Rent modelling of Swedish office markets
- Forecasting and rent effects

Hamza Harrami & Oscar Paulsson
Master of Science thesis

Title: Rent modelling of Swedish office markets – Forecasting and rent effects
Authors: Hamza Harrami & Oscar Paulsson
Department: Department of Real Estate and Construction Management
Master Thesis number: TRITA-FOB-ByF-MASTER-2017:06
Archive number: 464
Supervisor: Sviatlana Engerstam
Keywords: Swedish office market, vector autoregression, forecast modelling, rent effects

Abstract

The Swedish office markets has been emerging the last decade towards a higher rental level equilibrium. The aim of this study is to investigate the fundamental drivers of office rents and modelling of office rent forecasts in five Swedish office submarkets; Stockholm (2), Gothenburg (2) and Malmö (1). The methodology is a combination of economic theory and econometric analysis. The product is an econometric model. By using the estimated drivers, office rent forecasts are modelled and computed based on a vector autoregression-model. Our results show that office stock and vacancy, in lagged fashion, are statistically superior in explaining office rent development. OMX30 was evident to be the largest macro-driver in explaining office rent. The generated forecasts were significant and valid in the CBD-submarkets. However, the forecasts for the Rest of Inner City (RIC)-submarkets were not as precise. The results also show that the forecasts move more linearly compared to the actual office rent data that move more “step-wise”.
Acknowledgement

First and foremost we would like to express our deepest gratitude to our supervisor Sviatlana Engerstam for valuable guidance and input during the writing of this thesis.

We are thankful to prof. Mats Wilhelmsson for his input on the econometrics of this study.

We are thankful for the opportunity to do this thesis in collaboration with WSP (Analysis & Strategy) in Stockholm. We would also like to express our gratitude to all the consultants, especially our supervisor Monika Fridenäs, for sharing knowledge and providing facilities.

In order to conduct this study we needed data in which the availability was limited. We would like thank JLL that stepped forward and provided us valuable real estate data that made this study possible.

We would like thank Olof Netzell, model-expert at Finansinspektionen, for the discussion and initial thoughts of the econometric modelling in this study.
Contents

Abstract .............................................................................................................................. I
Acknowledgement ............................................................................................................. II
1 Introduction ...................................................................................................................... 1
2 Purpose and background ................................................................................................. 3
  2.1 Research questions ......................................................................................................... 3
  2.2 General real estate market characteristics ..................................................................... 3
  2.3 Submarkets ................................................................................................................... 5
3 Theoretical framework ..................................................................................................... 7
  3.1 Literature review on dynamic models of the office market .......................................... 7
  3.2 Previous studies on vector autoregression and forecast modelling ............................... 8
  3.3 Office market specifics ................................................................................................. 11
  3.4 DiPasquale-Wheaton four quadrant model ................................................................... 12
  3.5 Business- and real estate cycles .................................................................................. 14
4 Data .................................................................................................................................. 15
  4.1 Variables ....................................................................................................................... 15
  4.2 Historical trends of submarkets .................................................................................... 17
5 Econometric methodology ............................................................................................... 21
  5.1 Pooled regressions ........................................................................................................ 21
  5.2 Vector Autoregression (VAR) ....................................................................................... 24
  5.3 Forecasting .................................................................................................................... 26
  5.4 Forecasting performance and test statistics ................................................................... 28
6 Results and analysis ......................................................................................................... 30
  6.1 Pooled regression .......................................................................................................... 30
  6.2 Forecasting .................................................................................................................... 34
7 Concluding remarks ......................................................................................................... 41
  7.1 Suggestions for future research .................................................................................... 42
References ........................................................................................................................... 43
Appendix – Econometric dictionary .................................................................................... 46
1 Introduction
The Swedish office rental markets, particularly the large markets; Stockholm, Gothenburg and Malmö, has been emerging the last decade meeting an almost consistent robust demand (Nordic City Report, 2016). As a result, the office rents have been increasing steadily towards a higher equilibrium. The office market in Sweden is moving towards an interesting future where the interest rates have been declining, almost without exception, for the past 20 years and are positioned as negative today (Riksbanken, 2017). Within the recent two years, office rents have peaked up while vacancies has peaked down.

Considering real estate markets in general, office markets are more synchronized in terms of exposure to macro-effects and performance of the real estate within the market (Hekman, 1985). The heterogeneity of the residential and retail markets makes it more complex to analyze. We found it therefore interesting to study the drivers and future development of office rent markets in Sweden. Five submarkets have been investigated in this study; Stockholm (2), Gothenburg (2) and Malmö (1). The methodology of this study is a combination of economic theory and econometric analysis. The product is an econometric model. The office drivers are estimated through OLS dynamic regressions. Furthermore, the future office rent development is conducted by forecast modelling (in-sample and out-of-sample) using Vector Autoregression (VAR).

To our knowledge, forecasting office rents (using VAR) has never been conducted when applying it to the Swedish real estate market. On the other hand, studies applying VAR modelling in the real estate markets outside Sweden are numerous (see among others, Bönner (2009) and Karakozova (2005). Due to limitations on availability of office rents, they used average rent prices on cities and large markets. A belief of better data availability today motivated this study further where the three largest regions in Sweden were found to be superior in this aspect. Quarterly quoted data between 2000q3 and 2016q4 were used in the regressions.

The office markets in the chosen submarkets will be investigated based on two research questions. Firstly; What are the fundamental drivers of office rents in chosen submarkets? Following the results retrieved from the first section, a second research question was formulated to be answered; Does Vector Autoregression (VAR) as an underlying econometric model estimate significant office rent forecasts?
The estimates from the OLS pooled regressions showed that the drivers of office rents differed between the submarkets. In line with previous research, office stock and vacancy was shown to be drivers of office rents for all submarkets. However, the result indicated that there is a lagged effect on office rents on these variables in line with the dataset and economic theory. Considering the macro-variables, the effect of OMX30 was shown to be statistically superior. Diverging from previous studies, interest rate (both short and long) did not have any causal impact on office rents. It can be found interesting since office stock and interest rate were highly correlated in all submarkets. Moreover, one can argue that interest rates have an indirect impact on office rents (through the supply-side).

The forecast modelling based on VAR-econometrics generated significant and valid predictions, both in-sample and out-of-sample, for the CBD-submarkets (especially Gothenburg and Malmö). Conversely, the forecasts for the two RIC-submarkets (Rest of Inner City) were not as precise and significant. It can be argued that these submarkets are too large; the real estate are more heterogenic and the location factor has more impact in which makes it difficult to generate significant drivers and forecasts. The generated forecasts follows a linear development compared to the “step-wise” development of the observed data. Furthermore, the uncertainty bands of 95% were thin for all submarkets during the first year of the forecast horizon.

In line with Englund et.al. (2005), it can be argued that the “step-wise” development of the office rentals are generated by new contract-signings due to an illiquid office market. Therefore, there are arguments against the office rental pricing during periods since long rental contracts do not consider the change in market price (only marked against CPI) during the contract horizon. Hence, office rental price can be wrongly priced during longer periods.

The rest of the paper is organized as follows. Section 2 covers some fundamentals and geographic delimitations. Section 3 contains a review of previous literature and theoretical framework. A description of our data and historical trends of the submarkets are presented in Section 4. The econometric methodology constituting the basis of this study are presented in Section 5. Further, Section 6 covers our results and interpretations while Section 7 includes the most important conclusions and suggestions for future research.
2 Purpose and background

The aim of this study is not to evaluate the econometrical methodology regarding office rent forecasting. The econometric methods and models used in this paper are in line with previous research within this field of study. The purpose of this study is instead to investigate the fundamental drivers of office rent and modelling of office rent forecast models for chosen submarkets.

2.1 Research questions

In order to capture the scientific elements of this field of study, two research questions have been formulated to be answered in this study. Moreover, the questions will constitute the two parts of our purpose with this study.

I
- What are the fundamental drivers of office rents in chosen submarkets?

II
- Does Vector Autoregression (VAR) as an underlying econometric model estimate significant\(^1\) office rent forecasts?

This study solely investigates the fundamental drivers of office rents and do not take speculative factors into account. Speculations, such as political elections, that could affect office rents are not included in the forecast modelling.

