***
cluster_office										
Outside	0.00			0.00		0.00			0.00	
Katong	0.05			0.25	*	0.17		.	-0.10	
Orchard	0.26		***	0.27	.	0.22		***	0.10	
Eunos	0.02			0.04		0.02			-0.10	
CBD	0.29		***	0.37	***	0.33		***	0.51	***
Victoria St.	0.07		*	0.10		0.44		***	0.35	***
Adj. R^2/PseudoR^2:	0.42			0.74		0.55			0.92	
AIC	385.57			-526.16		1308.13			-1126.61	
Lambda				0.79	***				0.89	***
Moran's I	0.60		***	0.02	*	0.86		***	-0.02	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.5.2 Retail comparison model

Table 27 shows the results for the retail comparison model. This model contains the floor area, two access variables (distance_mrt, bus_number_unique), whether the property is located inside a mall or not, the tenure type and the cluster definition. The access and the floor area are the most important variables for retail, where the parameters for the transaction and asking data are comparable. The influence of the variable in_shopping_mall is very similar for the asking and the transaction data.

A leasehold duration of 999 years can be considered to be worth equal or more than freehold. For the asking data this class (999 years) is not significant, which means, that this class can be considered to be equal to freehold. A leasehold of 99 years is worth less in the asking data compared than in the transaction data.
The values of the cluster/location variable show some differences between the two data types. In the linear models the values are more similar than in the SAR$_{err}$ models. The data availability is

different for the two data types, which makes some of the locations to be not significant.
The model fit of the two models are comparable to each other and the spatial autocorrelation is removed entirely in both SAR_{err} models.

Table 27: Regression results for the comparison model retail

Dependent: log(price_psm)	**Transaction**					**Asking Sale**				
	LM			**SAR_{err}**		**LM**			**SAR_{err}**	
	Coef.	StCoef	Sign.	Coef.	Sign.	Coef.	StCoef	Sign.	Coef.	Sign.
Intercept	11.98			11.43	***	11.52			11.37	***
log(floor_area_sqm)	-0.25	-0.32	***	-0.14	***	-0.19	-0.28	***	-0.14	***
in_shopping_mall	0.54		***	0.71	***	0.53		***	0.77	***
log(distance_mrt)	-0.15	-0.21	***	-0.15	***	-0.21	-0.20	***	-0.19	***
bus_number_unique	0.01	0.30	***	0.01	***	0.01	0.23	***	0.01	***
tenure_type										
Freehold	0.00			0.00		0.00			0.00	
Not available	-			-		-1.02		***	-0.97	***
99 Years	-0.44		***	-0.33	***	-0.64		***	-0.72	***
999 Years	0.11		***	0.08		-0.09		.	-0.01	
cluster_retail										
Outside	0.00			0.00		0.00			0.00	
CBD	-0.35		***	-0.23	**	-0.35		***	-0.34	**
Orchard	0.19		***	-0.25	**	0.01			-0.73	***
Katong	-0.29		***	-0.19	**	-0.20		***	-0.21	*
Geylang Rd.	-0.28		***	-0.01		-0.22		**	-0.17	
Rochor	-0.08		**	0.16	*	-0.02			-0.10	
Adj. R^2/PseudoR^2:	0.49			0.69		0.43			0.66	
AIC	2626.00			1259.56		5367.32			3922.77	
Lambda				0.71	***				0.70	***
Moran's I	0.52		***	-0.05		0.54		***	-0.01	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.5.3 Industrial comparison model

The results of the industrial comparison model are reported in Table 28. The most important variables are the floor area as well as the access variables (distance_mrt, bus_number_unique). The tenure_type variable shows similar results and is also an important variable. The influence of the 999 years class is not significant for the transaction model. In the asking model this class is only significant on a low level, but the parameter value of the SAR_{err} model is surprisingly highly negative. A possible reason for this could be that only a few properties are sold with a 999 year leasehold. A leasehold duration of 60 and 99 years has approximately the same influence on the price. The results for the cluster/location variable are very dissimilar for some classes. Not all classes are significant for this variable. Further the gpr_num variable is included, but the influence on the price is rather small in both models.

The model fits of the transaction and asking data are very similar for the linear as well as for the spatial SAR_{err} models. The spatial autocorrelation is in both linear models equal and is removed by the use of the SAR_{err} models.

Table 28: Regression results for the comparison model industrial

Dependent:	**Transaction**					**Asking Sale**				
log(price_psm)	**LM**			**SAR_{err}**		**LM**			**SAR_{err}**	
	Coef.	StCoef	Sign.	Coef.	Sign.	Coef.	StCoef	Sign.	Coef.	Sign.
Intercept	10.65		***	10.21	***	10.40		***	10.83	***
log(floor_area_sqm)	-0.24	-0.29	***	-0.09	***	-0.17	-0.28	***	-0.09	***
log(distance_mrt)	-0.20	-0.27	***	-0.21	***	-0.16	-0.23	***	-0.24	***
log(bus_number_unique)	0.16	0.25	***	0.10	***	0.11	0.20	***	0.04	**
gpr_num	0.11	0.08	***	0.06		-0.03	-0.02	*	-0.17	***
tenure_type										
Freehold	0.00			0.00		0.00			0.00	
Not available	-			-		-0.39		*	-0.67	***
30 Years	-0.71		***	-0.92	***	-0.62		***	-0.99	***
60 Years	-0.37		***	-0.46	***	-0.48		***	-0.49	***
99 Years	-0.20		***	-0.53	***	-0.41		***	-0.46	***
999 Years	-0.01			-0.25	.	-0.11		*	-0.64	***
cluster_industrial										
Outside	0.00			0.00		0.00			0.00	
Bukit Merah	0.04			0.48	***	0.17		***	0.13	
Kallang	0.07		*	0.50	***	0.31		***	0.36	***
Boon Lay	-0.45		***	-0.49		0.22		***	0.03	
Pandan	-0.09		*	0.25	**	0.12		*	0.21	*
Eunos	-0.01			-0.02		0.04		***	-0.02	
Sembawang	-0.09		**	-0.12		0.02			-0.03	
Kallang Way	-0.09		***	-0.18	***	-0.11		***	-0.08	.
Clementi	0.15		***	0.09	.	0.07		***	-0.09	**
Adj. R^2/PseudoR^2:	0.68			0.88		0.69			0.87	
AIC	3097.97			-5676.83		617.96			-3732.69	
Lambda				0.85	***				0.80	***
Moran's I	0.67	***		0.00		0.67	***		-0.03	

Signif. codes: 0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

5.6 Results of the cluster analysis / models

This model takes the price per square meter as dependent variable and the clusters defined by the usage type as the only independent variable. This shows the differences between the different locations in the city of Singapore. As the cluster variables are often correlated to other variables, this model should show differences between the different regions in the city. This analysis shows the differences between the different areas in the city very well. The calculated cluster variable does not perform best in all models, but this variable is the best for analysing spatial differences in prices.

5.6.1 Office

Table 29 gives the results for the transaction and asking (sale and rent) data. The amount of listings in each cluster is also included in the results table. A map with the cluster areas is shown in Figure 19(a) on page 54.
The declarative power of the models is quite high at approximately 25%. In all models the cluster dummy variable is highly significant. As it is assumed, there are some areas with higher and some with lower prices. The difference between the different models (asking and transaction) is well visible. For some of the clusters the amount of data points is not large enough to make a statement, for example for the Katong cluster of the asking sale model.

Table 29: Regression results of the office cluster model

	Transaction			Asking Sale			Asking Rent		
	% of obs	Coef	Sign.	% of obs	Coef	Sign.	% of obs	Coef	Sign.
Intercept		9.65	***		9.40	***		3.87	***
Not in cluster	29.9%	0.00		22.4%	0.00		30.3%	0.00	
Katong	2.9%	0.20	***	0.9%	0.46	***	0.5%	-0.13	**
Orchard	1.6%	0.36	***	4.1%	0.69	***	7.5%	0.59	***
Eunos	9.2%	0.05		10.8%	0.44	***	0.1%	-0.24	
CBD	41.5%	0.29	***	43.5%	0.73	***	42.2%	0.67	***
Victoria St.	14.9%	-0.17	***	18.3%	0.59	***	19.4%	0.33	***
Adj. R^2		0.25			0.33			0.34	
Total Obs.:	1719			1208			4513		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; Dependent variable: log(price_psm)

5.6.2 Retail

The results of the cluster analysis for the retail market are listed in Table 30. The different clusters are shown in the Figure 19(b) on page 54. The model fit reached with the clusters is much lower compared to the office market. This means, that for retail such locations with higher prices do not exist or only at few spots. For the transaction data only the Orchard cluster shows a highly significant influence. The other investigated areas (except the CBD) are significant in the asking models, but the highest positive influence is found at the Orchard cluster. The CBD does not show a higher or lower price compared to the listings outside the cluster areas. The results of the Katong, Geylang Road and Rochor dummy are different between asking sale and asking rent.

Table 30: Regression results of the retail cluster model

	Transaction			Asking Sale			Asking Rent		
	% of obs	Coef	Sign.	% of obs	Coef	Sign.	% of obs	Coef	Sign.
Intercept		10.28	***		10.33	***		4.72	***
NA	51.5%	0.00		58.2%	0.00		62.8%	0.00	
CBD	16.6%	0.00		12.4%	0.08	.	11.7%	-0.01	
Orchard	2.6%	0.70	***	2.8%	0.95	***	6.1%	0.50	***
Katong	9.9%	-0.08	*	9.3%	0.28	***	3.1%	-0.23	***
Geylang Rd.	2.2%	-0.12	.	2.1%	0.34	***	2.8%	-0.67	***
Rochor	17.2%	-0.01		15.2%	0.12	**	13.5%	-0.16	***
Adj. R^2		0.12			0.09			0.10	
Total Obs.:	2162			1888			3307		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; Dependent variable: log(price_psm)

5.6.3 Industrial

Table 31 shows the results of the location analysis model for the industrial sector. The definition of the different cluster areas can be found in Figure 19(c) on page 54. The model fit shows that also for the industrial sector, the location is not as important as for office properties. This confirms the hypothesis. The results are similar for some locations across the different data types. Some of the cluster locations only have very few listings on it, which signifies that these coefficients are probably not very robust. Generally it can be said, that the Boon Lay cluster is the cheapest investigated location, the Kallang the priciest location.

Table 31: Regression results of the industrial cluster model

	Transaction			Asking Sale			Asking Rent		
	% of obs	Coef	Sign.	% of obs	Coef	Sign.	% of obs	Coef	Sign.
Intercept		8.32	***		8.70	***		3.10	***
Not in cluster	49.3%	0.00		43.8%	0.00		44.9%	0.00	
Bukit Merah	0.8%	0.46	***	2.1%	0.16	***	4.0%	0.42	***
Kallang	3.3%	0.58	***	1.2%	0.46	***	3.2%	0.50	***
Boon Lay	2.6%	-1.07	***	0.1%	-0.73	***	0.4%	-0.20	***
Pandan	2.8%	-0.03		4.3%	-0.07		4.0%	0.05	
Eunos	29.3%	0.14	***	31.2%	0.07	***	23.5%	0.21	***
Sembawang	1.0%	-0.45	***	1.2%	-0.49	***	1.4%	0.00	
Kallang Way	3.1%	0.38	***	6.1%	0.21	***	11.7%	0.16	***
Clementi	7.7%	-0.16	***	10.2%	-0.32	***	6.9%	0.18	***
Adj. R^2		0.13			0.10			0.13	
Total Obs.:	7211			2327			2583		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; Dependent variable: log(price_psm)

5.7 Results of accessibility analysis

Table 32 shows the results for the access/accessibility analysis model. The table contains the coefficients of the different data types for the three variables distance_mrt, bus_number_unique and accessibility_work. The model is specified as a log-log model. Each model is regressed with only one variable, so that the influence of the access and accessibility variables can be compared directly.

The office models show quite some differences in performance. Overall the number of unique bus lines works the best. The distance to the next MRT station and the workplace accessibility do not work with the office transaction data. For the asking data these two variables work similar to the number of bus lines.

The access and accessibility variables do not work very well for the retail models. The standardised coefficient values are similar for the number of bus lines. For the distance to the MRT station and the workplace accessibility some differences can be seen in terms of model fit and the value of the coefficients. Generally the accessibility variable works best for the retail models. If the two access variables are used together in a model, the model performs slightly better than a model with only the accessibility.

