AN APPROACH TO MODELING AND FORECASTING REAL ESTATE RESIDENTIAL PROPERTY MARKET

A Thesis Submitted for the Degree of

Doctor of Philosophy

by

HAMED AHMED AL- MARWANI

Brunel Business School

Brunel University

November 9, 2014

Acknowledgement

I am thankful to my research advisor, Dr.Tillal Eldabi, Senior Lecturer at Brunel Business School, Brunel University, UK and my supervisor Dr.Ahmad Gharaibeh. Dr. Tillal has been helpful in providing advice many times during all these years. Both of them remain my best role model for a scientist, mentor, and teacher.

I would also like to thank all the staffs of the Brunel University who have rendered their valuable support during all these years.

Abstract

This thesis aims to provide an approach to real estate residential modeling and forecasting covering property types' correlation, time series attributes within a region or a city, and socio-economic attributes of preferred real estate locations. The thesis covers residential estate markets and concentrates on property types, while previous studies that have considered country wide house price indices. There is a gap identified in the literature in the need to study correlations between property types within a region or a city and whether they will provide diversification benefits for real estate investors such as risk reduction per unit of returns. The thesis concentrates on property type seasonality in addition to modeling time series attributes within a region or city instead of real estate index seasonality. This thesis the first to combine modern information systems techniques such as geographic information systems (GIS) with socio-economic factors to help understanding causal relationships that can be used to forecast real estate prices. The results show that it is more achievable to forecast real estate prices within a city than for the real estate market of the entire country. The GIS and socio-economic modeling results show that higher property prices are awarded to real estate with more green spaces, residents with higher disposable incomes, lower council tax bands, fewer tax benefits claimants, and better health services. Previous studies have examined real estate price indices at the macro level (the general, all real estate house price indices). There has not been a study that examines real estate price forecasts by property types within a city. The contribution of this thesis is its focus on time series analysis as well as causal modeling within a city with the objective of providing a better understanding of the dynamics of real estate price changes.

Glossary of Terms

Terms	Description
MPT	Modern Portfolio Theory
NAREIT	National Association of Real Estate Investment Trust
NCREIF	National Council of Real Estate Investment Fiduciaries
ARCH	Autoregressive Conditional Heteroscedasticity
REIT	Real Estate Investment Trust
VAR	Vector Autoregressive Regression
GIS	Geographic Information System
NBS	Nationwide Building Society
ESRI	Environmental Systems Research Institute
ONS	Office for National Statistics
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
PIA	Property Industry Alliance
DCF	Discounted Cash Flow
ARIMA	Auto Regressive Integrated Moving Average
PRF	Population Regression Function
DGP	Date Generating Process
MSOA	Middle Super Output Area
LSOA	Lower Super Output Area
NOMIS	Official Labour Market Statistics
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
P/E	Price over Earning
NOI	Net Over Income

FTSE	Financial Times and the London Stock Exchange

TABLE OF CONTENTS

CHAPTER 4: REAL ESTATE TIME SERIES MODEL BUILDING AND TESTING

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

Chapter 1: Introduction

1.0 Background

Real estate forecasting has become an integral part of the larger process of business planning and strategic management real estate sector. Investors in real estate such as pension funds, insurance companies and real estate funds will make allocations, which they may maintain for some time, are looking to forecast future asset performance across real estate sectors and locations. Investors need to know where rents, yields and prices will be within a predetermined time frame. Investors with loan portfolios, including those secured on real estate, demand objectivity and transparency in the analysis of future trends. Scenario analysis concerning future rent and cash flows is also important. Real estate forecasting is the natural progression in real estate as more data become available for a larger number of markets especially with the extra details provided by Geographic Information Systems (GIS) on relevant real estate valuations factors such as locations, transportation, and infrastructure.

Worthington and Higgs, 2003 studied the correlation between UK regional property markets from the period 1976-2001 by using the quarterly property indices of Nationwide Building Society. They suggested that reduction in risk was limited because assets in a portfolio showed high positive correlations. Worthington and Higgs, 2003 outlined the fact that using the concept of diversification in portfolio theory, accurate forecasts of regional prices can be made. They suggested that forecasting techniques in regional property markets should consider both exogenous and endogenous variables. Worthington and Higgs, 2003 study did not consider the correlation or the effect of diversification within the property types that are found within a region or a city.

Mehmedović *et.al*, 2010 tried to identify what factors govern the first time buyers' investment decision. Their work was mostly related to the qualitative aspects of decision making. In their conclusion they highlighted the need for additional quantitative analysis methods such as correlation and regression analysis.