2.2 General real estate market characteristics

In this study, it is of importance to distinguish between the real estate markets, including the office market, from other economic markets. These market specifics will constitute a theoretical understanding in our forecasting modelling (Gantenbein, 1999; Chinloy, 1988; Schulte et al., 2000). Market inefficiencies can briefly be defined as follows.

*Heterogeneity*

Real estate is individualized by factors such as architecture, location, age, occupancy etc. This leads to weak comparability in terms of economic goods (other real estate), limited market transparency and low liquidity in the market.

\(^1\) Definition: a significant forecast has a MAPE-estimate (error in percentage) below 10% in this study.
**Immobility**

Real estate is immobile and connected to its original location. In this way, the piece of real estate is affected by the development of the surrounding area.

**High transaction costs**

The illiquidity of the real estate market can also be explained by high transaction costs. As a result of low transparency, there is an information asymmetry which derives high broker costs (especially in commercial real estate). Additionally, taxes and notary fees increase the transaction cost when executed.

**Illiquidity**

The real estate market is relatively illiquid compared to other economic markets. The high transaction costs and heterogeneity are, among others, the main affecters of this. Furthermore, the illiquidity, among others, of the real estate market creates busts and booms.

In economic theory, real estate cycles are lagged to the general business cycles. This can be explained by macroeconomic influences and the length of construction processes. Further, the real estate market takes time to react towards demand shocks. Real estate cycle determinants can be explained by:

**Macroeconomic influences**

As other economic markets the real estate market, especially the office market, are linked to the macroeconomic situation. A positive development of macro factors such as interest rate and GDP generates an increase in office space demand. For instance, the outcome of positive development will increase the demand for office employees thus, the demand for office space.

**Length of construction processes**

On average, it takes approximately 2-5 years from the planning stage to the completion of a real estate project. This creates a time lag in the market; from the positive demand shock of office space until the new office space eventually enters the market. This lag constitutes the underlying mechanism of the cyclical behaviour of the real estate market compared to the general business cycle. Continuously, this constitutes the fluctuations in rent and vacancy levels.
2.3 Submarkets
This study will be delimited by studying five submarkets in the three largest cities in Sweden; Stockholm, Gothenburg and Malmö. These submarkets were chosen because of the market scale, liquidity of markets and good data accessibility. Further, each submarket will be investigated individually with market-specific real estate data. Thus, the macro data will be the same for all submarkets.

We use geographical delimitations of the five submarkets set by JLL (provider of real estate data) in line with the common perception of these submarkets (JLL, 2016). The Stockholm market has been divided into two submarkets; Stockholm Rest of Inner City (denoted RIC, see figure 2.1) and Stockholm CBD (see figure 2.2). The market of Gothenburg has been divided into two submarkets as well; Gothenburg Rest of Inner City (denoted RIC, see figure 2.3) and Gothenburg CBD (see figure 2.4). Lastly, only Malmö CBD (see figures 2.5 and 2.6) has been included in our study. Physical boarders can be illustrated in the following figures sorted by order of market scale.

Submarkets - Stockholm

Figure 2.1 and 2.2: Illustrates the submarkets of Stockholm RIC and Stockholm CBD. (JLL, 2016)

Stockholm RIC is by marginal the largest submarket, in both quantity and area, investigated in this study. The market area captures all office real estate within the inner-city duties. Norra Bantorget in the north and the Swedish Parliament in the south constitutes the outer boundaries of the Stockholm CBD market.
Submarkets - Gothenburg

Gothenburg RIC is the second largest submarket in this study. Worth mentioning is that areas Norra Älvstranden and Mölndal (in Figure 2.3) coloured in grey are expanding rapidly with new construction and establishment of multinational corporations. Moreover, the Gothenburg CBD is concentrated around the area of Avenyn.

Submarket - Malmö

For Malmö only CBD was included in the revision. The main reason for this was that some regions in Malmö e.g. Hyllie have seen significant changes due to political decisions and huge investment in the area during the studied time period. Moreover, Hyllie were a nearly undeveloped area in the first 10 years of the 21th decade but has since seen the highest rates in entire Malmö. Due to this rapid change of rents, this area has been excluded from this study.
3 Theoretical framework
Our review of previous published research is divided into two sections. Section 3.1 covers a brief description of previous research on dynamic models of office markets. Continuously, Section 3.2 includes an introduction and literature review on Vector Autoregression (VAR) and forecast modelling. The focus of Sections 3.3 -3.5 is to increase the understanding of the theory constituting the foundation of this study.

3.1 Literature review on dynamic models of the office market
Previous literature on dynamic models of the office market are relatively limited and the associated research has not been published as frequent as other areas within the field of real estate economics, especially in Scandinavia. In general, published work within the subject emanates from two different points of views; research investigating the econometric methodology of office rent forecasting (comparing different econometric models) and research about a specific market area.

The causal relationship between changes in real rents and deviations in vacancy rates from the natural vacancy is a well-established subject when modelling within the real estate market. This linkage has its origin from the traditional labour economics where changes in real wage inflation has been linked to the deviation in employment rates from the natural employment rate. Inspired by this, Blank et al. (1953) was first to explicate the relationship between vacancy rates and changes in rents. Investigating the data of rents and vacancy rates in six US cities between 1932-1937, the authors prove the relationship between the two variables.

By using a two-stage least squares model for rent and demand, Ho (2003) forecasts the Hong Kong office market. Among his findings, he shows that demand of office space is positively correlated to office employment (in line with economic theory). However, his results show that demand is negatively correlated to rent. Further, on one hand, his study shows that office rent is positively related to office employment. One the other hand, office rent is inversely related to office stock.

Hekman (1985) investigated the rental price adjustment mechanism in 14 metropolitan office markets by using a two-equation model. Constructing one equation where supplied quantity is a function of rent and other variables, and another in which rent is a function of vacancy and other variables. His results show that construction of office space responds strongly to long-term office employment and real rents. Moreover, markets rents of building under construction responds strongly to market areas in both the suburbs and central parts of the city.
Diverging from other studies, Rosen (1984) presented a methodological framework for forecasting key variables of the office space market; flow of new construction, office stock, vacancy and office rent. The author conducted empirical estimations mainly based on data from the office market in San Francisco. Rosen retrieved statistically significant results for desired office stock and changes in office rents. However, the results regarding supply of new office space was not statistically significant.

Taking on the case of Helsinki, Karakozova (2005) examines short- and long run movements in CBD using data between 1971-2001. A reduced form of demand-supply approach is used in order to model real office rents. The author investigates three econometric methods with the aim of finding the relationship between rent and their determinants: a regression model, an error correction model (ECM) and an autoregressive-moving average model (ARIMAX). Continuously, the results indicate an impact of supply-side effects (new office building completions) on Helsinki office rents compared to demand-side drivers (GDP, office sector employment etc.). Her findings were in line with previous research, where different supply variables were found to be insignificant in explaining the short term movement in office rent (see among others, D’Arcy et al., 1997b; White et al., 2000 and Hendershott et al., 2002). Moreover, the study also shows that there exists a long-run relationship between office rent and demand drivers.

Englund et.al. (2005) investigated adjustments in property space of the Stockholm office market between 1977-2002, by using standard hedonic methods. Interestingly, the authors found that long-term contracts played an essential role in the aspect of office rents in Stockholm. Their study highlight that the property market is slow to adjust because tenants are constrained by long-term contracts. Consequently, this give rise to “hidden vacancys” meaning the difference between space occupancy and demand at current office rent.

3.2 Previous studies on vector autoregression and forecast modelling

Previous research on VAR as an underlying econometric model for rent forecast modelling are rare. Generally, publications within VAR and forecast modelling studies these subjects separately. Much of the publications within this field investigates the econometric methodology of VAR as a model. For a deeper review of VAR modelling, we refer to various textbooks, for example, Hamilton (1994); Hatanaka (1996); Lütkepohl & Krätzig (2004) and Lütkepohl (2005). To our knowledge, forecasting modelling of office rents has never been published in the case of the Swedish market.
Bönner (2009) studies underlying econometric models and forecasting of office rents in nine German office markets. The dissertation is motivated by the research gap, from a scientific point of view, in this field of study because it is mainly done by commercial organizations. By evaluating different econometric models, Bönner can generate different rent forecasting models for nine different office markets in Germany. Among his findings he concludes that univariate models outperform multivariate models in the short run. However, the author states that the case is vice versa regarding the long run forecasting. Interestingly, Bönner remarkably finds that increased volatility appears after forecasting horizons of three years within each city rent series.