Table 32: Regression results of accessibility analysis

Dependent: log(price_psm)

		log(bus_number_unique)				log(distance_mrt)				log(accessibility_work)			
	Model	Inter.	Coef	StCoef	R^2	Inter.	Coef	StCoef	R^2	Inter.	Coef	StCoef	R^2
Office	Trans.	8.95	0.24	0.49	0.24	10.08	-0.05	-0.11	0.01	7.58	0.19	0.19	0.04
	Sale	8.59	0.38	0.48	0.23	11.86	-0.33	-0.48	0.23	0.81	1.80	0.53	0.28
	Rent	3.10	0.34	0.55	0.30	6.03	-0.30	-0.52	0.27	-1.13	1.07	0.42	0.18
Retail	Trans.	9.57	0.24	0.28	0.08	11.47	-0.18	-0.26	0.07	5.83	0.89	0.30	0.09
	Sale	9.40	0.34	0.30	0.09	12.00	-0.25	-0.24	0.06	2.93	1.51	0.38	0.14
	Rent	3.93	0.26	0.29	0.09	6.20	-0.24	-0.34	0.12	1.87	0.58	0.23	0.06
Industrial	Trans.	7.63	0.32	0.52	0.27	11.22	-0.41	-0.54	0.29	2.97	1.14	0.54	0.29
	Sale	8.04	0.27	0.48	0.23	11.28	-0.37	-0.54	0.29	3.88	1.02	0.52	0.27
	Rent	2.88	0.15	0.51	0.26	4.68	-0.21	-0.35	0.12	-0.03	0.67	0.53	0.28

Model Type: log - log; Inter.= Intercept; StCoef = Standardized coefficient

All variables are significant at level 0

The influence of the workplace accessibility is highest for the industrial sector. The influence and the model fit is similar for the industrial models (transaction, asking sale and asking rent). The variables number of bus lines and distance to the next MRT station do not perform better if they are used alone. These two variables perform slightly better than the accessibility if they are used together in a model (same as for the retail models).

6 Discussion

This section provides a discussion and analysis for the models presented in the previous section. Section 6.1 shows the different specifications tested for the use of the dependent variable. Sections 6.2 and 6.3 discuss the results of the two one-variable models. Section 6.4 analyses the results of the comparison models and Section 6.5 those of the full models. Further the results are compared to other study regions presented in the literature review.

6.1 Dependent variable

All presented models are estimated with the price per square meter as dependent variable. The first estimation results were done using the price as dependent variable. This had lead to very high model fits. In these cases the floor area alone already explained about 70 - 80% of the variance in the price. Because of the big influence of the floor area other variables had only a marginal influence on the results. The resulting adjusted R^2 with the price as dependent variable is at about 95 to 98% for all models. In the literature for the office market, some models are found, which use the price as dependent variable. These models also have very high R^2, similar to the first results in this study.
To give more power to other determinants than the floor area, all models are regressed using the price per square meter as dependent variable. The floor area variable is still included in the model, to test if there are some bulk effects for renting properties.

Apart from the usual two dependent variables (price and price per square meter) another approach is introduced, where the tenure is included in the dependent variable. The dependent variable describes the price and the tenure conditions in one mixed variable. This is done by calculating the annual rent, which has to be paid for the property. The mixed dependent variable is calculated by dividing the price per square meter with the remaining years of the lease. Following equation shows the used definition:

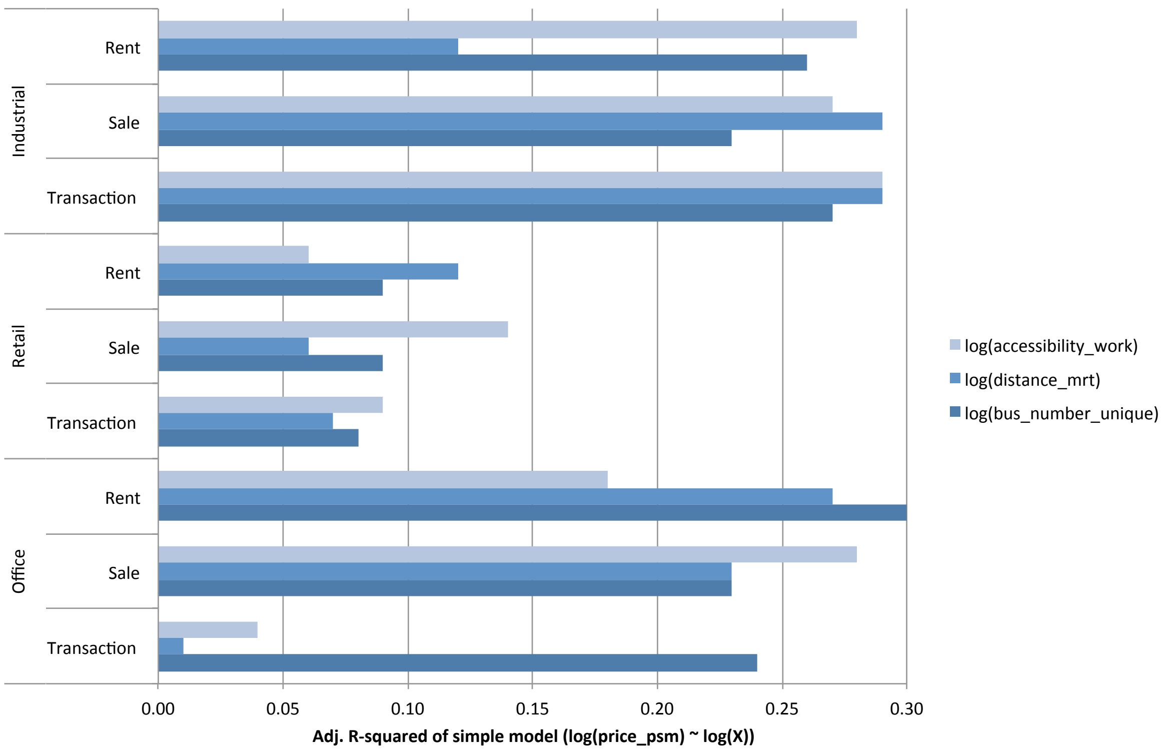
$$Y = \frac{P * r}{1 - (1 + r)^{-a_t}} \tag{12}$$

where Y is the annuity used as dependent variable, r the rate per period, P the present value of the property and a_t number of periods or remaining years of the lease (Finance Formulas, 2013). This variable did not performed better than the classical dependent variables and the interpretation of the results is more difficult by using this mixed variable. To include the tenure conditions as an independent variable works better and gives more reliable results. Results for such a mixed dependent variable are not reported in this thesis.

6.2 Access/Accessibility analysis

As the results showed, the access or accessibility has a large influence on the price of each property. Figure 20 shows the values of the adjusted R^2 for the distance to the next MRT station, the number of unique bus lines and the work accessibility. The analysis is done for the three markets, office, retail and industrial. It can be clearly seen, that overall the access or accessibility is not as important for the retail market as for the other two markets.

Figure 20: Access/Accessibility analysis



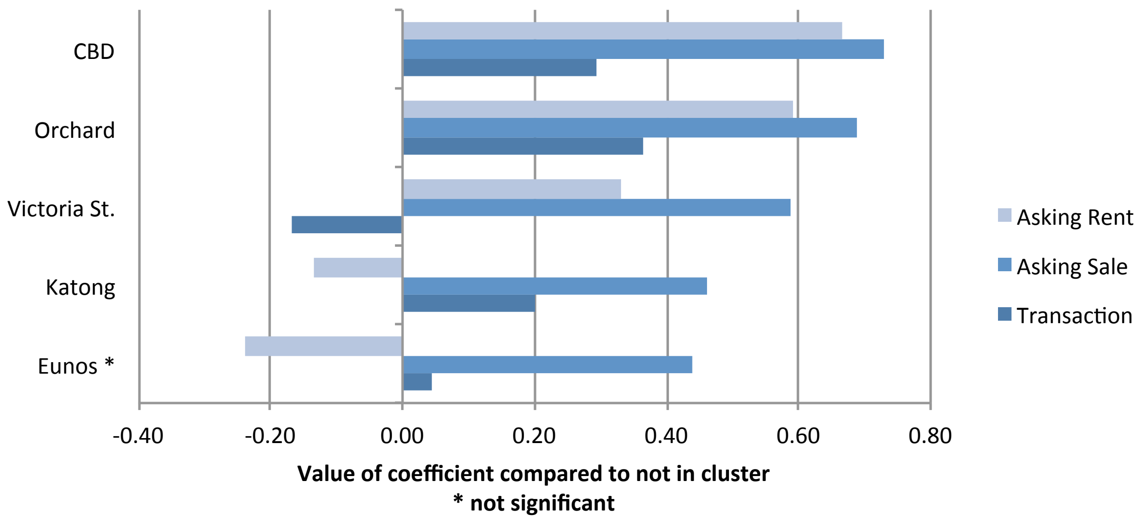
For the transaction office model the accessibility and the distance to the next MRT station does not work at all. A reason for this could be the high concentration of office space in the CBD. The distance to the MRT is equally distributed in the CBD. Also the accessibility in the city centre is overall quite high. The number of bus lines is especially high around Raffles Place. There are also some properties around Raffles Place that are much more pricey than the rest. Because of this fact, it is assumed, that the number of bus lines works best for the transaction office model. Overall the workplace accessibility and the number of bus lines yield to the best results. The effect of the distance to the next MRT station usually does not describes the price variances as well.

6.3 Cluster/Location analysis

6.3.1 Office

For office usages the location itself is assumed to be a very important location choice factor. This means, that some places should be pricier than other ones. Figure 21 shows the results of the cluster analysis for the office market. It can be seen, that for the transaction data the investigated locations do not show an effect as high as for the asking sale and asking rent data. The most important locations are the CBD and Orchard. The region around Victoria Street is also important.
The results for the locations Katong and Eunos have to be interpreted with caution, as either the number of listings in these areas are small or the parameter is not significant. The number of listings in the Orchard cluster is also not large. Because of these limitations, it is assumed, that the CBD and the Victoria Street cluster have the biggest impact on the price.

Figure 21: Cluster/location analysis for the office market

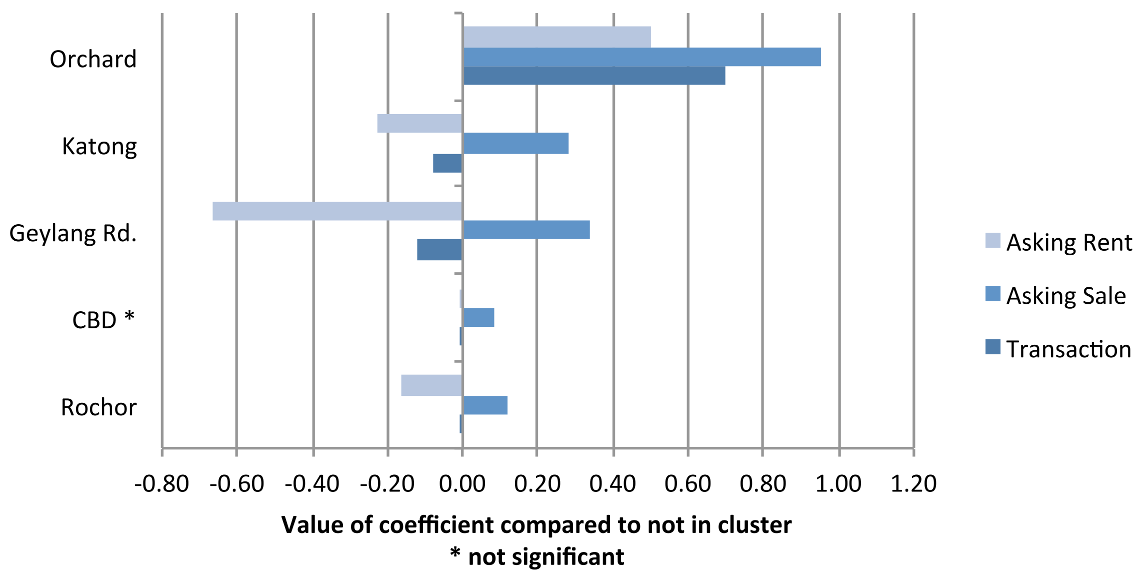


6.3.2 Retail

Figure 22 shows the effects of the different locations for the retail market compared to the listings outside these cluster areas. The results show, that for retail no real premium location exists (except Orchard Road) as it is the case for the office market. The location of Orchard Road

shows a clear overall positive effect on the price. The value for the coefficient varies between 0.5 and 0.95 in the different markets. This means, that a property located in Orchard Road has a 50 to 95% higher price per square meter (depending on the market), what can be considered as a high influence. The highest influence is found in the asking sale market. For the other locations the effect on the price seems to be more or less randomly distributed, as the parameters do not have the same direction of sign in the different data types.
Considering that Orchard Road is the main shopping area in Singapore, these results are reasonable. It is also probable, that another retail location like the Orchard Road with a clear positive or negative effect on the price does not exists in Singapore.