Brown, 1997 suggested that investment in real estate is a heterogeneous type of investment. Although market factors are important for the ups and downs in the market, forecasting returns are also dependent on many other factors such as type of tenant, age of property, location, socio-economic parameters and the neighborhood regarded status. Although Brown, 1997 considered the importance of socio-economic factors, he did not address these factors in detail or how they can be used in a quantitative manner.

Wilson and Okunev, 2001 studied the combination of forecasts for property markets in the USA, UK and Australia. They found out that pooling the forecast resulted in better out of sample results. Similar results were obtained by Diebold, 1989 who found that pooling forecasts from different models tends to yield better outcomes for shorter term forecasts than for component models.

Vishwakarma, 2013 studied the Canadian real estate price index using the Autoregressive Integrated Moving Average (ARIMA) family of models. He compared his results with more econometric advanced models than ARIMA, such as vector autoregression, co-integration and error corrections models. He found that ARIMA models are more efficient than more complex models. Crawford and Fratantoni, 2003 used ARIMA, GARCH, and regime switching models to forecast the real estate market in various parts of the US. They used state-level repeat transactions data for California, Florida, Massachusetts, Ohio, and Texas. The study found that ARIMA models are generally more suitable for out-of-sample point forecasts. Stevenson, 2007 applied the Ordinary Least Squares (OLS), ARIMA, and Vector Autoregressive (VAR) models for forecasting housing supply in the Irish market, using quarterly data from 1978 through 2003. He found that the ARIMA model had better forecasting ability than the others for the period 1998–2001.

The above studies have concentrated mainly on univariate time series. Another approach to real estate forecasting is to consider causal relationships. Feldstein, 1992 and Kearl, 1979 studied the relationship between inflation and house demand. They found that higher inflation leads to decreased demand for housing. However, Quan and Titman, 1999 obtained results indicating that inflation may increase demand for housing because housing is seen as a hedge against inflation. Paul and McNamara, 1997 found that forecasts in rental growth are important in forecasting prices in real estate market. Breedon and Joyce, 1993 analyzed the relationship between real estate prices, earnings, disposable income, demographic factors, the rate of repossessions by lenders, and dwellings stocks. Using quarterly data, they concluded that real estate prices are directly linked to real personal disposable income and demography. They also concluded that investment in residential units is also linked to mortgage lending with real estate prices rising when mortgages become easier to obtain.

Cheng and Han, 2013 used population forecasting in the city of Shanghai, China to explain price change in real estate. They found that the effect of long-term population growth on real estate prices is limited. They identified a positive correlation between real estate requirements and population volatility. Miller and Sklarz, 1986 concluded that forecasting in housing market analysis is both an art and a science. The art consists of reliable economic factors which capture the true behaviour of the market, and the science of sound statistical modeling approaches. Sklarz, 1986 reported that medium to longer term real estate price drivers were employment, income, supply constraints and interest rates. The study concluded that accuracy of forecasting depends on choosing the right market indicators apart from house price.

Miller and Sklarz, 1986 studied price seasonality for the house price index price changes for each year between 1987-2007, and found that the real rates of return during spring and summer months are not the same and there by mentioned that there existed a seasonality in house price.

There are two important facts with respect to the movement of house prices, first is that house price changes are highly persistent from one period to the next (Case and Shiller, 1990; Meese and Wallace, 1991; Glaeser and Gyourko, 2006), and the second is that the housing market is prone to high rate of return periods followed by extremely low or negative return periods (Muellbauer and Murphy, 1997; Glaeser and Gyourko, 2006). Malpezzi, 1999 and Gallin, 2006 suggested that short to medium term forecasts can suffer from boom and bust, seasonality and short term volatility while in the long term there should be an equilibrium relationship between the fundamentals of real estate and its value.

The novelty of the forecasting model presented in the thesis is in its integrated approach to forecasting real estate price changes. The approach combines established time series techniques such as the autoregressive integrated moving average models, causal explanatory models using socio-economic data with geographic information system (GIS). The thesis is the first to provide a framework that ties social and economic variables to the price of real estate property types within a city. There has not been a study that used GIS to map causal relationships between real estate prices and explanatory variables such as higher income, council tax claimants, good health and green space. The thesis contributes to the existing literature by providing a comprehensive framework for the study of the changes in real estate prices. The approach is comprehensive in the sense that, it provides flexibility with the use of available data; since there is always an alternative methodology in case if data of certain model is not available, for example; if time series data is not available then, causal models and GIS can be used and vice versa.