Our thesis is inspired by Bönner’s study, though with another approach. Firstly, the forecasting of office rents are applied on the Swedish office market. Compared to this study, he focuses on evaluation and testing of different econometric models before forecasting. Instead of looking at the rent forecasting for entire cities, this study focuses on submarkets within a city in order to retrieve more specific results on a micro-level. From the thorough research of Bönner (2009), a resolution have been made to focus on VAR as an underlying model for forecasting. Initially, a Vector Error Correction Model (VECM) was constructed. However, the error correction term (ECT) was found to be insignificant which indicates that the model’s short run adjustment behaviour is insignificant. As a result, the long run equilibrium relationship between the non-stationary variables will be spurious and statistically insignificant (Lütkepohl & Krätzig, 2004).

In recent years, ‘structural’ office rent models attempt to use both supply and demand indicators in which the changes of office rents are generated by the interaction of supply and demand of accommodation. In order to provide more detailed information regarding market dynamics, these models tend to be constructed using sophisticated econometric techniques such as ARIMA, VAR and cointegration models. For instance, VAR modelling was used by McGough and Tsolacos (1995) to forecast office rents at UK national level. Their result showed satisfactory forecasts of rents for four quarters ahead.

An inventory of explanatory variables used in previous research has been made and is illustrated in Table 3.1 (see next page) where variables that has been used the most are highlighted. It can be found that rents, vacancy, supply, economic activity, interest rates and employment has been used the most when modelling office rents. In line with previous research, all the highlighted variables in the table are used in this study.
Table 3.1: Illustrates the use of explanatory variables used in previous research and structured in chronological order. Most used variables are highlighted in the summarization.
3.3 Office market specifics
This study investigates the office market particularly in which can be distinguished from other real estate types in terms of relation between risk and yield/rent affection. Both micro- and macroeconomic factors that increases the risk affects the yield of various real estate segments differently. In general, the residential market is considered to have the lowest risk followed by the office market. In the event of economic recessions, rents of the office market are relatively less exploited to risk compared to retail and industrial real estate markets.

The real estate sector can be divided into two segments, residential and commercial properties. The commercial segments are broad and includes various property types. The office market has its own characteristics that distinguish it. The overall quality of the building will affect the rent level but also the flexibility. With the most influential factor for the rent level in the office segment to be the location. The classification for location can be in A-C or a similar descriptive system. Further, the most attractive location is often the CBD and that is reflected in rental prices. For the secondary area it can be more complicated. In general, that submarket is often next to the CBD but can in some cases be further away. This is due to that specific nodes can be created where the rent level increase sufficient with the clustering of companies there. Kista in Stockholm can be mentioned with the clustering of IT-companies; also Hyllie outside of Malmö is a node where the rent level has seen adequate rent upturns. Therefore, the secondary market has a greater complexity than the CBD. It is often a larger area with higher heterogeneity and larger range between highest and lowest rent within the area.

Figure 3.1: Illustrates different property types’ position to yield and risk. Based on Morena and Truppi (2016).
3.4 DiPasquale-Wheaton four quadrant model
The model created by DiPasquale and Wheaton (1992) can be used to explain the mechanisms of the real estate market. According to DiPasquale and Wheaton, the four quadrant model subdivide the general real estate market into two major markets; real estate space market and real estate asset market. The model illustrates the connection between the two real estate markets from a macroeconomic perspective and how they are affected by financial markets. Monetary factors such as Price or Rent is opposing to supply factors, Stock and Construction. With the Demand as an overall constraint in the model. Moreover, the model explains the implications derived from exogenous shocks in rents, construction, asset prices and the stock of real estate.

Figure 3.2: Illustrates office characteristics that affects the rents. Based on Morena and Truppi (2016).
This thesis theoretically focuses on Quadrant 1 in which indicates the demand function on the space market demanded by users. In this case, the users are represented by the office employees. The model states that office rent is a function of supply and can be derived where $E$ is number of office employees, $a$ and $b$ are demand parameters, $S$ is supply and $R$ is rent:

$$S = E(b - a \times R) \quad | \quad R = f(S)$$

(1)
3.5 Business- and real estate cycles
The business and real estate cycle can be particularly differentiated in terms of cycles. The fundamental structure of the booms and recessions in the real estate market are affected by the real and monetary economy. Furthermore, each step in the real estate conjuncture process can be linked to the parallel business cycle (see figure 3.4).

![Diagram of business and real estate cycles](image)

**Figure 3.4**: Illustrates the process and linking of business and real estate cycles. Barras (1994).

Theory and previous research, in line with the collected data in this study, indicates that the business and real estate cycle are unsynchronized; the real estate market is lagged (see Figure 3.5). Moreover, the effects from the business cycle reach the real estate in a lagged fashion due to illiquidity and tangible assets.

![Diagram of lagged real estate cycle](image)

**Figure 3.5**: Illustrates the lagged real estate cycle parallel to the business cycle. Based on Fanning (2014).
4 Data
The empirical foundation of this paper exclusively consists of time series data. The time period we chose was between 1st of October 2000 until the 31st of December 2016. Further, the data consists of variables from different sectors that constitute the forecast modelling; real estate market, macro-based and demographics (see Table 4.1). Data availability imposes even further restrictions on the analysis. We did not include the period before 2000 due to the lack of data regarding the real estate-variables. Despite this, with data going back to the millennium it is possible to cover two economic cycles that includes both recession and periods of booming economy. By reason of the slow movement of the real estate market, data are far from always available on high frequency. Therefore, we solely observe data on a quarterly basis. Data quoted on a daily or monthly frequency (macro-variables) were converted into quarterly quotes by using the arithmetic mean (average) during that period. This allowed us to obtain information of ten variables (inclusive Rents) at quarterly frequency across five submarkets (previously presented in Section 2.3); Stockholm CBD, Stockholm Rest of Inner City (RIC), Gothenburg CBD, Gothenburg Rest of Inner City (RIC) and Malmö CBD. Furthermore, each time series contains 66 observations (quarters) making it 2790 observations in total.

The real estate data used in this thesis was provided by JLL. The information contains of quarterly data regarding office rents, vacancy, new office space and office stock per submarket. Consequently, the geographical delimitation regarding the different submarkets set by JLL was used in order to simplify. Noteworthy is that office rent data is quoted as prime rents\(^2\) instead of average rent during a quarter.

4.1 Variables
Office rents (defined as prime rents) will be viewed as a dependent variable in this study. The rent reflects an object of high quality with preferable location. In general, prime rents are considered a driving force for the overall office rent market. Continuously, prime rents are measured in \textit{SEK per square meter and per year}, consistent with general market notation. The independent (explanatory) variables are chosen in line with previous research and economic theory (see Table 3.1). Moreover, all the independent variables are illustrated in Table 4.1.

---

\(^2\) Prime rent = The highest level of rent price obtained on the market for a locale with a high standard. (JLL, 2016)
**Explanatory variables** – real estate indicators

*Vacancy* is believed to be a strong independent variable with considerable effect on the office rent. This variable is measured in percentage of total sqm of a property. The underlying theory indicates that if the vacancy increase rent should decrease and vice versa. Further, vacancy can be seen as both supply and demand indicator normally striving against equilibrium (natural vacancy). *Office stock* constitutes the outstanding supply of the market and is a fundamental supply-variable. Further, *New Office Space* is included in the model in order to capture the increase in supply per year.

**Explanatory variables** – regional and macro

The interest rate, both short and long term, is included to capture risk-free rate on the market. Data was retrieved from the central bank (Riksbanken) where a treasury bill (*3 months*) and a bond (*10 years*) represents the short- and long run effect on investment strategies. A low rate eases the raising of capital in which creates beneficial times for companies and subsequently the office (rent) market. Logically, when high rate, the scenario is the contrary.