Figure 22: Cluster/location analysis for the retail market

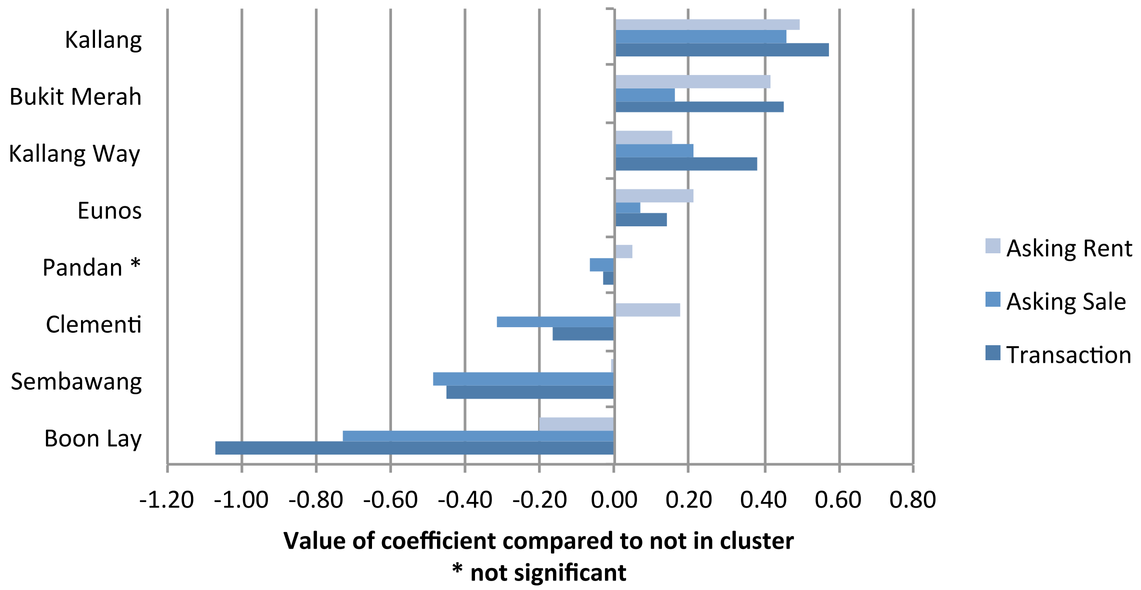


6.3.3 Industrial

The location analysis results for the industrial market are again more distinct as the results of the retail market. Figure 23 shows the results for the industrial market. Not all analysed locations have a significant and clear effect on the price. The locations Pandan, Eunos and Clementi show only little or nearly no effect on the price. The Pandan cluster is not significant.
The priciest location is at Kallang, where the coefficient is at about 0.5 for all data types. On the other side, the Boon Lay location is the cheapest one, but there the variation across the different data types is very high. For the transaction data the effect on the price is over -1, but for the asking rent data the value is just -0.2. A reason for this could be, that most properties in this

region of Jurong West are never listed on the commercial platform and are sold by the Jurong Town Corporation directly to interested parties.

Figure 23: Cluster/location analysis for the industrial market



6.4 Analysis comparison models

6.4.1 Cross-sectional comparison

Table 33 shows the results of the major determinants for the office, retail and industrial markets. The coefficients shown are from the linear comparison models presented in Section 5.5. The comparison of the coefficients are shown in the next subsections.

Between the different market similar preferences can be observed. The main determinants show the same sign, but the importance of the different variables is different. The biggest differences can be found in the tenure_type variable. There each market has its own charakteristic how important each category is compared to freehold properties.

Table 33: Comparison of the linear model results across the office, retail and industrial market

	Office				Retail				Industrial			
	Transaction		Asking Sale		Transaction		Asking Sale		Transaction		Asking Sale	
Variable	Coef.	St.Co.	Coef.	St.Co.	Coef.	St.Co.	Coef.	St.Co.	Coef.	St.Co.	Coef.	St.Co.
Intercept	9.05		10.67		11.98		11.52		10.65		10.40	
log(floor_area_sqm)	-0.01*	-0.03	-0.11	-0.17	-0.25	-0.32	-0.19	-0.28	-0.24	-0.29	-0.17	-0.28
log(distance_mrt)	-	-	-0.08	-0.12	-0.15	-0.21	-0.21	-0.20	-0.20	-0.27	-0.16	-0.23
bus_number_unique	0.01	0.44	0.00	0.17	0.01	0.30	0.01	0.23	0.16	0.25	0.11	0.20
tenure_type												
Freehold	0.00		0.00		0.00		0.00		0.00		0.00	
Not available	-		-0.22		-		-1.02		-		-0.39	
30 Years	-		-		-		-		-0.71		-0.62	
60 Years	-		-		-		-		-0.37		-0.48	
99 Years	-0.25		-0.16		-0.44		-0.64		-0.20		-0.41	
999 Years	0.00 *		0.12		0.11		-0.09	*	-0.01 *		-0.11	

* not significant; All the other variables are significant at a level 0.05; St.Co.: standardized coefficient

6.4.2 Floor area and access comparison

Figure 24 compares the influence of the number of unique bus lines, the distance to the next MRT station and the floor area on the price. These results are for the linear OLS models.

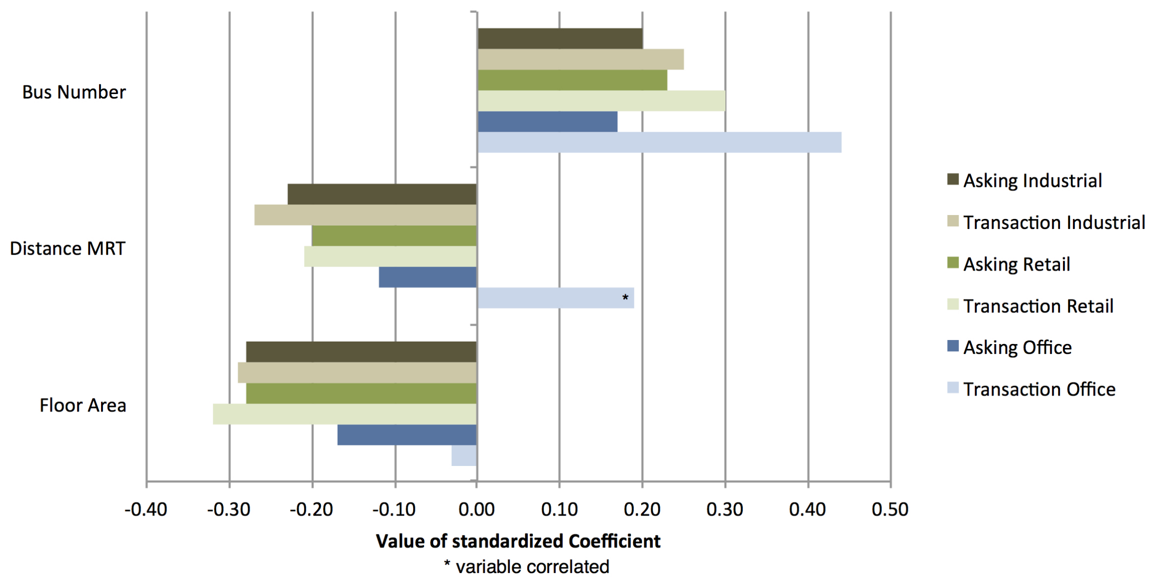
The floor area of the property has a negative effect on the price per square meter. The bigger the bought property, the cheaper the price per square meter. The more floor space is bought or rented, the higher are the bulk effects on the price per square meter. The effect is very similar for the retail and industrial market. If the floor area of a retail or industrial property is increased by one standard deviation, then the price per square meter will decrease approximately 0.3 times of the price standard deviation. For office properties this variable is not that important. Especially for the transaction office data, the floor area has only a low influence. A reason for this behaviour is not found.

The distance to the next MRT station also has a negative influence on the price. This means, if the distance is increased, the price will decrease. The values of the parameters are again similar for the retail and industrial properties. The office properties behave little different. The value of the transaction office is correlated to the number of bus lines and should because of these correlations be interpreted with caution. The overall influence of the distance to the next MRT station seems to be little lower than the floor area.

The number of unique bus numbers at the next five bus stations has a positive effect on the price, as expected. The value for the transaction data is slightly higher than for the asking data. The highest influence is found for the retail and industrial models. The value of the transaction office has to be interpreted with caution because of the correlation.

Overall it can be said, that all three variables are important. The floor area is the most important one, followed by the number of bus lines. These determinants do not work very well for the office market. The office properties are probably not as dependent from the public transport access, because the majority of the office properties have very good and similar access to public transport. This good access is due to the centrality of the office properties.

Figure 24: Influence analysis for the floor area and access variables



6.4.3 Influence of the leasehold duration/freehold

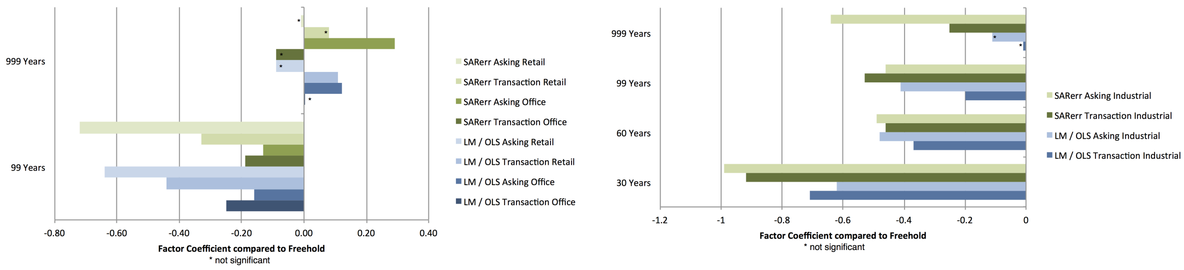
The different tenure conditions have a big influence on the price of the properties. Figure 25 compares the influence on the price per square meter of different leasehold durations to freehold properties.

For the office and retail market (Figure 25(a)) the two categories 99 and 999 years are compared to freehold properties. It can be seen, that a leasehold of 99 years has a negative influence on the price. In the spatial models, the influence is bigger than in the linear ones. There are also differences between the asking and transaction data. The biggest difference can be found in the retail market. A leasehold of 999 years does not show a distinctive positive or negative influence. As the influence in some models is very small, it is considered overall that the 999 year leasehold has nearly no effect on the price, compared to freehold properties.

Figure 25(b) shows the effect of the different leasehold categories compared to freehold for the industrial market. Industrial properties are sold more often with a leasehold duration of 30 and 60 years, than with a leasehold of 99 or even 999 years. The duration of 30 years shows a huge negative effect on the price compared to freehold. The asking and transaction data have similar results. The categories of 60 and 99 years have especially in the spatial SAR_{err} models a similar effect. They reduces the price by up to 50%. The 999 year leasehold has very different results. In the linear models, the effects are negligible, but in the SAR_{err} model the effect is for the asking data comparable to the results of the 99 year category.

Figure 25: Influence analysis for the tenure on the price

(a) Office/Retail (b) Industrial



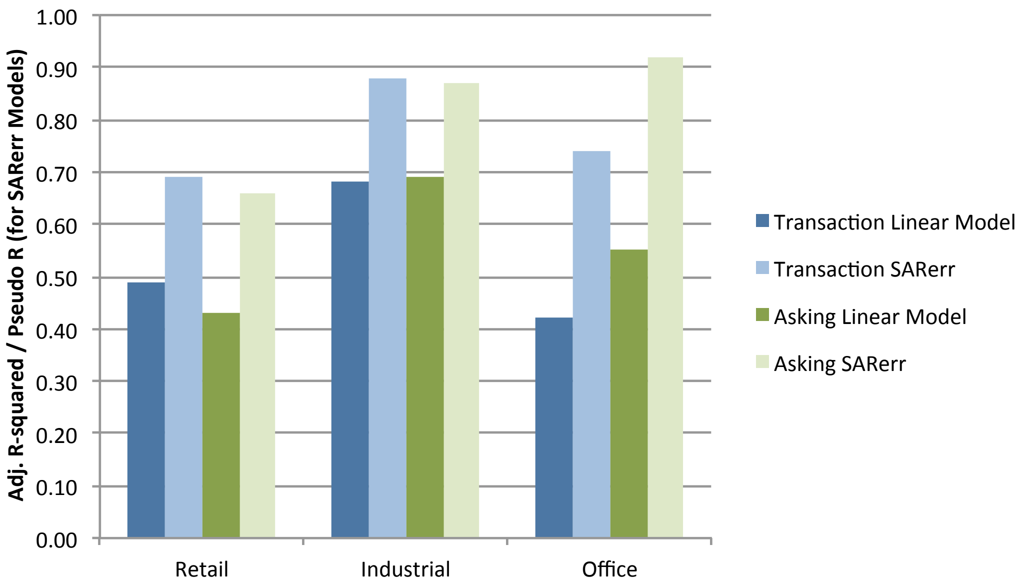
6.4.4 Model fit of the comparison models

Figure 26 shows the model fit of the different comparison models. The industrial models have overall the highest model fit (adj. R^2) of approximately 0.7 to 0.9. The retail and office models have similar model fits at around 0.5 to 0.7. The asking data models are performing similar compared to the transaction models.