The focus of the thesis is on forecasting real estate prices within the short to medium terms. Therefore, ARIMA and GARCH modeling as in Vishwakarma, 2013 and Fratantoni, 2003; in addition to causal regression models will be employed as in Feldstein, 1992; Kearl, 1979 and Worthington and Higgs, 2003. The thesis concentrates on a sector of the UK real estate market that has been relatively under researched despite its large size; the residential sector. Most of the research that has been carried out in the field of real estate has concentrated on commercial and industrial property where data have been collected from the FTSE UK commercial property index or other such indices which deal with office, retail and industrial properties. The Property Industry Alliance (PIA) Data Report 2012 states that the residential sector is far larger than any other asset class in the UK, worth 4,224 billion GBP. The report states that institutional investors and REITs managers only have a very small exposure of 0.54% (2.3 billion GBP) of residential real estate UK market. This is in contrast to 35% (292 out of an 820 billion GBP commercial real estate market) domination of the commercial real estate sector by institutional investors and REITs managers. The British Property Federation (2012) report states that by 2026, there will be hike of at least 26m residential units. These units give a better financial output in comparison to other investment assets. Over the last thirty years, it has been proved that residential units have not only been the best investment performing asset but also provide higher returns and less risk. Unlike stocks and bonds that have received the lion's share of research on price change predictability, less research has gone into forecasting residential real estate prices.

1.1 Research Problem

The concise literature review conducted in the previous section indicates three areas for research that need to be explored in more detail.

1.1.1 Problem 1: Lack of studies of real estate price determination by residential property types within a city.

A gap has clearly been identified in the literature review that is it is necessary to study correlations between property types within a region or city and ascertain whether they will provide diversification benefits for real estate investors such as risk reduction per unit of returns. Studies have considered the UK regional house price index of Nationwide Building Society and correlations within properties within regions. Worthington and Higgs, 2003 suggested that forecasting techniques in regional property markets should consider both exogenous and endogenous variables. Worthington and Higgs, 2003 study did not consider the correlation or the effect of diversification within the property types which are found within a region or a city. Mehmedović *et.al*, 2010 highlighted the need for additional quantitative analysis methods such as correlation and regression analysis. The main objective of investments is to maximize return given a level of risk, or minimize risk for a given level of return. This objective is essentially related to the correlation structure between individual investments within a portfolio. Low or negative correlations will lead to more diversification benefits as measured by lower risk for a given rate of return.

1.1.2 Problem 2: Lack of studies of seasonality of real estate prices by property types within a city.

There is a gap identified in the literature for work that concentrates on property type seasonality in addition to modeling its time series attributes within a region or a city instead of real estate index seasonality. A reason for this is that real estate buyers are location specific, and are therefore by definition specifically interested in locations. Previous studies have only considered seasonality within the all house price index. Miller and Sklarz, 1986 studied price seasonality for the house price index price changes within each year during 1987–2007 and identified the real rates of return during spring and summer months. Case and Shiller, 1990; Meese and Wallace, 1991; Glaeser and Gyourko, 2006 conclude that house price changes are highly persistent from one period to the next (Muellbauer and Murphy, 1997); Glaeser and Gyourko, 2006)). Malpezzi, 1999 and Gallin, 2006 found that the housing market is prone to high rate of return periods followed by extremely low or negative return periods. There is a difference between the interests of real estate investors who are buying many units of real estate with the objective of diversification as an objective, and individual real estate buyers who do not have the ability to diversify as real estate investors due to their limited budget and ability to borrow.

1.1.3 Problem 3: Lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting.

A clear gap has been identified in the literature that there is a need for a study connecting GIS maps and socio economic parameters that can help in identifying the most important socio-economic variables driving house prices. There has been no study combining modern information systems techniques such as geographic information systems (GIS) attributes with socio–economic factors to help in understanding causal relationships that can be used to forecast real estate prices. Socio-economic factors are the forces driving demand and supply for real estate. These forces include but are not limited to demography, education, income, inflation etc. GIS helps to map these forces to specific real estate area facilitating investment decision making. Feldstein, 1992 and Kearl, 1979 studied the relationship between inflation and house demand and found that higher inflation leads to decreased demand for housing. However, Quan and Titman, 1999 achieved different results, indicating that inflation may increase demand for housing because it is seen as a hedge against inflation. Paul and McNamara, 1997 found that forecasts in rental growth are important in forecasting prices in the real estate market. Breedon and Joyce, 1993 analyzed the relationship between real estate prices, earnings, disposable income, demographic factors, the rate of repossessions by lenders, and the dwellings stocks. Cheng and Han, 2013 used population forecasting in the city of Shanghai, China to explain price changes in real estate. They found that the effect of long-term population growth on real estate prices is limited. They found a positive correlation between real estate requirements and population volatility. GIS maps can identify and prioritize the locations that historically have been able to command the highest real estate prices in addition to the socio economic factors and attributes of these locations.

1.2 Research Aim and Objectives

1.2.1 Aim:

From the research problem as discussed in 1.1, the aim of the thesis is to provide an approach to real estate residential modeling and forecasting covering a property type's correlation and time series attributes within a region or a city instead of looking at only real estate index seasonality, and that can identify and prioritize locations in addition to the socio-economic attributes of these locations.