*OMX30* is an index, created by NASDAQ, of the thirty most traded shares in Stockholm Stock Exchange. The index is included to be an indicator of the overall Swedish companies’ performance. Further, this variable should be able to indicate were Sweden is on a conjuncture-scale, from recession to boom. *Consumer price index (CPI)* accounts for the development of prices over time (inflation) and is collected from Statistics Sweden.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit of measurement</th>
<th>Correlation</th>
<th>Characteristics</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacancy</td>
<td>Percentage (%)</td>
<td>Negative (S)</td>
<td>Indicator</td>
<td>Submarket</td>
</tr>
<tr>
<td>Office Stock</td>
<td>Square meter (m²)</td>
<td>Negative (S)</td>
<td>Supply of office space</td>
<td>Submarket</td>
</tr>
<tr>
<td>New Office</td>
<td>Square meter (m²)</td>
<td>Negative (M)</td>
<td>Increase in supply</td>
<td>Submarket</td>
</tr>
<tr>
<td>Interest rate 3M</td>
<td>Percentage (%)</td>
<td>Positive (W)</td>
<td>Cost of capital</td>
<td>National</td>
</tr>
<tr>
<td>Interest rate 10Y</td>
<td>Percentage (%)</td>
<td>Positive (M)</td>
<td>Cost of capital</td>
<td>National</td>
</tr>
<tr>
<td>OMX30</td>
<td>Index (1986=125)</td>
<td>Positive (M)</td>
<td>Stock market index</td>
<td>National</td>
</tr>
<tr>
<td>Consumer price index</td>
<td>Index (1980=100)</td>
<td>Positive (W)</td>
<td>Price index</td>
<td>National</td>
</tr>
<tr>
<td>GDP</td>
<td>Monetary (SEK)</td>
<td>Positive (M)</td>
<td>Economic performance</td>
<td>National</td>
</tr>
<tr>
<td>Unemployment</td>
<td>Absolute value</td>
<td>Negative (S)</td>
<td>Demand of work space</td>
<td>Regional</td>
</tr>
</tbody>
</table>

**Table 4.1:** Illustrates a descriptive table of the independent variables and their unit of measurement, correlation to rents, characteristics and data-level. (S) Strong >70%, (M) Medium = 30-70%, (W) Weak <30%
In order to capture the demand for office space, Unemployment is incorporated as a proxy for demand. Underlying theory states that if the unemployment is low, the need for workspace in the contrary will be high, i.e. demand for office space should be high and the rents should increase. The unemployment rate is quoted on a municipal level (Stockholm, Gothenburg and Malmö) and collected from Statistics Sweden.

The Gross Domestic Product (GDP) have been included in the study for two reasons; to capture the performance of the country over time and indicate where Sweden is in the economic cycle. Secondly, as Karakozova (2005) argues, to measure the demand for office in Sweden and subsequently the viewed submarkets. The data have been collected from Statistics Sweden and is of expenditure approach.

4.2 Historical trends of submarkets
This section provides graphical illustrations of historical trends in each submarket investigated in this study. OMX30 index is used as a conjuncture indicator. Moreover, rents are presented in nominal values, vacancy in percentage and OMX30 index in absolute values. Figures are constructed based on data provided from JLL (2016).

*Office rents*

![Nominal office rental price over time](image)

**Figure 4.1**: Illustrates historical office rents, quoted as SEK/sqm/year, for each submarket between 2000q3 and 2016q4.
As the graph (Figure 4.1) illustrates, office rents in Stockholm (especially CBD) has been deviated noticeably from Gothenburg and Malmö. At some quarters, it differs more than 4000 SEK between Stockholm CBD and Malmö CBD. Further, one can see that the Stockholm CBD and RIC markets are observably more volatile than the rest. Gothenburg and Malmö markets follow the opposed pattern where the movement in rents are slow. This can be seen in the two economic recessions\(^3\) where the rental prices in Stockholm were much more volatile.

Figure 4.2 above shows the average rent of all submarkets together with Stockholm CBD, benchmarked against OMX30. The graph clearly shows that the financial (non-tangible) market got perceptibly more affected by the recession compared to the real estate (tangible) market. One can see that the financial market is visibly more volatile than the real estate markets, especially the average rent market. Interestingly, one can observe that the effect of both recessions are lagged in the real estate market, in line with theory.

In general, the vacancies for all submarkets follow the same pattern. Stockholm RIC market is the most volatile to economic recessions which is visualized in the graph where the vacancy increased tremendously by the IT-bubble recession. Further, the vacancies move toward an equilibrium seemingly between 2-6%.

Figure 4.3: Illustrates historical vacancy in percentage for each submarket between 2000q3 and 2016q4.

Figure 4.4: Illustrates historical vacancies in Stockholm CBD and average submarket (left axis) vacancy against OMX index (right axis). (NASDAQ, 2016)
As Figure 4.4 shows, the average vacancy and Stockholm CBD vacancy follows the same pattern with marginal differences in absolute values. In line with economic theory, one can clearly observe that the real estate market cycle is lagged compared to the non-tangible business cycle. More precisely, the effect of the first recession is lagged two years in the real estate cycle. The latter six years follow the expected development of negative correlation between vacancies and OMX30, indicating a flourishing economy.

**Office stock**

The office stock grows approximately in the same (proportionally) pace for each submarket. Logically, the Stockholm RIC market hold a larger area than the rest, resulting in an observably larger stock. Data indicates that there will be more activity for Stockholm and Gothenburg CBD regarding future construction of office space.

![Accumulated office stock over time (in thousands m²)](image)

**Figure 4.5**: Illustrated accumulated office stock in each submarket between 2000q3 and 2016q4. The data is visualized through the constant order of Stockholm CBD, Stockholm RIC, Göteborg CBD, Göteborg RIC and Malmö CBD for each year.
5 Econometric methodology

In the analysis, each submarket will separately be investigated with the perspective of forecasted office rents. This procedure is chosen with the consideration that the different markets have different structures and price levels. Due to this, it is not considered that the submarkets are similar enough to compare or to group (constructing of a panel data set). Further, all variables containing monetary data consists of nominal values. Consequently, it is assumed that the inflation effect is the same over time on these variables.

The method is divided into two different subsections, both based on historical data. Firstly, to analyze which fundamental factors that affect office rents in chosen submarkets using Ordinary Least Squares (OLS) multivariate dynamic regressions. The regressions are based on grounded theory (economic theory) and variables that are expected to show to effect on office rents from previous studies. Continuously, the result from this empirical investigation will constitute the use of variables in the second part. Secondly, to forecast office rents in chosen submarkets using Vector Autoregression (VAR). Forecast scenarios of office rents are generated for each submarket. Further, both in-sample and out-of-sample forecasts will be generated. Econometric expressions (not explained in this section) are explained in Appendix.

5.1 Pooled regressions

In this section, the methodology behind the pooled regressions are presented. By using simplified equations and functions, the variables constituting supply and demand in the model are expressed. Initially, all the chosen variables are included in the fundamental regression and are dropped or modified continuously in order to improve the model. Acronyms for the variables used in this study are presented in the table (5.1) below. The table illustrates how each variable is related to one of the three components in the model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Capture</th>
<th>Acronym</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacancy</td>
<td>Demand/Supply</td>
<td>V</td>
</tr>
<tr>
<td>Office stock</td>
<td>Supply</td>
<td>OS</td>
</tr>
<tr>
<td>New Office Space</td>
<td>Supply</td>
<td>NOS</td>
</tr>
<tr>
<td>Unemployment</td>
<td>Demand</td>
<td>UE</td>
</tr>
<tr>
<td>Interest Rate 3 months</td>
<td>Economic Indicator</td>
<td>IR3M</td>
</tr>
<tr>
<td>Interest Rate 10 years</td>
<td>Economic Indicator</td>
<td>IR10Y</td>
</tr>
<tr>
<td>OMX30</td>
<td>Economic Indicator</td>
<td>OMX30</td>
</tr>
<tr>
<td>Consumer price index</td>
<td>Economic Indicator</td>
<td>CPI</td>
</tr>
<tr>
<td>Gross Domestic Product</td>
<td>Economic Indicator</td>
<td>GDP</td>
</tr>
</tbody>
</table>

Table 5.1: Independent variables with acronyms used in the equations.
In line with the theoretical framework in this study, the office market can be viewed as a dynamic model based on the pillars of a supply-demand approach. Vacancy (V) and number of unemployed people (UE) constitutes the demand-side. Therefore, demand (D) can be expressed as a function of V and UE

\[ f(D_t) = V_t + UE_t \]  \hspace{1cm} (2)

On the supply-side of the simplified model, supply (S) can be expressed as a function of office stock (OS) and new office space (NOS)

\[ f(S_t) = OS_t + NOS_t \]  \hspace{1cm} (3)

Economic indicators (EI) are the exogenous drive force in the three-component model. The segment for EI has several indicators and it can therefore be reasoned to leave some of the variables out since they plausibly have an equivalent capture. Emanating from economic theory, both Gross Domestic Product and OMX30 are measurements of general performance in Sweden. Based on this underlying statements and theory the data will be tested to see for correlation. Moreover, EI can be expressed as a function of following variables

\[ f(EI_t) = IR3M_t + IR10Y_t + OMX30_t + CPI_t + GDP_t \]  \hspace{1cm} (4)

Rosen (1984) and Heckman (1985) was among the first researchers to use a three-component model for their studies for rents in the office market. Following their work, a numerous amount of researchers have used similar triangulation model approach using different variables to explain the rent level in the office sector. Furthermore, the model used in this study follows the same approach.