The fit of all spatial SAR_{err} models are higher, because the spatial autocorrelation is removed. The highest increase in model fit between a linear and a spatial model is found in the office category. An explanation for this high increase could be, that the spatial autocorrelation is very high (0.6 for the transaction and 0.86 for the asking data) in the linear model. Therefore the increase in model fit is bigger if the autocorrelation can be removed completely. The office properties are highly concentrated in the city centre. This could be a reason for a high spatial autocorrelation of the listings.

The fit for the office models are comparable with the results of the literature review. The fit of the linear models are a bit lower than seen in the literature review. For the retail market, the model fit of both (linear and spatial) models are higher than reported in previous studies. This difference could originate from the fact, that for the Singapore case good data is available, especially the tenure conditions have a big influence on the price for the retail market. The fit for the industrial market is comparable to other studies, however the model fit is slightly higher for the models presented in the literature review.

Figure 26: Model fit for the comparison models



6.5 Analysis of the full models

The major determinants of the models are already discussed in the previous section (6.4). This section shows some additional determinants that could not be included in the comparison models.

The model fit is generally higher for the full models compared to the comparison models discussed in the previous section. A graphic with the results can be found in Appendix B.1 in Figure 29.

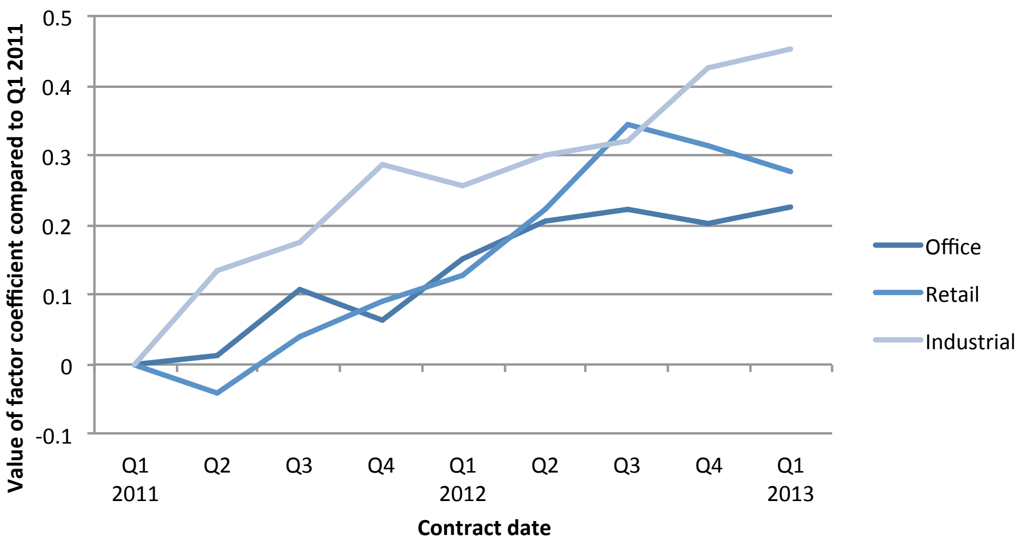
As seen in the comparison models, the model fit of the retail market is the smallest. The best performing models are the spatial models for the industrial sector.

6.5.1 Temporal effects on the price

For the transaction data the contract date is available. The effects of the contract date are shown in Figure 27. The price is increasing approximately linear since the first quarter of 2011. The highest increase in the price is visible for the industrial market. The price of the office and

retail market is evolving similar. The overall influence can be considered as quite large. For the industrial market the price in the first quarter of 2013 is nearly 45% higher than two years ago in the first quarter of 2011.
The results of the models seem to be too high, compared to the findings in the reported price analysis in Section 2.2.7 on page 17.

Figure 27: Influence of the contract date for the transaction data



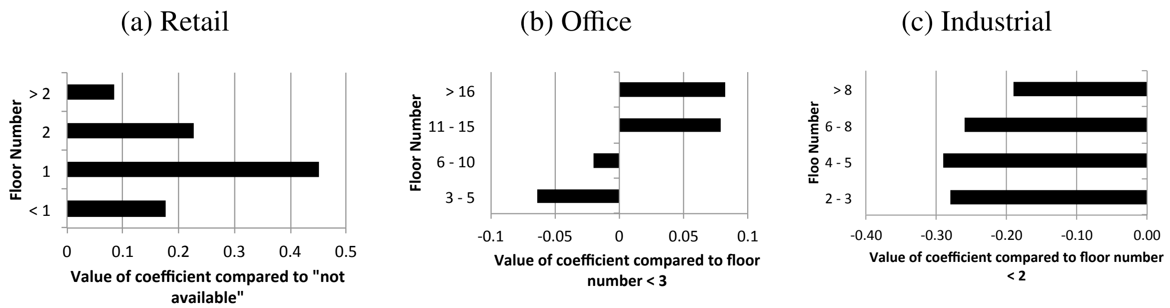
The Q2/Q3 2011 of the retail market and Q2 2011 of the office market are not significant

The TOP year (Year of temporary occupation period) is available for the asking office and retail categories. This variable describes when the first owner can move into the unit. As this attribute is only available for about 10% of the office sale entries and about 20% of the retail sale entries, the influence is measured by a factor variable. If the property's TOP year is situated in the future, a positive effect on the price is measured. For the office sale category, the price is increased by approximately 10% and for retail the increase in the price per square meter is nearly 6750 S$. The different units of the parameters are because of the different model specifications (the retail sale model is not a logarithmic model). On the other side, if the TOP year is in the past, the effect is negative. For office the influence on the price is -11% and for retail -3000 S$ per square meter.
These results show, that the building age has an influence on the price, as it is reported in the literature reviewed.

6.5.2 Floor number

The floor number is especially for the retail market an important determinant. The results shown in Figure 28 confirm this hypothesis. The results shown are only for the transaction market (strata), as for the asking data no floor number is available.

Figure 28: Effect of the floor number on the price



For retail the priciest floor is the ground floor, followed by the second floor and the basement. These results show the importance of the micro-location of a shop. Therefore it is very important for a shop or restaurant on which floor and probably also where on this floor the property is located. What could not be tested is what effect a floor which is directly connected to a MRT station has on the price.
For the office market the influence is rather small. There is a tendency visible, that a higher floor number increases the price. This can be explained with the fact, that for large high rise buildings the construction costs are higher and therefore the price is by tendency higher. It is also possible, that a small prestige effect exists for the high floors.
The influence of the floor in the industrial sector is quite different. The priciest floor is the ground floor. Industrial properties need usually a good access for the delivery of goods. This could be an explanation why the ground floor is the priciest. Figure 28(c) also shows that the higher the floor is, the less negative the effect on the price is. Here the same prestige influence could take effect as described for the office market, just much weaker.

6.5.3 Office: literature comparison

The vacancy rate in the transaction model has a positive effect. Farooq *et al.* (2010) and Ozus (2009) report the opposite effect for the vacancy rate in the neighbourhood of the property. The vacancy rate of the building has a positive effect on the price as reported by Ozus (2009). In this study the public vacancy rate is significant, the private vacancy rate is not. A reason for the

different behaviour could be that the government in Singapore has a bigger influence, because of 80% of the land is owned by the state.
The study of Jennen and Brounen (2009) show a positive effect of the percentage of the population with a tertiary education degree. Same is the case for the asking sale model and partially for the asking rent model. Further state Jennen and Brounen (2009) that a higher contract/building size ratio increases the price, because of possible privileges introduced by the buyer (reception area, putting the corporate identity on building). These results cannot be confirmed. The parameter in the asking rent model shows different results.

The results of Tu *et al.* (2004) are comparable to the results in this study. They report that the floor level of the office units has a positive effect on the price. The variable is included as numeric and not as a factor variable as done in this study. In this study the higher floors generate a higher rent than the lower ones, but the effect is small.
Tu *et al.* (2004) defined the tenure variable as a simple dummy variable indicating leasehold or not (1/0). It is assumed if the property is not leasehold, that it has to be freehold. The reported effect of the leasehold dummy variable has a negative effect, similar as in this study. The strength of the effect cannot be directly compared, as the two studies use different model specifications. The floor area has a positive effect in the study of Tu *et al.* (2004). This is because of the dependent variable. The study uses the price as dependent variable. In this study the price per square meter is used.

6.5.4 Retail: literature comparison

Liang and Wilhelmsson (2011) report a positive effect on the rent if the property is located in a shopping centre. The same effect is found in this study in all retail models (except the transaction land model). Further Liang and Wilhelmsson (2011) report a positive effect for a location dummy describing the main shopping area in the city of Shanghai. Such effects are comparable to the results found for the cluster/location variable in this study, where Orchard Road has a positive effect on the price.
Nase (2013) reports a positive effect of the vacancy rate on the price, but the effect is unexpected. In this study the private and public vacancy rate are used. The private vacancy rate in the asking rent model also shows a positive effect. The effect of the public vacancy rate in the asking sale model and in the transaction strata model is negative. Reasons for the unexpected positive effect of the vacancy rate in the asking rent model are not known.

For the retail models the location/cluster was assumed to be more important. The results show, that only in Orchard Road the prices are significantly higher than in the other regions. It also can be concluded that the location is not one of the major determinants for the retail market.

The prices around and inside big and new shopping malls in the outer districts are similar as compared to the city centre. Because of this not as much clusters (with higher or lower prices) exist for the retail market as for the office and industrial market.

6.5.5 Industrial: literature comparison

Dunse and Jones (2005) conclude that for the city of Glasgow the prices are higher closer to the CBD. They measure this influence with the distance to the CBD. In this study similar results could be found, but the effect is not measured by the distance to the CBD, however by the location dummy variable. This variable shows, that the locations closer to the CBD (Kallang) are pricier than the locations further away (like Boon Lay and Sembawang).
Dunse and Jones (2005) and Dunse *et al.* (2000) report, that properties located close to highway junctions are pricier. In the study of Dunse and Jones (2005) the variable describing the proximity to the highway is the most important one. In this study the distance to the next highway junction is not significant in any industrial model. This is surprising according to the results seen in the two studies of Dunse and Jones (2005) and Dunse *et al.* (2000). A reason for this behaviour could be that in Singapore each location is close to a highway. There is no region with a bad connection to highways. In other study regions this is probably the case.

7 Conclusions and outlook

7.1 Conclusions

Overall the industrial models worked best with a R^2 of 0.7 (OLS models) to 0.9 (spatial models). The transaction models have higher model fits than the asking models. The model fit of the office models varies much more, namely between 0.55 and 0.9. But the office models regressed with asking data worked better. The model fit of the retail sector varies between 0.5 and 0.8, where the transaction models works tendentiously better.
The results show that the SAR_{err} model specification works best. The SAR_{mix} models have slightly better fit than the SAR_{err} models, but the interpretation is more difficult for the SAR_{mix} models. The coefficient of the SAR_{mix} determinants are sometimes unexpected. Because of this reason the SAR_{err} models are preferred. The SAR_{lag} model specification does not work for the case of Singapore. Through the use of the SAR_{lag} model the spatial autocorrelation can not be removed.
The k-nearest neighbours weighted by distance approach for calculating the spatial neighbours matrix produced the best results. The count of the neighbours to include in the method has to be optimized for every model. In the models presented 6 to 16 neighbours are used.

The results of the floor area variable confirm the hypothesis. The high importance for some models is however surprising.

The access and accessibility variables for public transport are important for every market. The smallest influence on the price can be found at the retail market. The highest influence is found at the office and industrial market. Best are the variables describing the distance to the next MRT station and the count of unique bus lines at the next five bus stations working together. The working accessibility of the public transport achieve similar results. The performance of the models with the accessibility are slightly lower and other variables would have needed to be excluded due to more correlations. Because of this reason all models are regressed with the two access variables.

The tenure conditions of the property show to be a major determinant. Generally a leasehold of 999 years is assumed to generate similar prices as a property with freehold conditions. A leasehold of 99 years shows a strong negative influence on the price. The strength of the influence is dependent on the market. In the industrial market most properties are sold with a leasehold of 60 years. The results show that the effect of a 60 year leasehold can be compared to 99 years. The high importance of the determinant is surprising, especially the fact, that properties with a 999 year leasehold are rated similar as freehold properties.