The background in section 1.0 provides justification that most of the research done on real estate markets to date has concentrated on aggregate indices and correlations between regional property assets. The research also shows that the residential real estate market is less studied compared to commercial real estate despite figures showing huge potential growth in the residential real estate market. This thesis covers residential real estate markets of different types with regard to time varying parameters, causal relationships and correlations between property types.

1.2.2 Research Objectives

The content of the research problem and the aim of this study have been broken down into the three objectives of the research, which are as follows:

Objective 1:

Investigate the time series properties of different property types (Flats, Terraced, Semi-Detached, and Detached) and compare their time series characteristics with each other as well as the UK all house price index.

This objective aims at building appropriate statistical models based on time varying parameters that can help in improving real estate forecasts within a city and by property types beyond simple naïve or seasonal models. Box-Jenkins ARIMA (Autoregressive Integrated Moving Average) approach to model identification, estimation and forecasting real estate prices is employed in chapter four. ARIMA models are found to be able to indicate short-term market direction. ARIMA models also do well when compared to other model classes as it will be discussed in details in chapter two.

Objective 2:

Introduce a methodology that combines published data from the UK office for national statistics with the geographic maps (GIS) to identify causal relationships that can be used in forecasting real estate prices.

This objective helps with building causal models that can provide insight into cross sectional model building, estimation and forecasting. The methodology used for model validation and hypothesis testing is based on multiple regression technique.

Objective 3:

To provide a new approach utilizing time series techniques, causal models and GIS tool to model and forecast real estate price changes.

The performance of time series models as identified in objective one will be compared with the performance of causal models with regard to forecasting accuracy of real estate prices.

This objective tests the forecasting accuracy of time series models and causal models on out of sample data. This objective is important for real estate managers and REITs mangers in building their forecast for future performance of real estate investment in terms of return/risk trade off.

1.3 Research Plan

The thesis proceeds by reviewing the literature in chapter two. Chapter three introduces the time series and causal models that are going to be combined and used in chapters four and five. Chapter five introduces GIS and shows how it can help in providing deeper understanding of the forces that cause changes in real estate prices. Chapter five also studies causal relationship models and compares its forecasting ability with the time series models built in chapter four. Chapter six concludes the thesis, provides recommendations and suggests areas for future research.

1.4 Organization of the thesis

Chapter 1 presents aims, objectives and research plans.
Chapter 2 outlines a review of the existing literature with a final focus on the research gap.

Chapter two will provide a review of real estate research The Chapter reviews the discounted cash flow and price to earnings methods of evaluating real estates. The socio-economic factors that shape the decision making process of real estate buyers will be discussed and connected with proper locations for real estate development using the geographic information system (GIS). The Chapter will then determine the research gap that the thesis intends to bridge.

Chapter 3 deals with research methodology and statistical models in the area of real estate forecasting.
The models used are time series models, cross sectional multiple regression models, GIS tools. The six steps methodology will be introduced and discussed in Chapter three. The steps are statement of research problem, gathering relevant data, estimation methods, diagnostic tests for the residuals, model evaluation, and out of sample forecasting.

Chapter 4 investigates and builds appropriate time series models for real estate prices in the UK.
This chapter deals with thesis research problems one and two which are lack of studies of real estate

price determination by property types within a city and lack of studies of seasonality of real estate prices by property types within a city. The chapter deals with objective one which is to *investigate the time series properties of different property types (Flats, Terraced, Semi-Detached, and Detached) and compare their time series characteristics with each other as well as the UK all house price index.* This objective aims at building appropriate statistical models based on time varying parameters that can help in improving real estate forecasts beyond that of simple naïve or seasonal models. The contribution of this chapter is in applying the methodology and models by property types and within a city in contrast with previous research studies that examined market indices.

Chapter 5 estimates causal relationships between real estate prices and potential explanatory variables such as disposable income, council tax, and employment levels. The chapter deals with the thesis research problem number three which is lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting. The chapter provides a new model to the literature by combining GIS tools with multiple regression techniques.

Chapter 6 presents the Conclusions, Contributions, Limitations and Recommendations. This chapter will present findings, novelty with respect to forecasting and managerial relevance.