The underlying theory in a function constructed with supply and demand supported by EI as explanatory exogenous features. There is a close linkage between D and the EI. Many of the economic variables can also be seen as contributors to D, since a strong economic market should be a driving factor for the demand.
Office rents \( y_t \) can therefore be written as a function of D, S and EI at time t.

\[
Y_t = \alpha_t + f(D + S + EI)_t + \varepsilon_t
\]  \hfill (5)

Lastly, we can rewrite the fundamental regression equation as following

\[
y_t = \alpha_t + \beta_1 V_t + \beta_2 OSt + \beta_3 NOS_t + \beta_4 UE_t + \beta_5 IR3M_t + \beta_6 IR10Y_t + \beta_7 OMX30_t + \beta_8 CPI_t \\
+ \beta_9 GDP_t + \varepsilon_t
\]  \hfill (6)

The objective of the fundamental regression is to find a model that reach a high explanation value of the model variance (adj.\( R^2 \)) and with statistically significant variables. To accomplish this, variables can be modified to a dynamic form. In this study, some variables are lagged and differentiated in order to reach statistical significance. The rent variable will also be used, in a lagged variable, as an explanatory factor. By doing this, the OLS regression becomes dynamic. The reasoning behind the lagging technique is that effect on \( y \) (office rents) can be lagging \( t4 \) periods due to the slow real estate market.

This can be particularly convenient in the context of real estate since that is a segment impartially deliberate moving. Furthermore, a variable lagged \( x \) periods, from time \( t \), can be written as

\[
LAG (x)GDP = GDP_{t-x}
\]  \hfill (7)

Differentiation can be used when it is believed that the change (\( \Delta \)) is more significant than the absolute value itself, given time \( t \). First difference is the delta value between two time periods (illustrated in equation 8). Moreover, variables can be used in second difference etc. if the correlation coefficients are more statistically significant.

\[
\Delta Y = \frac{Y_t}{Y_{t-1}}
\]  \hfill (8)

\footnote{Time period \( t \) = quarters}
5.2 Vector Autoregression (VAR)
Sims (1980) initially advocated VAR models as being alternatives by using multivariate simultaneous equations for macroeconomic analysis. At that time, longer and more frequent data called for a model which described the dynamic structure of variables. VAR models fitted for this purpose where they typically treat all variables as priori endogenous. Further, Lütkepohl (2011) underlines that VAR models are natural tools for forecasting. Simplifying the intuition behind the mathematical derivation of the model; “The idea underlying forecasting with a VAR model is first to summarize the dynamic correlation patterns among observed data series and then use it do predict likely future values for each series” (Robertson & Allman, p.7, 1999).

Generally, VAR models are constructed for stationary variables without time trends. The important feature of stochastic trends has shown to be captured by the model (Johansen, 1995 and Granger, 1981 among others). The advantage by using VAR level models compared to vector error correction\(^5\) (VEC) models is the fact that they can be used even though the cointegration structure is unknown (Lütkepohl, 2011). Moreover, the relatively small sample used in this study constitutes another reason for the use of VAR instead of VECM. The process scheme of a VAR analysis is described in Figure 5.6.

---

\(^5\) VEC models separate the effect of short run and long run movements (Lütkepohl, 2011).
A VAR can be described as a model where \( K \) variables is specified as linear functions of \( p \) of their own lags, \( p \) lags of the other \( K-1 \) variables and eventually additional exogenous variables (Stock & Watson, 2001). Algebraically, a \( p \)-order VAR model, written \( \text{VAR}(p) \) and with exogenous variables \( x_t \) can be expressed as,

\[
y_t = v + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B_0 x_t + B_1 x_{t-1} + \cdots + B_s x_{t-s} + u_t \quad t \in \{-\infty, \infty\} \tag{9}
\]

where

- \( y_t = (y_{1t}, \ldots, y_{Kt})' \) is a \( K \times 1 \) random vector,
- \( A_1 \) through \( A_p \) are \( K \times K \) matrices of parameters,
- \( x_t \) is an \( M \times 1 \) vector of exogenous variables,
- \( B_0 \) through \( B_s \) are \( K \times M \) matrices of coefficients,
- \( v \) is a \( K \times 1 \) vector of parameters, and
- \( u_t \) is assumed to white noise; that is,
  - \( E(u_t) = 0 \),
  - \( E(u_t u'_s) = \Sigma \), and
  - \( E(u_t u'_s) = 0 \) for \( t \neq s \)

Continuously, there are \( K^2 \times p + K \times (M(s + 1) + 1) \) in the equation for \( y_t \) and \( \{K \times (K + 1) \}/2 \) in the covariance matrix \( \Sigma \). Due to the fact that no priori assumptions are imposed, fitting a VAR model allows the data to speak for itself. However, some restrictions on the structure of \( \Sigma \) needs to be imposed in order to make causal interpretations of the results. Furthermore, by making some technical assumptions, we can derive another representation of the first equation (9). If the model states that VAR is stable, \( y_t \) can be expressed as

\[
y_t = \mu + \sum_{i=0}^{\infty} D_i x_{t-i} + \sum_{i=0}^{\infty} \Phi_i u_{t-i} \tag{10}
\]

where \( \mu \) the \( K \times 1 \) time-invariant mean of the process. \( D_i \) and \( \Phi_i \) are \( K \times M \) and \( K \times K \) matrices of parameters, separately. \( \Phi_i \) are the moving-average coefficient (also known as \( \text{IRF}^6 \)) and \( D_i \) are the dynamic-multiplier functions. In other words, equation (2) indicates that the process in which variables in \( y_t \) fluctuate about their time-invariant means, \( \mu \), if fully determined by the

---

*IRF = Impulsive response function*
parameters in $D_i$, $\Phi_i$, the (infinite) past history of the exogenous variables $x_t$ and the independent and identically distributed (i.i.d.) shocks.

In the econometric modelling, lag-selection criteria were used with the purpose of choosing the optimal lag length. There are several information criteria used in research today. However, research indicates that the selection of criteria is contingent with the number of observations; generally, if a data sample is under or above 60 observations (Liew, 2004). In the case of this study, Aikake Information Criteria (AIC) was primarily used due to the condition of the data sample.

5.3 Forecasting
In this section, the methodology constituting the forecasting part of this study will be presented. Both in-sample and out-of-sample forecasts will be conducted based on underlying VAR-model of each submarket. By employing model-based forecasts, the accuracy of point forecasts from the VAR-model can be statistically evaluated prior using the forecasts (Robertson & Allman, 1999). Due to the limitation of this study, we use an evaluation, by Bönner (2009), of how different econometric models fits different forecasting scenarios and parameters. Forecasting can be performed in various ways depending on preferred outcome. Due to the data set and forecasting preferences, the vector Autoregression (VAR) is found most useful. The summary is presented in Table 5.2 below.