The tenure conditions have the highest influence on the retail market. In the industrial market the influence is especially high for the short leaseholds of 30 years. The data quality of the tenure conditions is unique for Singapore and is usually not available in other regions. Therefore other studies in the literature review usually do not use the tenure in the models.

The literature review showed, that the a variable describing the proximity to the CBD is very important, especially for the office market. Such a variable could not be integrated in this study, because a lot of variables are correlated to a measure describing the distance or the belonging to the CBD. For example is the number of bus lines at the next five bus station correlated with the distance to the CBD, as in the CBD more bus lines are available than at other locations. The Singapore city centre, with a large number of offices, MRT and bus lines promotes the correlation of a lot of variables, because nearly every measure introduced is better or higher in the city centre. Therefore CBD variables are not included into a model. The influence of the CBD is split up into multiple variables, which are describing the prices better than just a simple distance to CBD variable, as seen in a lot of past studies.

Parking measures are not working very well. The biggest problem with the parking measures is the data availability. All private car and truck parking are not available and could not be included in the analysis. That is probably the main reason why the parking measures are not working. A second reason is that in Singapore parking is not a big issue, as the car licences are limited by the COE (Certificate of Entitlement). Because of these restrictions to car ownership, enough parking is available at the majority of the buildings. If someone is using the car it is more a question of his willingness to pay the COE. This makes it very hard to include variables describing the accessibility for car users.

The contract date of the listings has a significant influence on the price. The prices are rising since the first quarter of 2011. A certain amount of variation over time is normal, but the influence in the regressed models is too high compared to the stock data, which is available for Singapore and analysed in Section 2.2.7.

The neighbourhood variables (average land use, percentage of tertiary education, working density, etc.) did not yield results as expected. The working and population density are nearly in all models correlated to other major determinants and can because of this not be included in the models. The average land use, the average income and the percentage of tertiary education in the neighbourhood work only in a few models. But in these models the influence is small or negligible. Probable is that in other regions the neighbourhoods at different locations differ more than in Singapore.

The vacancy rate worked differently in the models and the results are sometimes not explicable. This fact corresponds to the findings of the literature review.

7.1.1 Office

The main determinants for the office market are the access, the location and the tenure conditions. In the office some car access variables (distance to the next highway junction and the number of car lots in the neighbourhood) are significant, but the influence is small.

Clusters and locations are especially quite important for the office market. In Singapore the highest prices are paid in the CBD. Other locations with generally higher prices are the areas around Orchard Road and Victoria Street.
The floor area has a negative influence on the price, as expected. The influence is not as big as for the retail and industrial market. The strength is similar in all the office models, but is largest in the renting model.
The floor number plays only a subsidiary role for the office market. The results show that higher floors are pricier than lower ones, but the differences are small and are not significant for all tested categories.

7.1.2 Retail

The floor number is an important determinant for the retail sector. The results show that the priciest floor is the ground floor. The basement and the second floor are rated similar. The floor number has the highest effect compared to the office and industrial sector. These results show, that the location inside a building or shopping mall is very important for retail properties.

The location/cluster variables do not work very well for the retail market. It can be seen that the price in the major shopping street of Singapore (Orchard Road) is higher than in other locations. But another location with significant higher or lower prices could not be identified. A reason for this could be, that retail properties are distributed uniformly over the whole island.

7.1.3 Industrial

The most important determinant is the access, followed by the tenure conditions and the floor area of the unit. The distance to the next truck parking and the GPR number (ground plot ratio) are in some models significant. Their influence is however small.

The location of the industrial property is an important determinant. In Singapore different locations with higher and lower prices exists. The main industrial estates (like Jurong West and Sembawang) are the cheapest. The closer the locations are to the city centre the pricier the properties. These effects are shown by the cluster variables and not with a variable describing the distance from the CBD.

The road and parking access are not significant. The literature shows that the distance to highways or airports are important. In Singapore this is not the case. Probably the access of a property to the highways in Singapore is too similar across the whole island, which could explain why there is no significant influence.

7.2 Outlook

Most car access variables did not worked as expected. For some models a small effect is found. A similar approach as used to calculate the public transport accessibilities could be used to calculate a roadside accessibility. The use of both accessibilities should increase the explanatory power, as the public transport accessibility worked quite well in the models.
The literature review showed that the building age is a major price determinant, especially for retail and office usages. Including such information in the regression could improve the explanation power of the models.

Office: For the office sector a subdivision of the market could be useful. It should be tested, if the property prices are determined by other attributes in the CBD. Since in this area a lot of properties are available and the determinants differ only very little, probably other determinants are important. A subdivision by industry type (Finance etc.) could probably also show different results.

Retail: The retail sector is directly dependent on the amount of customers. To include this better in the models, the passenger flows and the micro-locations of the property would probably generate better results. The micro-location describes where exactly the shop is located. In a shopping centre it is assumed, that the locations close to the entrance or to escalators are pricier than other ones. Similar the locations at a street corner could generate higher prices. Furthermore the visibility could be included in the model. The better the visibility and the design of the buildings façade, the higher the price should be. In Orchard Road the newer shopping malls have very costly façade designs to receive the attention of the pedestrians and drivers.

Industrial: The Master Plan includes “Business Parks” in the industrial land use. As such business parks are not like the classic industrial estates, which need a lot of land, a separate market should be used for them. Another possibility is to include them into the office market and test the influences there. A reliable distinction of the normal industrial park and the “Business Parks” is not possible.

The results show, that the SAR_{err} models generate the best results. This corresponds to the results of Lehner (2011) for the residential market in Singapore. To describe the neighbours of a listing, the k-nearest neighbours algorithm works best. The neighbours should be weighted by distance to apply different weights in the neighbours list. A possible improvement could be to weigh the neighbours not by the straight line distance as done in this thesis, but by the travel time, similar as done in the accessibility calculations. The straight line weighting method does not account for different spatial structures. Examples are obstacles (like a lake) which are not considered in the calculation.

8 References

Allproperty Media Pte Ltd (2013) Sinpagore's Commercial Property Expert, accessed on 12-05-2013, `http://www.commercialguru.com.sg`.

Anselin, L. (1988) *Spatial econometrics: methods and models*, Kluwer Academic Publishers, Dordrecht.

Anselin, L. (2002) Under the hood issues in the specification and interpretation of spatial regression models, *Agricultural economics*, **27** (3) 247–267.

Anselin, L. (2003) An Introduction to Spatial Regression Analysis in R, *Technical Report*, University of Illinois, Urbana-Champaign.

Balamurugan, A. (2004) Jurong Town Corporation, accessed on 11-03-2013, `http://infopedia.nl.sg/articles/SIP_553_2004-12-31.html`.

Benjamin, J. D., G. W. Boyle and C. F. Sirmans (1990) Retail Leasing: The Determinants of Shopping Center Rents, *Real Estate Economics*, **18** (3) 302–312.

Bollinger, C., K. Ihlanfeldt and D. Bowes (1998) Spatial Variation in Office Rents within the Atlanta Region, *Urban Studies*, **35** (7) 1097–1118.

Breusch, T. and A. Pagan (1979) A simple test for heteroscedasticity and random coefficient variation, *Econometrica: Journal of the Econometric Society*, **47** (5) 1287–1294.

Burnham, K. P. and D. R. Anderson (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection, *Sociological Methods & Research*, **33** (2) 261–304.

Buttimer, R., R. Rutherford and R. President (1997) Industrial warehouse rent determinants in the Dallas/Fort Worth area, *Journal of Real Estate Research*, **13** (1) 47–54.

Chalermpong, S. and K. Wattana (2009) Rent capitalization of access to rail transit stations: Spatial hedonic models of office rent in Bangkok, *Journal of the Eastern Asia Society for Transportation Studies*, **8**.

Department of statistics Singapore (2013) Statistics Singapore, accessed on 13-05-2013, `http://www.singstat.gov.sg/`.

Downs, D. and B. Slade (1999) Characteristics of a full-disclosure, transaction-based index of commercial real estate, *Journal of Real Estate Portfolio Management*, **5** (1) 95–104.

Dunse, N. and C. Jones (1998) A hedonic price model of office rents, *Journal of Property Valuation and Investment*, **16** (3) 297–312.

Dunse, N. and C. Jones (2002) The existence of office submarkets in cities, *Journal of Property Research*, **19** (2) 159–182.

Dunse, N. and C. Jones (2005) The spatial pattern of industrial rents and the role of distance, *Journal of Property Investment & Finance*, **23** (4) 329–341.

Dunse, N., C. Jones, J. Brown and W. Fraser (2000) The Determination of Rents of Industrial Property in a Sub-regional Context, paper presented at the *RICS Cutting Edge Research Conference*, 11–28.

Eng, T. (1986) New towns planning and development in Singapore, *Third World Planning Review*, **8** (3).

Enström, R. and O. Netzell (2007) Can Space Syntax Help Us in Understanding the Intraurban Office Rent Pattern? Accessibility and Rents in Downtown Stockholm, *The Journal of Real Estate Finance and Economics*, **36** (3) 289–305.

Esri Inc. (2013) Light Gray Canvas, accessed on 02-06-2013, `http://www.arcgis.com/home/item.html?id=8b3d38c0819547faa83f7b7aca80bd76`.

Farooq, B., E. J. Miller and M. Haider (2010) Hedonic Analysis of Office Space Rent, *Transportation Research Record: Journal of the Transportation Research Board*, **2174**, 118–127.

Fehribach, F., R. Rutherford and M. Eakin (1993) An analysis of the determinants of industrial property valuation, *Journal of Real Estate Research*, **8** (3) 365–376.

Finance Formulas (2013) Annuity Payment (PV), accessed on 31-05-2013, `http://www.financeformulas.net/Annuity_Payment_Formula.html`.

Fortin, M. and M. Dale (2005) *Spatial analysis: a guide for ecologists*, Cambridge University Press.

Fotheringham, A., C. Brunsdon and M. Charlton (2002) *Geographically weighted regression: the analysis of spatially varying relationships*, Wiley.

François, R., T. Marius and M. Laurent (2005) Spatial versus non-spatial Determinants of shopping center rents: Modelling location and neighborhood-related factors, *Journal of Real Estate Research*, **27** (3) 293–319.

Future Cities Laboratory (2013) Future Cities Laboratory (FCL), accessed on 13-05-2013, `www.futurecities.ethz.ch`.

Global Property Guide (2012) Singapore's rental yields are worryingly low, `http://www.globalpropertyguide.com/Asia/Singapore/Rental-Yields`.

Gorr, W. and A. Olligschlaeger (1994) Weighted Spatial Adaptive Filtering: Monte Carlo Studies and Application to Illicit Drug Market Modeling, *Geographical Analysis*, **26** (1) 67–87.

Haase, R. (2011) Ertragspotenziale - Hedonische Mietpreismodellierungen am Beispiel von Büroimmobilien, Dissertation, ETH Zürich.

Hackney, J., B. Vitins and B. R. Bodenmann (2013) Market Clearing Models in FaLC, paper presented at the *Swiss Transport Research Conference (STRC)*, Monte Verità, Ascona.

Ho, K. H. D. (2007) Modeling the structure of CV formation and expectations: The commercial retail real estate sector, *Journal of Property Investment & Finance*, **25** (2) 179–206.

Housing Development Board (2013) About Us: Our Role, accessed on 13-03-2013, `http://www.hdb.gov.sg/fi10/fi10320p.nsf/w/AboutUsOurRole?OpenDocument`.

Hui, E., C. Yiu and Y. Yau (2007) Retail properties in Hong Kong: a rental analysis, *Journal of Property Investment & Finance*, **25** (2) 136–146.

Hui, E. C. M., C. Liang, W.-C. Ip and D. K. H. Ho (2013) Examining structural changes in Asian offices market, *Land Use Policy*, **32**, 375–380.

Jackson, T. O. (2002) Environmental contamination and industrial real estate prices, *Journal of Real Estate Research*, **23** (1/2) 179–199.

Jennen, M. G. and D. Brounen (2009) The Effect of Clustering on Office Rents: Evidence from the Amsterdam Market, *Real Estate Economics*, **37** (2) 185–208.

JTC Corporation (2012) About JTC, Our History, accessed on 12-03-2013, `http://www.jtc.gov.sg/About-JTC/Pages/Our-History.aspx`.

Kim, J. (2007) Discriminant impact of transit station location on office rent and land value in Seoul: An application of spatial econometrics, *Journal of Transport Economics and Policy*, **41** (2) 219–245.

Kissling, W. and G. Carl (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models, *Global Ecology and Biogeography*, **17** (1) 59–71.