Chapter 2: Literature Review

2.0 Introduction

Chapter one identified the three objectives to be covered in the thesis. The first objective is the gap identified in the need to study correlations between property types within a region or a city and whether they will provide diversification benefits for real estate investors such as risk reduction per unit of returns. The main objective of investments is to maximize return given a level of risk or minimize risk for a give level of return. This objective is primarily related to the correlation between individual investments within a portfolio. Low or negative correlations will lead to more diversification benefits as measured by lower risk for a given rate of return. The second objective is to study a property type's seasonality in addition to modeling its time series attributes within a region or a city instead of real estate index seasonality. The reason for this is that real estate buyers are location specific and are therefore by definition interested in specific locations. There is a difference between the interests of real estate investors who are buying many units of real estates with the objective of diversification and individual real estate buyers who do not have the ability to diversify as real estate investors due to their limited budget and ability to borrow. The third objective is to provide a new approach utilizing time series techniques, causal models and GIS tool to model and forecast real estate price changes. Socio-economic factors are the forces driving demand and supply for real estate. These forces include but are not limited to demography, education, income, inflation etc. GIS helps to map these forces to a specific real estate area facilitating investment decision making.

Mehmedović *et.al*, 2010 defined real estate property as a financial asset that brings interest, benefits and encumbrances inherent in the ownership of the land and all improvements that are permanently related to it. This Chapter attempts to review the published literature on real estate investments, real estate evaluation and real estate forecasting. Most research on real estate investments concentrates on real estate portfolios that are traded in secondary markets, namely REITs (Real Estate Investments Trusts). Research on REITs focuses on using time series models to understand the price of volatility and its effects on average REITs returns. This Chapter will provide a review of research done on REITs volatility and average returns. The Chapter also reviews the discounted cash flow and price to earnings methods of evaluating real estates. The socio-economic factors that shape the decision making process of real estate buyers will be discussed and connected with proper locations for real estate development using the geographic information system (GIS). The Chapter will then determine the research gap that the thesis intends to bridge.

2.1 Returns on Investment

Barras, 1983 outlined that, whenever the dominant factor is user demand automatically cyclical fluctuations will be seen in real estate development. According to Barras, 1983; the external factors which either increase or decrease the price trend according to change in environment are; construction costs, variation in user activity, availability of credit etc. In real estate, persistent buying and selling increase liquidity of the market, although the market is also prone to interventions by the government.

The UK Property Industry Alliance (PIA) Data Report 2012 shows the importance of real estate assets in UK. The report points out the fact that even when the real estate market is worth 4,224 billion GBP, institutional investors transaction has been only of 2.3 out of the 4,224 billion. The report expects a change in direction in the long run as residential real estate offers better return and incurs less risk too. Also, the report states that in the period to come, there will be huge demand for residential units in the UK. Thus, the shift is inclining toward residential units due to future demand. Unlike stocks and bonds, forecasts cannot be estimated with accuracy in this domain because of the lack of appropriate research.

In general, investment in real estate involves a huge amount of funds thus it is essential to devise a technique which will guard investors' interests. This section will help to understand the nature of return from real estate investment. For example, real estate return is considerably influenced by geographical location, which does not influence the stocks and bonds market.

2.1.1 Returns from Housing Property

A key category of real estate assets are housing assets and the prices of these often reflects the volatility of the market. Case *et.al*, 1990 studied single-family home prices amongst housing assets and found that returns on investment are location dependent. Their empirical study of different metropolitan areas globally showed that housing assets (Flats/single housing) demonstrate a better return on investments than rural areas. Their results are dependent on other factors such as income, population growth and material/construction costs.

In an ideal real estate market, price would always signify the agreed value which is determined by both sellers and buyers. The deviations from the normal trend of real estate prices can be either positive or negative. Real estate markets are supposed to have a lower liquidity than financial markets.

2.1.2 Inflation and real estate returns

There is a common belief that the performance of real estate value depends on inflation. Wurtzebach *et.al,* 1991; studied the impact of inflation on the value of assets. They showed that real estate does provide an inflation hedge. They concluded that when market imbalance occurs, the risk increases and the returns suffer regardless of inflation. Rubens *et.al*, 1989 stated that the real estate hedge against inflation depends on the type of real estate. Their results have shown that a mix of commercial and residential properties provide an even better inflation hedge. The example given below shows how real estate investment beats inflation by a wide range in most of the time periods.