<table>
<thead>
<tr>
<th>Assumptions</th>
<th>Regression</th>
<th>ARIMA</th>
<th>GARCH</th>
<th>VAR</th>
<th>VECM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicable with time series data</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Applicable with panel data</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Minimum of data observations</td>
<td>30</td>
<td>30</td>
<td>30+X</td>
<td>30+X</td>
<td>30+X</td>
</tr>
<tr>
<td>Consideration of cyclical motions</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Consideration exogenous variables</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>(Yes)</td>
<td>(Yes)</td>
</tr>
<tr>
<td>Consideration of economic theory</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Consideration of autocorrelation</td>
<td>(Yes)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Consideration of volatility cluster</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Short forecasting period preference</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Long forecasting period preference</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 5.2: Illustrates an evaluation of econometric models fitted for different forecast scenarios. Based on Bönner (p.32, 2009)
The forecasting will generate predicted values, one period (quarter) at the time, based on VAR estimates. By presupposing the VAR-equation (9), the mathematical derivation of the forecasting dynamics can be expressed using notation from Lütkepohl (2005)

\[ y_t(h) = \tilde{\theta} + \tilde{A}_1 y_t(h-1) + \cdots + \tilde{A}_p y_t(h-p) + \tilde{B} x_t \]  

(11)

Conditional on \( p \) initial values, the forecasting mechanism will dynamically predict variables in vector \( y_t \) of the endogenous variables and any exogenous variables \( x_t \). The latter equation estimates the optimal \( h \)-step-ahead (\( h= \) one quarter) forecast of of \( y_{t+h} \) conditional on exogenous variable \( x_t \). Moreover, when there are no exogenous variables (as in this study), the equation (11) can be rewritten as follows

\[ y_t(h) = \tilde{\theta} + \tilde{A}_1 y_t(h-1) + \cdots + \tilde{A}_p y_t(h-p) \]  

(12)

By excluding exogenous variables, asymptotic confidence bounds can be computed. Confidence bounds of 95% will be used in this study in line with theoretic praxis (Lütkepohl, 2005). Initially, in-sample forecasts will be executed. This type of forecasting predicts rents whose data is within the data sample i.e. include forecasting within the existing data set i.e. before 2016q4. In this study, in-sample will be executed 8 periods within the actual data (2014q1) and 16 periods ahead to 2018q1. This forecast methodology enables opportunities to evaluate performance and comparing of results.

Dynamic out-of-sample forecasting will be executed in each submarket. Out-of-sample implies that the forecasting period starts after the last period included in the data sample used in the VAR-modelling i.e. out-of-sample exclusively forecasts periods outside the existing data set i.e. after 2016q4. Further, 8 step-ahead (two years) forecasting in terms will be executed in the terms of the forecasting in each submarket. This is done in accordance with the general approach within the real estate field. Contracts is often between 1-3 years and believed to be shorter in the future, therefore two year forecasting was chosen.
5.4 Forecasting performance and test statistics

In order to measure the performance and accurateness that the VAR-model has in forecasting, standard statistical metrics can be executed. In previous literature (Diebold, 2001; Kirchgässner, 2007 and Wooldridge, 2009) there are numerous of examples that illustrates the performance of a certain forecast. In line with evaluation of in-sample forecasting, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used. These measures will answer if the forecasts are reasonable in general and hence whether the underlying model is correctly specified. Within this field of research, mentioned methods are widely used as statistical testing tools. Common for these measurements are the use of the error term (residuals) and the calculation of how narrow it is to the actual value. They differ in explaining the performance, RMSE use an absolute value while MAPE expresses it in percentage.

The equation for MAPE can be written as

\[ MAPE = 100 \sum_{t=1}^{n \to \infty} \left| \frac{\hat{y}_t - y_t}{y_t} \right| / n \]  

(13)

where \( \hat{y}_t \) is the forecasted value of the endogenous variable \( y_t \) (office rents) and \( n \) is the number of observations that stretches to infinity. Hence, \( \hat{y}_t - y_t \) is the forecasting error. Using the same notation as the equation (13) above, RMSE can be written as

\[ RMSE = \sqrt{\frac{\sum_{t=1}^{n \to \infty} (\hat{y}_t - y_t)^2}{n}} \]  

(14)

RMSE is performed by testing the average root square of the forecasted values in comparison to the best fitted line. While MAPE investigates the average percent that the forecast is from the actual observation. A critique for RMSE is that an underlying assumption that the errors is normally distributed and unbiased must be followed for the performance test to successfully operate.
In 1950, Durbin & Watson (1950) presented the statistical test for autocorrelation. Since then the Durbin-Watson test has been widely used for testing if the regression is auto-correlated and is used to test the OLS-regressions in this study. Further, this is done by predicting the error term from the regression and then test it. The test has a spread between zero and four, with a value of two signifying that there is no autocorrelation which often is preferable. With a result of zero, it indicates that the values is approaching a positive autocorrelation and the contrary for a result of four. Moreover, the equation for the d-statistic can be expressed as follows

\[ d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=2}^{T} e_t^2} \]  

The Dickey-Fuller test was developed by David Dickey and Wayne Fuller in 1979 and is used to test if time series are stationary. Stationarity means that the variable has a constant mean and variance over time and is used to test our variables in the initial OLS-regressions. Commonly a time series variable can show a non-stationarity i.e. random walk. To deal with the problem of random walk, first differentiation can be used. With first difference, the time series is transformed in the difference between the value for a given time \(y_t\) and the previous observation \(y_{t-1}\). This can be explained by the following equation \(\Delta y = y_t - y_{t-1}\) where \(y_t\) is integrated of order one i.e. I(1). If the null hypotheses that \(y_t = I(1)\) not can be rejected the solution is test a higher order of integration. To continue the regression of a time series, stationarity is compulsory.

In order to test for the direction of the causal relationship between variables, Granger causality tests are conducted. The method is a probabilistic account of causality; it uses empirical data sets to find patterns of correlation (Granger, 1969). In addition to test for causality, the direction of the linkage is tested. Which variable that is driving respectively following in this relationship. The underlying hypothesis can be expressed as if “x granger-cause y” and/or “y granger-cause x” (Leamer, 1985). Granger causality tests are conducted in this study in order to statistically secure the direction of the cause-effect on office rents.
6 Results and analysis

The results from the pooled regressions and forecasts are presented per submarket. Section 6.1 provides results from the pooled regressions i.e. the statistically significant variables. Further, out-of-sample and in-sample forecasts are presented in Section 6.2.

6.1 Pooled regression

The OLS dynamic regression output are presented in tables for each submarket. The variables have been lagged and differentiated in order to retrieve statistically significant estimates. 95% confidence interval has been used for all the regressions. Furthermore, regressions were executed accordingly to our office rent model i.e. a function of supply, demand and macroeconomic indicators. The office rent-variable, in all submarkets, showed signs of multicollinearity in level and lagged form. However, the office rent-variable showed no signs of stationarity when being first-differenced. Test statistics\(^7\) are not presented in table-format.

Stockholm CBD

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Stockholm CBD Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent</td>
<td>Adj. (R^2 = 96.4) Number of observations = 62</td>
</tr>
<tr>
<td>Independent variables</td>
<td>(\beta)</td>
</tr>
<tr>
<td>Rent (lag1)</td>
<td>0,95</td>
</tr>
<tr>
<td>Vacancy (lag2)</td>
<td>2873,49</td>
</tr>
<tr>
<td>Stock (lag2)</td>
<td>3,78</td>
</tr>
<tr>
<td>UE</td>
<td>-0,03</td>
</tr>
<tr>
<td>GDP (lag4)</td>
<td>-0,0009</td>
</tr>
</tbody>
</table>

Table 6.1: Illustrates regression results with independent variables for Stockholm CBD. * indicates statistical significance at 95% confidence level.

Stockholm CBD was problematic and the variables had to be modified to make a functional model. The final model includes lag rent, vacancy and stock both lagged two periods, unemployment and GDP lagged four periods. This model covers the demand, supply and Economic indicators, the lag rent is included as a variable that is believed to have strong influence on the overall explanation of rent level.

The test for Durbin-Watson showed a value of 1.69 which is acceptable. For the VIF-test, the result was 10.58 which is significantly higher than the acceptable limit of five meaning that the model has a problem with multicollinearity. However, it will not affect causal interpretations of the model.

\(^7\) Durbin-Watson d-statistic, Dickey-fuller test statistics and Variation inflation factor are available upon request.
**Stockholm RIC**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Stockholm RIC Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent</td>
<td>Adj. $R^2 = 0.88$ Number of observations = 60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>$\beta$</th>
<th>Std.Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent (lag1)</td>
<td>0.94</td>
<td>0.072</td>
<td>13.02</td>
<td>0.000*</td>
</tr>
<tr>
<td>Vacancy (lag6)</td>
<td>692.47</td>
<td>339.98</td>
<td>2.04</td>
<td>0.046*</td>
</tr>
<tr>
<td>Stock</td>
<td>0.39</td>
<td>0.125</td>
<td>3.16</td>
<td>0.003*</td>
</tr>
<tr>
<td>OMX30 ($\Delta$)</td>
<td>-0.25</td>
<td>0.155</td>
<td>-1.65</td>
<td>0.105**</td>
</tr>
</tbody>
</table>

Table 6.2: Illustrates regression results with independent variables for Stockholm RIC. * indicates statistical significance at 95% confidence level and ** for 90% confidence level.