Lancaster, K. J. (1966) A New Approach to Consumer Theory, *The journal of political economy*, **74** (2) 132–157.

Lehner, M. (2011) Modelling housing prices in Singapore applying spatial hedonic regression, Master thesis, ETH Zürich.

Liang, J. and M. Wilhelmsson (2011) The value of retail rents with regression models: a case study of Shanghai, *Journal of Property Investment & Finance*, **29** (6) 630–643.

Löchl, M. (2010) Application of spatial analysis methods for understanding geographic variation of prices, demand and market success, Doctor thesis, ETH Zurich.

MATSim (2013) Agent Based Transport Simulation: MATSim, `www.matsim.org`.

McKelvey, R. D. and W. Zavoina (1975) A statistical model for the analysis of ordinal level dependent variables, *Journal of Mathematical Sociology*, **4** (1) 103–120.

Nagai, K., Y. Kondo and M. Ohta (2000) An Hedonic Analysis of the Rental Office Market in the Tokyo Central Business District: 1985-1994 Fiscal Years, *The Japanese Economic Review*, **51** (1) 130–154.

Nappi-Choulet, I., I. Maleyre and T. Maury (2007) A Hedonic Model of Office Prices in Paris and its Immediate Suburbs, *Journal of Property Research*, **24** (3) 241–263.

Nase, I. (2013) Hedonic modelling of high street retail properties: a quality design perspective, *Journal of Property Investment & Finance*, **31** (2) 160–178.

Nitsch, H. (2006) Pricing Location: A Case Study of the Munich Office Market, *Journal of Property Research*, **23** (2) 93–107.

Ortúzar, J. d. D. and L. Willumsen (2011) *Modelling transport*, 4 edn., Wiley, Chirenster.

Ozus, E. (2009) Determinants of Office Rents in the Istanbul Metropolitan Area, *European Planning Studies*, **17** (4) 621–633.

Phang, S. (2000) How Singapore regulates urban transportation and land use, *Local dynamics in an era of globalization: 21st Century Catalysts for Development*, 159–1.

R Development Core Team (2013) The R Project for Statistical Computing, `http://www.r-project.org`.

Rebelo, E. M. (2011) Urban planning in office markets: A methodological approach, *Land Use Policy*, **28** (1) 83–95.

Rönz, B. and E. Förster (1992) Heteroskedastizität, in *Regressions- und Korrelationsanalyse*, 235–239, Gabler Verlag, Wiesbaden.

Rosen, K. T. (1984) Toward a Model of the Office Building Sector, *Real Estate Economics*, **12** (3) 261–269.

Rosen, S. (1974) Hedonic prices and implicit markets: product differentiation in pure competition, *The journal of political economy*, **82** (1) 34–55.

Shapiro, E., S. Ni and X. Ang (2008) Singapore's new towns as models of sustainable built environment for China, Ph.D. Thesis, Cornell University.

Stack Exchange Inc. (2013) Spatial clustering wiht PostGIS, accessed on 12-05-2013, `http://gis.stackexchange.com/questions/11567/spatial-clustering-with-postgis`.

Stahel, W. A. (2013) The R-Function regr and Package regr0 for an Augmented Regression Analysis, accessed on 01-05-2013, `http://stat.ethz.ch/~stahel/courses/regression/`.

Tobler, W. (1970) A computer movie simulating urban growth in the Detroit region, *Economic geography*, **46**, 234–240.

Tonelli, M., M. Cowley and T. Boyd (2004) Forecasting office building rental growth using a dynamic approach, *Change*, **20** (09).

Tu, Y., S.-M. Yu and H. Sun (2004) Transaction-Based Office Price Indexes: A Spatiotemporal Modeling Approach, *Real Estate Economics*, **32** (2) 297–328.

Urban Redevelopment Authority (1995) *Changing the face of Singapore: through the URA sale of sites*, Urban Redevelopment Authority, Singapore.

Urban Redevelopment Authority (2010a) About Land Use Planning, accessed on 12-03-2013, `http://www.ura.gov.sg/land_use_planning/`.

Urban Redevelopment Authority (2010b) Sale of Sites, accessed on 12-03-2013, `http://www.ura.gov.sg/about/ura-saleofsites.htm`.

Urban Redevelopment Authority (2011) Development control parameters for non-residential development, *Technical Report*, Singapore.

Urban Redevelopment Authority (2013a) Real estate information system REALIS, accessed on 15-03-2013, `http://www.ura.gov.sg/realis/`.

Urban Redevelopment Authority (2013b) Urban Redevelopment Authority, accessed on 30-05-2013, `http://www.ura.gov.sg/`.

Van Eggermond, M. A. (2013) Data for hedonic pricing models of the commercial market in Singapore.

Wee, K. N. L. and K. W. Tong (2005) *The 4Rs of Asian shopping centre management*, Business studies., Marshall Cavendish Academic.

Weisberg, S. (2005) *Applied linear regression*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Wikipedia (2013a) Real Estate Bubble, accessed on 25-04-2013, `http://en.wikipedia.org/wiki/Real_estate_bubble`.

Wikipedia (2013b) Singapore, accessed on 08-03-2013, http://en.wikipedia.org/wiki/Singapore.

Wikipedia (2013c) Strata Title, accessed on 25-04-2013, http://en.wikipedia.org/wiki/Strata_title.

A Descriptive statistics - tables

A.1 Descriptive statistics office

A.1.1 Transaction office

Table 34: Descriptive statistics for the transaction office model

N	1,271				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 2.98	Mio. 1.11	Mio. 11.27	Mio. 0.02	Mio. 228.81
price_psm	19064.54	18117.78	7051.31	4033.14	57894.74
floor_area_sqm	148.37	65.00	559.15	3.00	10909.00
gpr_num	5.12	4.20	3.24	1.00	12.60
distance_bus	83.36	73.15	46.08	13.75	291.17
distance_mrt	535.83	380.86	560.59	82.23	3208.96
distance_cbd	2687.63	1803.48	2289.41	52.71	9986.31
distance_hgw	847.60	840.03	380.89	130.17	2454.18
distance_car_parking	53.55	40.70	53.11	0.88	561.30
bus_number_unique	38.67	32.00	23.99	3.00	81.00
perc_open	0.12	0.07	0.12	0.00	0.48
perc_res	0.25	0.13	0.22	0.00	0.65
perc_ind	0.00	0.00	0.01	0.00	0.18
perc_com	0.25	0.23	0.11	0.01	0.52
car_lots	1298.30	1191.00	823.91	0.00	4505.00
work_density_500	25835.50	23791.49	23683.50	838.16	101257.09
totpax_500	7750.94	8073.00	4598.36	0.00	18122.00
office_vac_rate_private	9.92	9.02	5.21	0.00	21.36
office_vac_rate_public	4.64	2.93	11.20	0.00	95.42
cluster_size_office	0.21	0.29	0.14	0.00	0.36
contract_build_area_ratio	0.18	0.03	0.84	0.00	9.31
floor_plot_ratio	0.10	0.02	0.31	0.00	5.99
accessibility_work	123594.07	131430.80	36742.59	210.48	188647.63
tenure_ongoing	3799.70	97.00	4723.09	54.00	10000.00
Dummy Name	**%**	**Dummy Name**			**%**
contract_date_factor_2011 Q1	10.9%	floor_number_fact_Floor <= 2			16.4%
contract_date_factor_2011 Q2	11.7%	floor_number_fact_Floor 3-5			28.8%
contract_date_factor_2011 Q3	9.8%	floor_number_fact_Floor 6-10			24.5%
contract_date_factor_2011 Q4	8.7%	floor_number_fact_Floor 11-15			13.3%
contract_date_factor_2012 Q1	8.5%	floor_number_fact_Floor > 16			17.0%
contract_date_factor_2012 Q2	17.8%	cluster_office_NA			13.1%
contract_date_factor_2012 Q3	10.9%	cluster_office_Katong			5.7%
contract_date_factor_2012 Q4	12.6%	cluster_office_Orchard			3.5%
contract_date_factor_2013 Q1	9.1%	cluster_office_Eunos			12.4%
bus_number_unique_fact_[3,22)	26.4%	cluster_office_CBD			46.6%
bus_number_unique_fact_[22,33)	25.5%	cluster_office_Victoria St.			18.6%
bus_number_unique_fact_[33,60)	23.5%	tenure_type_99 Years			51.1%
bus_number_unique_fact_[60,81]	24.6%	tenure_type_999 Years			15.4%
in_cbd	54.0%	tenure_type_Freehold			33.4%
perc_tertiary_edu_0	4.6%	perc_tertiary_edu_[0.152,0.262)			49.0%
perc_tertiary_edu_[0.085,0.152)	40.9%	perc_tertiary_edu_[0.262,0.311]			5.4%

A.1.2 Asking sale office

Table 35: Descriptive statistics for the asking sale office model

N	1,576				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 3.63	Mio. 2.00	Mio. 5.07	Mio. 0.28	Mio. 67.20
price_psm	21860.28	20499.39	11377.76	2583.31	80743.62
floor_area_sqm	172.94	108.00	196.86	13.01	2229.65
distance_bus	97.89	80.55	67.92	22.99	462.60
distance_mrt	563.47	388.24	524.97	48.12	2954.83
distance_cbd	2996.90	1777.15	2648.18	119.10	17489.29
distance_hgw	818.81	854.17	451.19	70.52	2529.12
distance_car_parking	122.83	68.09	155.74	6.41	1135.97
num_of_bus_lines	53.76	41.00	34.29	5.00	127.00
bus_number_unique	36.87	32.00	24.22	1.00	87.00
perc_open	0.10	0.07	0.10	0.00	0.40
perc_res	0.22	0.12	0.20	0.00	0.69
perc_ind	0.05	0.00	0.14	0.00	0.78
perc_com	0.25	0.24	0.15	0.00	0.54
car_lots	1127.19	896.00	847.85	0.00	4539.00
work_density_500	28148.48	23844.58	24115.37	628.96	101210.90
totpax_500	7715.65	7724.00	5822.57	0.00	29341.00
office_vac_rate_private	10.15	9.02	6.91	0.00	21.36
office_vac_rate_public	3.43	0.90	6.33	0.00	95.42
contract_build_area_ratio	2.21	0.09	8.37	0.00	86.07
floor_plot_ratio	0.10	0.02	0.25	0.00	3.66
cluster_size_office	0.18	0.29	0.15	0.00	0.36
accessibility_work	116604.27	124416.52	37080.11	0.00	200724.46
tenure_ongoing	3589.13	75.00	4736.39	1.00	10000.00
Dummy Name	**%**	**Dummy Name**			**%**
aircon_available	18.3%	cluster_office_NA			30.5%
aircon_central	2.5%	cluster_office_Katong			1.0%
cargo_lift	3.0%	cluster_office_Orchard			4.8%
ceiling_height	8.1%	cluster_office_Eunos			8.3%
electricity_supply	5.4%	cluster_office_CBD			39.1%
floor_loading	5.3%	cluster_office_Victoria St.			16.3%
lift_capacity	0.4%	tenure_type_N.A.			1.3%
bus_number_unique_fact_[1,21)	26.2%	tenure_type_30/60 Years			7.4%
bus_number_unique_fact_[21,33)	27.5%	tenure_type_99/999 Years			56.0%
bus_number_unique_fact_[33,50)	21.8%	tenure_type_Freehold			35.3%
bus_number_unique_fact_[50,87]	24.4%				
top_year_fact_(-1,1]	90.6%	perc_tertiary_edu_[0.085,0.124)			33.2%
top_year_fact_(1,2.01e+03]	4.4%	perc_tertiary_edu_0			5.1%
top_year_fact_(2.01e+03,2.02e+03]	4.9%	perc_tertiary_edu_[0.124,0.262)			58.7%
in_cbd	54.7%	perc_tertiary_edu_[0.262,0.311]			2.9%