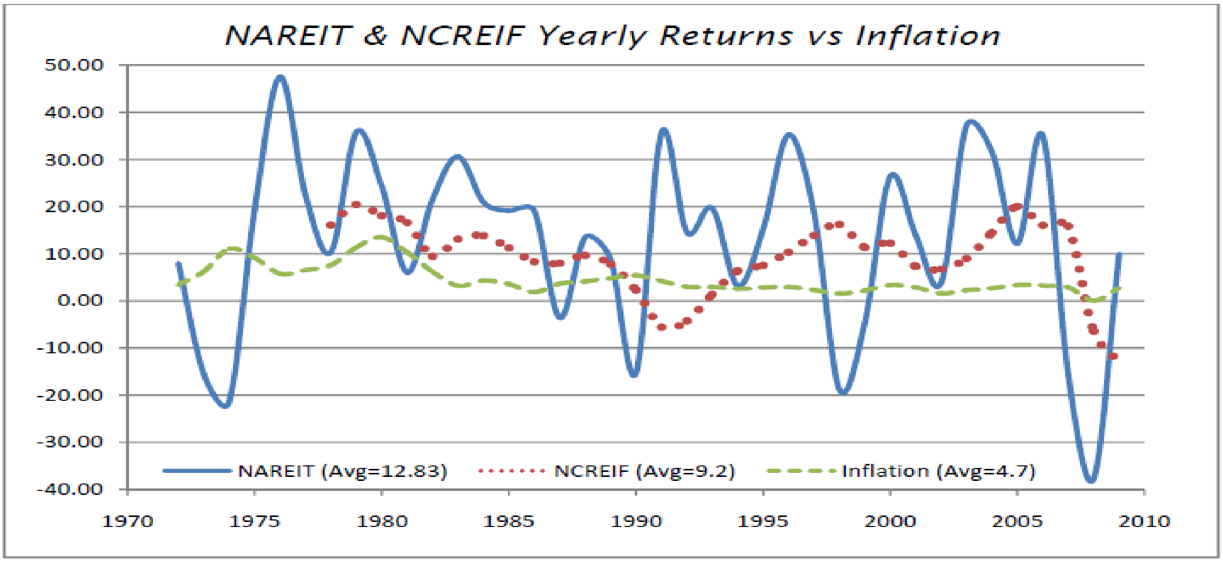

Figure 2.1: How Real Estate Beats Inflation (Source: wwww.NCREIF.org)

NAREIT refers to the National Association of Real Estate Investment Trust while NCREIF refers to the National Council of Real Estate Investment Fiduciaries.

2.2 Risk Classification, Spreading Risks with REITs, and Volatility in Real Estate Investment

2.2.1 Risk Classification

Any investment strategy that aims to provide operating rules for the buying or selling of real estate requires certain strategies to handle risks. The risk of investment in properties depends on the type, location and status of the property in terms of its development. Giliberto, 1993 classified risk as:

- Low Risk – properties which are associated with stable long-term cash flow.
- Moderate Risk – properties which are associated with less predictable cash flow.
- High Risk – properties which are associated with limited or no cash flow. Their return is maximum over the long term.

When formulating a diversified investment portfolio, investors are certain to encounter different types of risks attached to different types of real estate assets, but being able to prioritize risk is a skill that will surely help to achieve optimum portfolio returns.

2.2.2 Spreading Risk with Real Estate Investment Trusts (REITs)

A Real Estate Investment Trust (REIT) is a type of real estate company modeled after mutual funds. Land is considered to be an income generating real estate and improvements – such as apartments, offices or hotels are an integral part of it. (www.reit.com). Dividends of a REIT are equally shared among shareholders.

According to the modern portfolio theory of Markowitz, 1959; overall risk can be minimized through diversification as long as the correlations are negative or low. As suggested by Martin A. Smith, 2010; "it is widely accepted that within the total real estate portfolio, REITs tend to offer a strong feature".

Distinct from land property that is directly owned by an investor, REITs are a liquid asset since they are traded on organized exchanges. By using REITs, investors have the power to diversify all their holdings spanning across various geographical locations and specializations of property. Waggle and Agrawal, 2006 show that mixed portfolio of stocks, bonds and real estate decisions depend more on the expected return of REITs relative to that of stocks. The REITs portion of a mixed portfolio will depend on its correlations with stocks and bonds. Lower or negative correlations will be preferred as they will improve the risk adjusted returns efficient set.

2.2.3 Volatilities in Real Estate investment

There have been many studies undertaken on the fluctuations or volatilities of the real estate market in order to minimize risk in real estate investment. Hung *et.al*, 2009 studied the relationship between the different types of volatilities by using the Generalized Autoregressive Conditional Heteroscedasticity in Mean (GARCH-M) model for the REITs returns. Their findings showed that REITs returns are higher when volatility is greater. In essence, when volatility is high investors require higher returns on their investment. As noted above the standard practice in real estate risk management has been to work with REITs indices mixed with various equity indices. However, this does not consider the basic question of when to buy or sell a piece of real estate. Buying and selling operating rules with the objective of maximizing end of period wealth will be developed in the thesis. The thesis will develop a quantitative model that will provide us with the tools upon which operating rules can be based.