For the Stockholm RIC, the model encountered some problems. The variables Newspace and Rate3M could be dropped after the correlation control. OMX30 was chosen over GDP as economic indicators by significance reasons. Vacancy is a variable that is highly regarded and has therefore been included despite six lags was needed to get it significant. The regression that showed the best result had lag rent, Stock, Vacancy with six lags and OMX30 in first difference. The general approach has been to only include variables at a 95% significance level but for Stockholm RIC, OMX30 with first difference was still included, despite a p-value of 0.105. The rest of the variables reached the significance target and the adj. $R^2$ was 0.88. Both the Durbin-Watson and the VIF-test had pleasant result 2.23 respectively 1.62.

**Gothenburg CBD**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Gothenburg CBD Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent</td>
<td>Adj. $R^2 = 0.96$ Number of observations = 64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>$\beta$</th>
<th>Std.Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent (lag1)</td>
<td>0.90</td>
<td>0.056</td>
<td>16.08</td>
<td>0.000*</td>
</tr>
<tr>
<td>Stock (lag2)</td>
<td>-1.28</td>
<td>0.650</td>
<td>-1.97</td>
<td>0.054**</td>
</tr>
<tr>
<td>Rate10Y</td>
<td>-27.60</td>
<td>11.761</td>
<td>-2.35</td>
<td>0.022*</td>
</tr>
<tr>
<td>OMX30</td>
<td>0.08</td>
<td>0.036</td>
<td>2.23</td>
<td>0.029*</td>
</tr>
<tr>
<td>UE (lag2)</td>
<td>-0.006</td>
<td>0.003</td>
<td>-2.08</td>
<td>0.018*</td>
</tr>
</tbody>
</table>

Table 6.3: Illustrates regression results with independent variables for Gothenburg CBD. * indicates statistical significance at 95% confidence level and ** for 90% confidence level.
The regression output in the table above are considered as the variables that affect office rents in Gothenburg CBD. The Newspace-variable from the original equation were dropped due to low correlation. Notably is that the vacancy-variable were excluded due to low statistical significance (p-values) even when supposed to different lags and differentiations. Stock and unemployment were lagged two periods. Both the 10-year bond and OMX30-index were used as macroeconomic indicators to explain office rent in Gothenburg CBD that presents an Adj. $R^2$ of 0.96. Moreover, the Durbin-Watson d-statistic showed 2.01, which indicates no autocorrelation. The VIF-estimate showed 4.87 indicating that the variables are moderately correlated.

**Gothenburg RIC**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Gothenburg RIC Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent</td>
<td>$Adj. R^2 = 0.94$ Number of observations = 62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>$\beta$</th>
<th>Std.Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent (lag1)</td>
<td>0.88</td>
<td>0.050</td>
<td>17.69</td>
<td>0.000*</td>
</tr>
<tr>
<td>Vacancy (lag4)</td>
<td>-775.90</td>
<td>268.35</td>
<td>-2.89</td>
<td>0.005*</td>
</tr>
<tr>
<td>OMX30</td>
<td>0.107</td>
<td>0.039</td>
<td>2.82</td>
<td>0.007*</td>
</tr>
</tbody>
</table>

*Table 6.4:* Illustrates regression results with independent variables for Gothenburg CBD. * indicates statistical significance at 95% confidence level.

The explanatory variables of this simple regression equation consist of lagged office rent, vacancy (lagged 4 periods) and OMX30 i.e. the submarket of Gothenburg RIC is a function of only three variables. The rest of the variables were excluded due to low statistical significance. Noticeable is that a highly lagged vacancy tends to increase the explanation power of this submarket. Approximately 94% of the variance is explained in this pooled regression. Furthermore, the d-statistic show an estimate of 2.16 in which indicates no autocorrelation. VIF is estimated to 1.87 in which can be interpreted to lowly moderately correlated variables.
Malmö CBD

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Malmö CBD Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent</td>
<td>Adj. R² = 0.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>β</th>
<th>Std.Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent (lag1)</td>
<td>0.778</td>
<td>0.620</td>
<td>12.54</td>
<td>0.000*</td>
</tr>
<tr>
<td>Vacancy (lag4)</td>
<td>520.690</td>
<td>244.290</td>
<td>2.13</td>
<td>0.038*</td>
</tr>
<tr>
<td>Stock (lag6)</td>
<td>0.540</td>
<td>0.212</td>
<td>2.54</td>
<td>0.014*</td>
</tr>
<tr>
<td>UE (lag2)</td>
<td>-0.008</td>
<td>0.004</td>
<td>-2.08</td>
<td>0.042*</td>
</tr>
<tr>
<td>GDP</td>
<td>0.002</td>
<td>0.000</td>
<td>2.70</td>
<td>0.009*</td>
</tr>
</tbody>
</table>

Table 6.5: Illustrates regression results with independent variables for Gothenburg CBD. * indicates statistical significance at 95% confidence level.

Generally, explanatory variables in this submarket have a lagged effect on office rents. Noticable is that office supply was lagged six periods and vacancy four periods. Further, the result of the pooled regression showed that unemployment tends to effect the rent two periods after in our data sample for Malmö CBD. GDP were used as an external macroeconomic indicator of this market. Continuously, d-statistic was estimated 2.08 indicating no autocorrelation of the time series together with a significant VIF-estimate of 3.48 indicating no multicollinearity.
6.2 Forecasting
The forecast result of this study is presented in this section. The forecasted office rents (both out-of-sample and in-sample) are presented in graphs per submarket with SEK/year on the y-axis and time on the x-axis. 95% uncertainty bands (interval) were used in all out-of-sample forecasts. Some in-sample forecasts are presented starting from 2013q1 in order to strengthen the validity of the forecasts.

Stockholm CBD

It is from this VAR model that a forecast can evolve. Firstly, an out-of-sample forecast is visualizing how the model predict the rent level to move from the last observation and 2 years forward (eight steps). The result shows that the rent level keep increasing and reach above 8000 SEK in two years’ time. The increase is of step wise progress. It seems to be a rise in rent and then for some period be stable at that level and then rise again, make the line look like a upward stairway. The confidence interval after two years is substantially large, around 6000 SEK. Despite this, the confidence interval indicates that the rent level will not be lower than it is today. The most likely scenario is an escalation and the opposite is unlikely to occur. Moreover, all the included variables significantly granger-causes office rents.

Figure 6.1: Illustrates diagrammatic 8 period out-of-sample forecast for Stockholm CBD. The grey area indicates all possible outcomes within 95% confidence interval.
For the second part of the forecast from the VAR model, an in-sample forecast is performed. A total of 16 periods is predicted, eight in-sample and then eight out-of-sample. The in-sample forecast is then compared to the actual observed data. The forecasting starts from 2015q1 and the first year is a successful prediction but 2016 the observation has a higher level of rent than the prediction. The forecast line appears to be volatile; with significant up and down movements, the overall trend is still growing. The MAPE estimates showed a result of 6.9% error over the viewed 8 periods within the in-sample forecast.

**Stockholm RIC**

**Figure 6.2:** Illustrates diagrammatic 16 period in-sample forecast for Stockholm CBD. Observed data are dotted.

**Figure 6.3:** Illustrates diagrammatic 8 period out-of-sample forecast for Stockholm RIC. The grey area indicates all possible outcomes within 95% confidence interval.
From the pooled regression, a VAR model were constructed, see Table 6.2. The lag length was set to one. The forecast for the Stockholm RIC submarket can be seen below. The prediction line indicates a stable increase in rent level over the coming two years, approximately a 200 SEK increase. A conclusion from the illustration is that the confidence interval is quite broad. The entire Stockholm market has been volatile during the viewed period, making it harder to forecast. Granger causality test shows that all the included variables granger-causes the office rent variable in this submarket.

From the in-sample data, the problem with volatile data is clearly visible. The forecasted data is a linear predicts while the actual observed data has a heavily step wise progress. The forecast seems to have underestimated the pace that the rent level increases in this case. Looking at the increase from 2016q2 to 2016q3, almost 400 SEK. The model is not able to foresee such a sufficient upturn. For the overall percentage error, MAPE, the result for Stockholm RIC was 3.6% during the viewed period.