A.1.3 Asking rent office

Table 36: Descriptive statistics for the asking rent office model

N	7,333				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	13427.87	6800.00	21115.27	300.00	324000.00
price_psm	78.23	75.35	39.01	10.09	269.11
floor_area_sqm	182.37	110.46	257.23	4.65	4227.61
distance_bus	110.30	87.01	80.30	8.70	821.13
distance_mrt	569.74	368.57	653.80	14.06	5966.45
distance_cbd	3087.60	1613.87	3514.14	71.99	24777.25
distance_hgw	845.70	815.19	471.52	66.32	3983.95
distance_car_parking	127.71	65.43	165.16	1.84	1117.91
num_of_bus_lines	52.38	41.00	32.28	4.00	127.00
num_bus_station_500	13.46	13.00	4.76	0.00	30.00
num_bus_line_500	129.08	128.00	62.23	0.00	291.00
bus_number_unique	36.54	33.00	22.41	1.00	87.00
perc_open	0.12	0.08	0.11	0.00	0.56
perc_res	0.19	0.16	0.17	0.00	0.78
perc_ind	0.06	0.00	0.15	0.00	0.81
perc_com	0.25	0.28	0.14	0.00	0.61
car_lots	860.73	596.00	802.29	0.00	6319.00
work_density_500	29755.60	23104.43	26331.89	85.36	105851.46
totpax_500	6717.35	5909.00	5772.80	0.00	37067.00
office_vac_rate_private	8.18	7.09	6.71	0.00	100.00
office_vac_rate_public	14.66	0.90	29.81	0.00	95.42
contract_build_area_ratio	1.80	0.11	7.72	0.00	148.13
floor_plot_ratio	0.09	0.02	0.26	0.00	6.38
cluster_size_office	0.18	0.29	0.15	0.00	0.36
accessibility_work	116266.84	118847.06	44815.21	0.00	199876.87

Dummy Name	**%**	**Dummy Name**	**%**
aircon_available	21.3%	cluster_office_NA	36.9%
aircon_central	3.3%	cluster_office_Katong	0.9%
cargo_lift	1.5%	cluster_office_Orchard	9.0%
ceiling_height	2.1%	cluster_office_Eunos	0.0%
electricity_supply	2.0%	cluster_office_CBD	39.5%
floor_loading	2.8%	cluster_office_Victoria St.	13.6%
lift_capacity	0.7%	perc_tertiary_edu_[0.085,0.124)	12.5%
bus_number_unique_fact_[1,21)	28.0%	perc_tertiary_edu_0	32.7%
bus_number_unique_fact_[21,33)	31.4%	perc_tertiary_edu_[0.124,0.262)	39.0%
bus_number_unique_fact_[33,50)	16.8%	perc_tertiary_edu_[0.262,0.311]	15.8%
bus_number_unique_fact_[50,87]	23.8%	in_cbd	54.6%

A.2 Descriptive statistics retail

A.2.1 Transaction strata retail

Table 37: Descriptive statistics for the transaction retail model

N	2,756				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 1.35	Mio. 0.84	Mio. 2.92	Mio. 0.01	Mio. 119.11
price_psm	35108.62	30478.18	20080.99	5046.51	150537.63
floor_area_sqm	46.42	28.00	143.26	1.00	6512.00
gpr_num	3.57	3.00	2.36	0.00	12.60
distance_bus	92.65	78.25	63.01	9.69	485.12
distance_mrt	744.99	472.51	705.82	57.04	3360.15
distance_hgw	955.09	882.44	562.03	58.24	2650.96
distance_car_parking	115.03	40.70	202.04	3.13	1841.23
bus_number_unique	29.08	26.00	17.35	2.00	81.00
perc_open	0.08	0.05	0.09	0.00	0.51
perc_res	0.38	0.44	0.21	0.00	0.78
perc_ind	0.01	0.00	0.04	0.00	0.44
perc_com	0.21	0.21	0.14	0.00	0.51
car_lots	1146.46	910.00	1102.16	0.00	5062.00
work_density_500	16363.90	6272.43	19662.61	56.23	78053.10
totpax_500	10713.86	10478.00	6453.73	0.00	33308.00
retail_vac_rate_private	9.49	5.10	12.16	0.00	43.26
retail_vac_rate_public	1.70	0.40	4.11	0.00	57.14
cluster_size_retail	0.12	0.03	0.14	0.00	0.35
contract_build_area_ratio	0.37	0.02	1.16	0.00	15.39
floor_plot_ratio	0.03	0.01	0.14	0.00	2.89
accessibility_work	111691.47	108510.99	41108.61	0.00	188647.63
tenure_ongoing	6391.90	10000.00	4680.68	34.00	10000.00

Dummy Name	**%**	**Dummy Name**	**%**
contract_date_factor_2011 Q1	7.1%	bus_number_unique_fact_[2,21)	29.0%
contract_date_factor_2011 Q2	10.2%	bus_number_unique_fact_[21,27)	26.8%
contract_date_factor_2011 Q3	8.6%	bus_number_unique_fact_[27,36)	22.4%
contract_date_factor_2011 Q4	5.6%	bus_number_unique_fact_[36,81]	21.8%
contract_date_factor_2012 Q1	9.3%	tenure_type_Freehold	57.8%
contract_date_factor_2012 Q2	23.6%	tenure_type_99 Years	24.2%
contract_date_factor_2012 Q3	19.8%	tenure_type_999 Years	18.0%
contract_date_factor_2012 Q4	9.4%	floor_number_fact_NotAvail	3.0%
contract_date_factor_2013 Q1	6.4%	floor_number_fact_Basement	13.5%
cluster_retail_NA	49.9%	floor_number_fact_Floor 1	33.9%
cluster_retail_CBD	13.0%	floor_number_fact_Floor 2	22.0%
cluster_retail_Orchard	7.5%	floor_number_fact_Floor >2	27.7%
cluster_retail_Katong	10.8%	in_shopping_mall	4.6%
cluster_retail_Geylang Rd.	1.8%	in_orchard	10.8%
cluster_retail_Rochor	17.1%		

A.2.2 Transaction land retail

Table 38: Descriptive statistics for the transaction retail land model

N	670				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 4.63	Mio. 3.50	Mio. 4.01	Mio. 0.49	Mio. 34.43
price_psm	28188.64	25784.62	14016.72	2215.19	115384.62
floor_area_sqm	170.65	142.00	123.37	47.00	1307.00
distance_bus	107.47	94.93	62.78	2.00	314.55
distance_mrt	625.23	482.36	493.97	76.60	3180.34
distance_hgw	971.49	911.11	542.85	33.68	5246.12
distance_car_parking	105.62	52.16	176.84	3.95	1322.93
bus_number_unique	22.12	20.00	13.03	1.00	81.00
perc_open	0.06	0.05	0.05	0.00	0.30
perc_res	0.38	0.35	0.21	0.01	0.76
perc_ind	0.01	0.00	0.05	0.00	0.41
perc_com	0.21	0.19	0.16	0.00	0.53
car_lots	2044.33	1967.00	1676.33	0.00	6212.00
work_density_500	15529.20	8736.31	18702.06	96.13	102716.73
totpax_500	11466.35	11621.00	5258.97	149.00	27389.00
retail_vac_rate_private	8.80	9.31	7.29	0.00	43.26
retail_vac_rate_public	3.83	1.94	5.79	0.00	100.00
cluster_size_retail	0.19	0.24	0.14	0.00	0.35
contract_build_area_ratio	1.99	1.25	11.52	0.23	291.87
floor_plot_ratio	1.09	1.00	0.69	0.01	11.43
accessibility_work	96701.55	97050.07	37132.75	0.00	196370.34
tenure_ongoing	6413.24	10000.00	4658.12	14.00	10000.00

Dummy Name	**%**	**Dummy Name**	**%**
contract_date_factor_2011 Q1	15.4%	bus_number_unique_fact_[2,21)	30.9%
contract_date_factor_2011 Q2	15.1%	bus_number_unique_fact_[21,27)	20.4%
contract_date_factor_2011 Q3	10.1%	bus_number_unique_fact_[27,36)	23.9%
contract_date_factor_2011 Q4	7.8%	bus_number_unique_fact_[36,81]	24.8%
contract_date_factor_2012 Q1	8.8%	cluster_retail_NA	30.9%
contract_date_factor_2012 Q2	12.4%	cluster_retail_CBD	22.4%
contract_date_factor_2012 Q3	8.8%	cluster_retail_Katong	10.0%
contract_date_factor_2012 Q4	10.6%	cluster_retail_Geylang Rd.	8.1%
contract_date_factor_2013 Q1	11.0%	cluster_retail_Rochor	28.7%
tenure_type_Freehold	62.7%		
tenure_type_99 Years	21.5%		
tenure_type_999 Years	15.8%		

A.2.3 Asking sale retail

Table 39: Descriptive statistics for the asking sale retail model

N	3,049				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 3.42	Mio. 2.00	Mio. 3.81	Mio. 0.01	Mio. 42.00
price_psm	43494.28	39559.34	29796.66	222.01	217438.57
floor_area_sqm	110.00	55.00	137.43	4.65	1950.95
gpr_num	3.13	3.00	2.01	0.00	12.60
distance_bus	92.81	66.74	63.37	10.41	462.60
distance_mrt	703.21	413.02	694.69	74.89	3454.40
distance_hgw	875.89	843.19	489.48	40.77	4300.62
distance_car_parking	183.51	88.00	222.41	1.82	1703.04
bus_number_unique	25.82	23.00	16.66	1.00	81.00
perc_open	0.07	0.05	0.08	0.00	0.55
perc_res	0.42	0.48	0.19	0.00	0.75
perc_ind	0.05	0.00	0.09	0.00	0.69
perc_com	0.16	0.12	0.14	0.00	0.52
car_lots	876.20	350.00	1168.70	0.00	6567.00
work_density_500	12668.44	4451.09	16096.34	134.47	78893.91
totpax_500	14069.57	12344.00	8003.37	0.00	41068.00
retail_vac_rate_private	13.11	5.27	16.87	0.00	100.00
retail_vac_rate_public	1.83	0.40	5.09	0.00	57.14
cluster_size_retail	0.09	0.00	0.13	0.00	0.35
contract_build_area_ratio	1.27	0.06	4.59	0.00	53.64
floor_plot_ratio	0.13	0.01	0.45	0.00	5.12
accessibility_work	105600.86	113930.73	39029.17	0.00	188647.63
tenure_ongoing	6046.66	10000.00	4837.69	1.00	10000.00

Dummy Name	**%**	**Dummy Name**	**%**
cluster_retail_NA	63.0%	bus_number_unique_fact_[2,21)	25.7%
cluster_retail_CBD	8.3%	bus_number_unique_fact_[21,27)	26.1%
cluster_retail_Orchard	6.3%	bus_number_unique_fact_[27,36)	23.7%
cluster_retail_Katong	9.4%	bus_number_unique_fact_[36,81]	24.4%
cluster_retail_Geylang Rd.	2.3%	tenure_type_N.A.	5.3%
cluster_retail_Rochor	10.7%	tenure_type_99 Years	29.3%
in_shopping_mall	3.9%	tenure_type_999 Years	5.5%
in_orchard	8.5%	tenure_type_Freehold	59.9%
aircon_available	18.2%	electricity_supply	5.9%
aircon_central	1.3%	floor_loading	1.0%
cargo_lift	0.4%	lift_capacity	0.3%
ceiling_height	2.9%	condition	29.6%
top_year_fact_NA	82.6%	top_year_fact_2013,2020	9.3%
top_year_fact_-2013	8.1%		

A.2.4 Asking rent retail

Table 40: Descriptive statistics for the asking rent retail model

N	5,860				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	10407.27	6200.00	12639.35	400.00	284145.00
price_psm	140.68	114.16	93.75	12.92	645.62
floor_area_sqm	97.37	55.74	123.67	4.65	1759.85
gpr_num	3.10	3.00	1.83	0.00	12.60
distance_bus	100.04	87.71	69.44	5.82	1305.00
distance_mrt	702.50	413.02	707.32	4.09	3515.51
distance_hgw	987.44	910.93	581.40	32.61	4351.75
distance_car_parking	176.39	89.39	242.49	1.82	1963.36
bus_number_unique	28.18	24.00	17.77	1.00	87.00
perc_open	0.09	0.05	0.11	0.00	0.63
perc_res	0.37	0.36	0.20	0.00	0.97
perc_ind	0.04	0.00	0.09	0.00	0.75
perc_com	0.18	0.18	0.14	0.00	0.56
car_lots	730.78	302.00	1099.77	0.00	6567.00
work_density_500	14151.20	6877.48	15030.94	33.65	94872.97
totpax_500	11493.90	9960.00	8447.55	0.00	45401.00
retail_vac_rate_private	8.77	4.92	12.49	0.00	100.00
retail_vac_rate_public	3.75	0.00	10.58	0.00	57.14
cluster_size_retail	0.07	0.00	0.12	0.00	0.35
contract_build_area_ratio	1.01	0.05	4.39	0.00	76.63
floor_plot_ratio	0.08	0.01	0.29	0.00	5.90
accessibility_work	105248.26	109511.80	48291.34	0.00	200724.46

Dummy Name	**%**	**Dummy Name**	**%**
cluster_retail_NA	66.0%	bus_number_unique_fact_[2,21)	25.7%
cluster_retail_CBD	8.8%	bus_number_unique_fact_[21,27)	27.7%
cluster_retail_Orchard	11.3%	bus_number_unique_fact_[27,36)	21.7%
cluster_retail_Katong	3.6%	bus_number_unique_fact_[36,81]	24.9%
cluster_retail_Geylang Rd.	1.8%	in_shopping_mall	15.3%
cluster_retail_Rochor	8.4%	in_orchard	18.1%
aircon_available	16.2%	electricity_supply	5.2%
aircon_central	1.9%	floor_loading	1.6%
cargo_lift	0.3%	lift_capacity	0.2%
ceiling_height	1.1%	condition	28.5%