2.3 Diversification and Portfolio Optimization Benefits

The seminal work of Markowitz, 1959 and Seiler *et.al*, 1999 attempted to answer the important question, "Do returns from real estate behave like stocks and bonds?" The research demonstrates that these returns have low or negative correlation with other types of investment returns, so in fact real estate adds significant diversification benefits to a mixed portfolio. Worthington and Higgs, 2003 studied the correlation between UK regional property markets between 1976-2001. Worthington and Higgs, 2003 used the data from nationwide to do their analysis. They used Granger causality tests to measure causal relationships in the short-term, while Wald test statistics in a level VAR approach for long-run causality. Their study pointed to the fact that UK property market was segmented rather than integrated. They also pointed out that diversification across different property classes would facilitate investors to reduce portfolio risk while assuming the expected return to be constant. They suggested that risk reduction was limited because of high positivity in asset class correlation. Worthington and Higgs, 2003 outlined the fact that using the concept of diversification in portfolio theory, accurate forecasts of regional prices can be made. They argued that forecasting techniques in regional property markets should consider both exogenous and endogenous variables. Worthington and Higgs, 2003 concluded that if the series were co-integrated, long relationships were ought be considered in order to enhance accuracy.

2.3.1 Countering Real Estate Volatility with a Diversified Investment Portfolio

It is well established by now that holding a well-diversified portfolio tends to reduce risk. According to Addae-Dapaah *et.al*, 2002, "investment diversification (usually when it comprises international investments) has the potential to cater to a wider range of options for investment, provide developed returns of risk-adjustment and minimize volatility when the investment is made in assets that are real in nature. Usually, the merits of diversification are highest whenever the correlation between the assets is relatively low, hence a portfolio of a well-diversified nature would necessarily be inclusive of poorly or negatively correlated assets".

Brown, 1997 found that high levels of risk reduction are possible if low correlation exists between the returns on individual properties. Brown argues that, due to this correlation it is hard to build a portfolio with high diversification. He argues that portfolio return can be divided into two parts; systematic and unsystematic. His work points out to the fact that as a portfolio increases in size and reaches the market expectation the variation in the portfolio returns will approach the systematic level or there is a declining asymptotic function between portfolio size and standard deviation. Brown concluded that it is an empirical matter whether it is possible for investors to forecast successfully positive abnormal returns. He also notes that if perfect forecasting were not possible then some policy of diversification may benefit the investor to achieve a positive return. Farinella *et.al*, 2013 studied the relationships between the real estate and stock markets in Poland, and found that stock returns and real estate prices are directly related. They also argue that stock returns do not seem to directly affect rents. There existed a low correlation between real estate and stocks in Poland, which led to towards diversification opportunities for investors.

2.3.2 International Diversification

Goetzmann *et.al*, 1995 suggested that on a global scale, real estate market crashes have a lot to do with economics and monetary factors. Lee, 2005 outlined that the addition of real estate to a mixed asset portfolio not only enhances the compound return of the portfolio, but also potentially greatly reduces the risk. He tested this hypothesis by using yearly profits in the U.S. between the years 1951-2001. He revealed that land property has the potential to improve the return distribution on the portfolio compared to return on its individual composites. Lee, 2005 next outlined the method of Booth and Fama, 1992 for approximating returns due to diversification of a portfolio investment. Modern portfolio theory shows that if the correlation of any given property with a specific portfolio happens to be relatively low then its role in minimizing the portfolio's risk will be substantial. The study of the diversification effects of worldwide land property securities by Gordon *et.al*, 1998 explains that there are risk reduction benefits to having a mixed-asset portfolio of foreign land property securities, U.S. stocks, corporate bonds and global common stocks. The study used efficient frontiers to demonstrate the risk reduction potential of foreign land property mixed in with US land assets. However when combined with Chua's 1999 study of the function of global land property in a mixed-asset portfolio, the optimism Gordon *et.al*, 1998 had for foreign diversification is cooled as investors are reminded that they may also need to manage different taxes, operation expenditures, asset management fees, as well as foreign real estate valuation accuracy and valuation variance.

2.3.3 REITs in Comparison to Diversifying with Private Real Estate

The land available for real estate property is tentatively fixed in supply. The supply is constrained even more severely when in a metropolitan setting with expansion controls in place. Real estate investors may lengthen their investment periods as a result of high transaction expenses and site specific funding. This long holding period is closely associated with location demographics. REITs provide a better option for diversification than accumulating different private real estate investments according to Brown's 2004 analysis. As REITs diversify in the sense that they share risk, there are reasons now to suggest that sharing risk might be a superior option for a portfolio than spreading risk among different real estate assets. Ziering and McIntosh, 1997 study the merits of including both REITs and core land property (using the National Council of Real Estate Investment Fiduciaries' returns) in a mixed-asset portfolio of stocks and bonds within the years 1972-1995. The mixture of spreading risk and sharing risk is suggested by the authors to be the key to achieving optimal returns on a portfolio, as the maximum option of mitigating risk will be applied in investment.