Figure 6.4: Illustrates diagrammatic 16 period in-sample forecast for Stockholm RIC. Observed data are dotted.
Gothenburg CBD

Figure 6.5: Illustrates diagrammatic 8 period out-of-sample forecast for Gothenburg CBD. The grey area indicates all possible outcomes within 95% confidence interval.

The lag-selection criteria of underlying VAR-regression indicated a lag-order of one for this submarket. The forecast estimates are in line with the historical trend for Gothenburg CBD were the office rents have increased with approximately 100 SEK/year. Coherent with Gothenburg RIC, the forecast show a steady non-volatile increase two years ahead. However, this concentrated market is faster growing given that the slope is steeper. In the end of the forecast-horizon, there is a difference of 500 SEK within the 95% uncertainty bands. Furthermore, included variables granger-cause office rent to a probability of 94% (p-value: 0.059).

Figure 6.6 and 6.7: Illustrates diagrammatic 16 period in-sample forecasts starting from 2015q1 and 2013q1 for Gothenburg CBD. Observed data are dotted.
The in-sample forecast results for Figure 6.7 indicate a movement of the same direction as Figure 6.5 and 6.6 show. Notable is that the observed data (quarterly quoted) logically increases step-wise while the forecast-estimates increase linearly. The stepwise increase of the actual data can be fitted into a linear line over time consistent with the forecasting estimates indicating precise results. The 8 periods between 2015q1 and 2017q1 of the in-sample forecast for the Gothenburg CBD displayed a percentage error (MAPE) of 2.8%.

**Gothenburg RIC**

![Out-of-sample forecast - 2 years](image)

**Figure 6.8**: Illustrates diagrammatic 8 period out-of-sample forecast for Gothenburg RIC. The grey area indicates all possible outcomes within 95% confidence interval.

The lag-selection criteria’s in the underlying VAR-model indicated a lag of one. The forecasted estimates are in line with the historical development with slowly increasing non-volatile office rents that has been over the latest 17 years. The uncertainty bounds are narrow compared to the other submarkets. The in-sample forecast follows the observed data over time, however, in a more linear manner. Lastly, Granger-causality tests show that independent variables significantly granger-cause office rents in Gothenburg RIC. The calculated MAPE indicated the overall percentage error to be 4.7% for the Gothenburg RIC market.
AIC lag-selection criteria indicates a lag of four periods for the underlying VAR-model. The out-of-sample forecast shows a volatile increase in rents over two years. In line with Gothenburg submarkets, an increase of approximately 100 SEK/year over the forecast horizon. Similar to Stockholm CBD the forecasting trend of Malmö CBD shows a non-linear development over the periods due to the historical trend. The confidence bounds are narrow the initial four periods indicating a narrow and precise forecast. In line with previous submarkets, it can be concluded that the explanatory variables granger-cause office rents in Malmö CBD.
The forecast line follows the observed data under the forecast horizon. Between the periods of 2015q3 and 2016q2, forecast estimates tangent observed data indicating a precise forecast. The observed data increases in a step-wise manner compared to the linear increase the forecast model generated. Perceptibly, the forecast line develops in a volatile fashion the period out-of-sample (after 2016q4). During the forecast period of two years, there were a maximum difference of 50 SEK/year between the predicted values and observed data. Subsequently the MAPE for the Malmö CBD area revealed a percentage error of 1%.

**Figure 6.11**: Illustrates diagrammatic 16 period in-sample forecast for Malmö CBD; 8 periods in-sample and 8 periods out-of-sample. Observed data are dotted.

**Figure 6.12**: Illustrates diagrammatic 16 period in-sample forecast for Malmö CBD. Observed data are dotted.
7 Concluding remarks
The future development of the Swedish office market constitutes an interesting study object. In this study, we have mapped out office rent drivers for each submarket. Logically, the variable of lagged rent had the highest explanatory power. In line with previous studies (Bönner, 2009), office stock and vacancy showed to be explanatory drivers of office rent for the clear majority of the submarkets. However, the effect of these variables are lagged and do not impact on office rents in the same period. This empirically strengthens the theory of a lagged real estate-cycle and the obtained dataset.

Considering the macro-variables in our regression-model, OMX30 was shown to be the most explanatory variable that had a causal impact on rental levels. In contradiction to previous studies, interest rate did not have any causal relationship with office rent in our model. It can be found interesting since office stock and interest rate were highly correlated in all submarkets. Moreover, one can argue that interest rates have an indirect impact on office rents (through the supply-side) during this relatively low interest rate period.

The in-sample and out-of-sample forecasts, based on our VAR-model, generated significant and valid forecast estimates for the three CBD-submarkets (particularly Gothenburg and Malmö). On the contrary, the forecast estimates for the two RIC-submarkets were not as precise. It can be argued that the geographical delimitation of the latter submarkets are too large which makes space for a larger spectrum of heterogeneity and location effects. The out-of-sample forecasts indicated that Stockholm CBD and Malmö CBD have a more volatile development than the rest of the submarkets. Furthermore, the use of office employees as a demand-variable would strengthen the office rent model. Instead, number of employees was used as a proxy in this study due to the lack of availability.

An interesting result is that the observed data is moving “step-wise” while the forecasts are moving in a more linear manner. Nevertheless, the slope of the forecasts follows the actual change in office rents. One can argue, consistent with Englund et.al. (2005), that the “step-wise” development of the office rentals are generated by new contract-signings as a result of an illiquid office market. Therefore, there are arguments against the office rental pricing during periods since long rental contracts do not consider the change in market price (only marked against CPI) during the contract horizon. Hence, our results suggest that office rental price can be wrongly priced during longer periods.
7.1 Suggestions for future research
Future researchers interested in forecasting office rents in Sweden could continue to develop our analysis even further. An interesting study object would be to use average (arithmetic mean) office rent instead and the use of office employees instead of the used proxy of number of employees. Further, the dataset could be modified and improved if data availability is not an obstacle. More frequently quoted data generates better predictions and allows researchers to conduct a similar study with more sophisticated econometric models such as VECM and ARIMA. Moreover, it would be interesting to conduct similar study with panel data in order to make market generalizations. Alternatively, apply the model on the residential and retail market.
References


43


Appendix – Econometric dictionary

Adjusted $R^2$ - The $R^2$ indicates how well the observed data points fits the line, also known as Coefficient of determination. When adjusted it accounts for the number of terms or variables in the model.

ARIMA - AutoRegressive Integrated Moving Average is a univariate approach that explains development of one variables through the historic data of the same variable. The basic assumption is that previous pattern will continue in the future.

Autocorrelation – Is the serial correlation of the error for one period and is transferred and correlated the succeeding period.

Cointegration – First, the used series must be integrated of order 1. Further if these non-stationary variables is in a linear combination seen as integrated of order zero they are seen as cointegrated. This also means that there is a long-run stochastic trend between the variables.

Confidence Interval – Is a statistic tool to measure the probability that a data point will be within this interval. The interval usually has a percentage, 95% or 99% to guarantee that the data will with 95% certainty be between the upper and lower bound.

Endogenous - An endogenous change comes from within the model. An endogenous variable is a variable that can be explained by the other variables that is included in the model.

Exogenous - An exogenous change comes from outside the model and is unexplained by the model. An exogenous variable is independent of the random error term in the linear model.

GARCH - Generalized AutoRegressive Conditional Heteroscedasticity. Generally, a GARCH model involves three steps; (1) estimation of a best-fitting autoregressive model, (2) computing autocorrelations of the error term, (3) testing for significance.

Multicollinearity – when two or more explanatory variables in a multiple regression model are highly correlated indicating that one can be linearly predicted from the others.

OLS – Ordinary Least Squares. A regression method that finds the line of best fit for a dataset that generates a visual illustration of the relationship between the data points.

Pooled regression – Pooled regressions are most suitable when each observation is independent from each other. Starting the analysis by including all variables in the data set.

Stationarity - A data set (time series) has stationarity if a shift in time do not cause a change in the shape of the distribution.

Time series - A time series is a series of data points indexed or listed in time order.

VECM – Vector Error Correction Model. A multiple time series model particularly used for data sets where the underlying variables have a long-run stochastic trend (cointegration). The model estimates short-term and long-term effects of one time series on another.