A.3 Descriptive statistics industrial

A.3.1 Transaction industrial strata

Table 41: Descriptive statistics for the transaction industrial model

N	9,292				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 1.06	Mio. 0.81	Mio. 2.09	Mio. 0.12	Mio. 99.28
price_psm	5043.84	4832.95	2420.93	675.00	22604.44
floor_area_sqm	239.85	181.00	288.72	19.00	11175.00
gpr_num	2.43	2.50	0.34	0.00	3.50
distance_bus	149.13	140.57	83.69	14.85	462.63
distance_mrt	1165.20	1178.19	586.35	96.93	6413.85
distance_hgw	1083.68	769.52	986.02	37.13	4629.01
bus_number_unique	14.09	13.00	8.45	1.00	44.00
perc_open	0.05	0.03	0.06	0.00	0.44
perc_res	0.18	0.15	0.15	0.00	0.76
perc_ind	0.39	0.40	0.17	0.00	0.85
perc_com	0.02	0.00	0.04	0.00	0.28
work_density_500	7138.00	6843.97	3210.06	81.12	16518.23
contract_build_area_ratio	3.55	0.07	19.78	0.00	886.85
floor_plot_ratio	0.03	0.01	0.14	0.00	5.14
total_lots	332.06	116.00	634.88	0.00	6463.00
distance_car_parking	363.88	212.95	310.92	26.58	1161.36
distance_truck_parking	818.35	827.41	520.31	17.93	2399.92
industrial_vac_rate_private	7.97	7.96	3.19	0.00	29.39
industrial_vac_rate_public	1.17	0.50	1.76	0.00	11.65
accessibility_work	64452.24	60754.21	30360.99	3191.74	151820.06
tenure_ongoing	3144.27	54.00	4602.41	7.00	10000.00

Dummy Name	**%**	**Dummy Name**	**%**
contract_date_factor_2011 Q1	10.1%	bus_number_unique_fact_[1, 8)	29.2%
contract_date_factor_2011 Q2	14.2%	bus_number_unique_fact_[8,14)	21.5%
contract_date_factor_2011 Q3	11.0%	bus_number_unique_fact_[14,21)	25.9%
contract_date_factor_2011 Q4	6.5%	bus_number_unique_fact_[21,44]	23.4%
contract_date_factor_2012 Q1	9.5%	tenure_type_Freehold	31.1%
contract_date_factor_2012 Q2	13.7%	tenure_type_30 Years	4.8%
contract_date_factor_2012 Q3	14.4%	tenure_type_60 Years	54.8%
contract_date_factor_2012 Q4	15.1%	tenure_type_99 Years	8.7%
contract_date_factor_2013 Q1	5.6%	tenure_type_999 Years	0.8%
cluster_industrial_NA	49.7%	floor_number_fact_Floor < 1	10.8%
cluster_industrial_Bukit Merah	1.6%	floor_number_fact_Floor 2-3	22.0%
cluster_industrial_Kallang	5.2%	floor_number_fact_Floor 4-5	25.8%
cluster_industrial_Boon Lay	0.1%	floor_number_fact_Floor 6-8	30.7%
cluster_industrial_Pandan	2.2%	floor_number_fact_Floor > 8	10.6%
cluster_industrial_Eunos	25.4%		
cluster_industrial_Sembawang	0.9%		
cluster_industrial_Kallang Way	8.0%		
cluster_industrial_Clementi	7.0%		

A.3.2 Transaction industrial land

Table 42: Descriptive statistics for the transaction industrial land model

N	1,545				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 10.65	Mio. 5.50	Mio. 16.47	Mio. 0.25	Mio. 193.00
price_psm	2671.76	1372.78	3224.21	78.18	20501.29
floor_area_sqm	9066.47	4805.00	17589.71	218.00	127293.00
gpr_num	1.94	2.00	0.58	0.00	4.60
distance_bus	221.13	169.84	209.66	3.67	1282.88
distance_mrt	2054.79	1687.33	1475.69	123.51	7334.66
distance_hgw	1523.22	1165.09	1159.25	72.69	5627.17
bus_number_unique	5.60	3.00	6.49	1.00	33.00
accessibility_work	24926.30	17889.24	23288.24	0.00	145183.41
perc_open	0.06	0.02	0.11	0.00	0.75
perc_res	0.06	0.00	0.12	0.00	0.56
perc_ind	0.59	0.66	0.22	0.00	0.88
perc_com	0.01	0.00	0.04	0.00	0.55
work_density_500	3307.92	2005.95	3514.59	9.35	18154.23
contract_build_area_ratio	113.33	4.27	290.33	0.09	2380.64
floor_plot_ratio	1.03	1.00	0.25	0.00	6.29
total_lots	1220.30	150.00	2539.44	0.00	9330.00
distance_car_parking	326.66	239.97	284.22	16.55	1604.46
distance_truck_parking	404.92	260.72	403.27	10.93	2455.55
industrial_vac_rate_private	6.57	5.46	4.95	0.00	29.39
industrial_vac_rate_public	2.86	2.14	3.84	0.00	26.90
tenure_ongoing	788.81	14.00	2595.20	1.00	10000.00
accessibility_work	24926.30	17889.24	23288.24	0.00	145183.41

Dummy Name	**%**	**Dummy Name**	**%**
contract_date_factor_2011 Q1	12.6%	cluster_industrial_NA	71.6%
contract_date_factor_2011 Q2	14.9%	cluster_industrial_Bukit Merah	0.5%
contract_date_factor_2011 Q3	13.2%	cluster_industrial_Kallang	0.7%
contract_date_factor_2011 Q4	10.6%	cluster_industrial_Boon Lay	12.6%
contract_date_factor_2012 Q1	8.4%	cluster_industrial_Pandan	0.4%
contract_date_factor_2012 Q2	11.8%	cluster_industrial_Eunos	8.2%
contract_date_factor_2012 Q3	12.6%	cluster_industrial_Sembawang	3.6%
contract_date_factor_2012 Q4	9.7%	cluster_industrial_Kallang Way	1.1%
contract_date_factor_2013 Q1	6.2%	cluster_industrial_Clementi	1.3%
bus_number_unique_fact_[1, 8)	48.8%	tenure_type_Freehold	7.3%
bus_number_unique_fact_[8,14)	9.8%	tenure_type_30 Years	61.0%
bus_number_unique_fact_[14,21)	16.4%	tenure_type_60 Years	20.3%
bus_number_unique_fact_[21,44]	24.9%	tenure_type_99 Years	7.1%
		tenure_type_999 Years	4.3%

A.3.3 Asking industrial sale

Table 43: Descriptive statistics for the asking sale industrial model

N	5,402				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	Mio. 2.07	Mio. 1.30	Mio. 3.10	Mio. 0.08	Mio. 30.00
price_psm	6766.97	6458.35	3076.39	790.60	26912.93
floor_area_sqm	404.54	216.88	1003.12	18.95	11705.69
gpr_num	2.37	2.50	0.36	1.00	3.50
distance_bus	154.69	127.78	99.99	15.12	2522.89
distance_mrt	1172.38	1176.90	628.26	55.26	6304.37
distance_hgw	1090.80	857.00	969.40	42.14	5888.68
bus_number_unique	15.48	16.00	8.79	1.00	36.00
perc_open	0.04	0.03	0.05	0.00	0.37
perc_res	0.20	0.16	0.16	0.00	0.78
perc_ind	0.35	0.36	0.18	0.00	0.93
perc_com	0.02	0.00	0.04	0.00	0.24
work_density_500	6965.44	6843.97	3446.91	0.00	16518.23
totpax_500	6643.07	4561.00	6888.40	0.00	30654.00
contract_build_area_ratio	13.58	0.84	72.74	0.00	1352.68
floor_plot_ratio	0.08	0.02	0.33	0.00	6.01
car_lots	252.26	61.00	450.27	0.00	3182.00
lorry_lots	15.84	0.00	42.38	0.00	630.00
total_lots	304.92	116.00	572.77	0.00	8199.00
distance_car_parking	355.99	206.59	310.86	20.77	2458.19
distance_truck_parking	781.32	795.82	486.93	23.24	2853.09
industrial_vac_rate_private	8.45	7.96	3.45	0.00	29.39
industrial_vac_rate_public	1.15	0.09	1.96	0.00	11.65
accessibility_work	65062.74	60008.57	30217.32	0.00	139609.39
tenure_ongoing	4061.56	54.00	4881.80	1.00	10000.00

Dummy Name	**%**	**Dummy Name**	**%**
cluster_industrial_NA	51.9%	bus_number_unique_fact_[1, 8)	28.7%
cluster_industrial_Bukit Merah	3.1%	bus_number_unique_fact_[8,14)	22.4%
cluster_industrial_Kallang	1.9%	bus_number_unique_fact_[14,21)	24.6%
cluster_industrial_Boon Lay	0.5%	bus_number_unique_fact_[21,44]	24.3%
cluster_industrial_Pandan	1.9%	tenure_type_Freehold	1.8%
cluster_industrial_Eunos	26.4%	tenure_type_30 Years	7.2%
cluster_industrial_Sembawang	0.8%	tenure_type_60 Years	43.7%
cluster_industrial_Kallang Way	7.0%	tenure_type_99 Years	6.4%
cluster_industrial_Clementi	6.5%	tenure_type_999 Years	0.6%
aircon_available	14.1%	electricity_supply	14.5%
aircon_central	0.8%	floor_loading	16.6%
cargo_lift	6.6%	lift_capacity	4.0%
ceiling_height	23.0%		

A.3.4 Asking industrial rent

Table 44: Descriptive statistics for the asking rent industrial model

N	7,635				
Variable Name	**Mean**	**Median**	**St.Dev.**	**Min**	**Max**
price	13134.88	6250.00	16265.02	240.00	286527.00
price_psm	25.91	22.64	11.16	8.28	159.31
floor_area_sqm	641.64	260.87	869.93	4.65	9290.23
gpr_num	2.36	2.50	0.42	0.00	4.00
distance_bus	179.40	150.59	101.71	9.59	751.52
distance_mrt	1289.59	1398.13	677.78	108.61	6424.48
distance_hgw	1467.80	903.16	1361.67	14.00	5013.67
bus_number_unique	12.53	10.00	10.14	1.00	48.00
perc_open	0.06	0.05	0.06	0.00	0.56
perc_res	0.15	0.12	0.15	0.00	0.78
perc_ind	0.43	0.43	0.22	0.00	0.85
perc_com	0.02	0.00	0.04	0.00	0.48
work_density_500	7174.12	7649.45	3488.97	70.84	18154.23
totpax_500	5143.37	2708.00	6778.80	0.00	38108.00
contract_build_area_ratio	12.51	2.22	34.21	0.00	998.93
floor_plot_ratio	0.05	0.02	0.20	0.00	5.14
car_lots	267.15	106.00	452.02	0.00	3128.00
lorry_lots	27.74	0.00	57.70	0.00	630.00
total_lots	336.91	166.00	527.25	0.00	7497.00
distance_car_parking	342.13	258.19	271.94	12.86	1445.56
distance_truck_parking	610.49	518.41	488.87	26.85	2918.42
industrial_vac_rate_private	7.96	7.96	3.89	0.00	29.39
industrial_vac_rate_public	1.29	0.50	2.17	0.00	11.65
accessibility_work	58834.76	55946.07	31866.79	0.00	149166.93

Dummy Name	**%**	**Dummy Name**	**%**
cluster_industrial_NA	57.1%	bus_number_unique_fact_[1, 8)	27.3%
cluster_industrial_Bukit Merah	2.9%	bus_number_unique_fact_[8,14)	23.9%
cluster_industrial_Kallang	3.1%	bus_number_unique_fact_[14,21)	26.4%
cluster_industrial_Boon Lay	0.8%	bus_number_unique_fact_[21,44]	22.5%
cluster_industrial_Pandan	1.4%	electricity_supply	9.6%
cluster_industrial_Eunos	16.2%	floor_loading	16.5%
cluster_industrial_Sembawang	4.0%	lift_capacity	1.6%
cluster_industrial_Kallang Way	10.7%	aircon_central	1.0%
cluster_industrial_Clementi	3.7%	cargo_lift	3.8%
aircon_available	11.7%	ceiling_height	19.7%

B Further results of the full models

B.1 Model fit full models

Figure 29: Model fit for the full models