Johnson and Weber, 2008 explain that REITs have credibility as a strong investment option because investors can sell certain shares of their assets when prices drop without actually removing themselves completely from an asset's potential returns. Other general reasons for investment in REITs are poor performance of a physical real estate portfolio and rapidly declining property income and asset values. Buckley, 1994 outlines that it is pressure on management from both the public and private sector in the real estate market that brings disposition of real estate assets into their decision calculus. Buckley, 1994 provides a clear insight into the investor's decision-making process, and this is useful as it translates to how much risk an investor is willing to accept before taking actual selling steps. Overall, Buckley's work serves to justify diversification, and substantiates the claims of modern portfolio theory in developing diversified investment portfolios across markets.

2.4 Real Estate Investment Specific Risk

Hutchison *et.al,* 2005 argues that the risk in real estate investment is attributed to the valuation of property carried out by evaluators. They suggest that an investor should consider many risks, notably valuation accuracy and valuation variance. Earlier both empirical and theoretical researchers analyzed risks and returns. In order to minimize the risks of investment Nitsch, 2006 posited that choosing a prime location is the main component for minimizing investment risks. To substantiate his argument he developed a price model based on the structure, location and rent of various cities in Germany. However, this model does not assess the elements of risk fully since it ignores correlations. Even though Markowitz, 1959 modern portfolio theory is considered a major breakthrough in the financial world, some real estate researchers have reservations regarding the theory. Portfolio selection and development too often relies on past performance. Three types of forecast are required with regard to future expected return, expected correlation and expected volatility.

Ali, 2006 carried out an extensive investigation and found that modern portfolio theory without major modifications cannot be used in risk assessment for real estate. Markowitz, 1959 and Hung *et.al,* 2009 both acknowledge that volatilities in the price market significantly affect investment portfolios returns, but it is generally observed that higher volatility in the prices lead investors to pursue higher returns. Markowitz, 1959 suggested that the only way for investors to pre-empt these prices was to observe market trends and predict future trends by interpolating past data. Moreover, during the past decade a great deal of interest has also been generated in attempting to better understand the performance of this asset class for the purposes of benchmarking and manager performance. Knowledge of the risk and return characteristics of commercial real estate can help us more fully understand the future performance characteristics of this type of asset. Attempting to explain forecasting factors in the real estate market, Armstrong, 2009 noted that everything that surrounds us is in a constant state of movement. There is a constant change in interrelationships which explains the volatility of markets where an asset that is worth a great deal one moment might become of less worth the next. The largest asset component in all other markers is real estate; being so, this draws a lot of investors to the real estate market. Armstrong, 2009 defines the key cycle for real estate as the determining factor which affects the sustainable length of a profound economic depression.

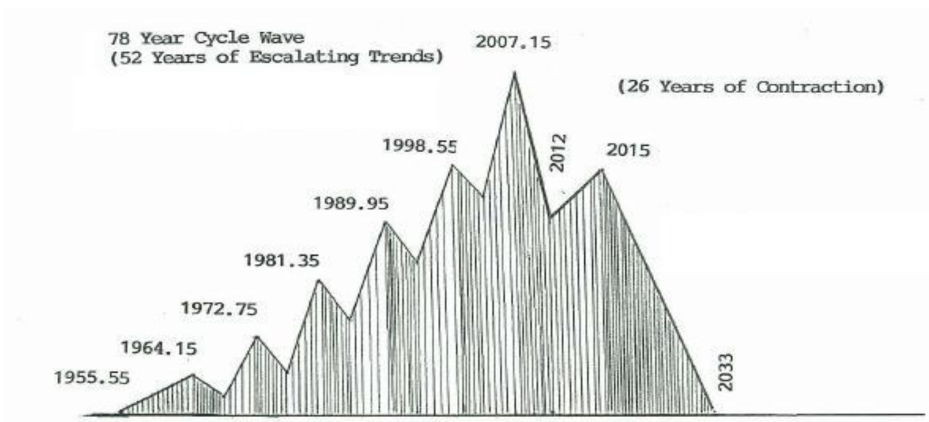

Figure 2.2: 78 Year Cycle Wave for Real Estate; Source: Armstrong (2009)

According to Armstrong, 2009; decline from the major high covers a time span of 26 years. This leads to a wave of rising prices of roughly 52 years.

There is a popular belief that investors are impatient about choices in the short term but highly patient when it comes to investing in long-term options (Grenadier, 2006). Price behaviour in real estate has always been a key area of interest to investors. As real estate investment heavily involves temporal factors, various stakeholders and portfolio managers have always been interested in predicting the price behaviour of real estate assets. Over the years there have been many studies by researchers on price behavior, which have led to the development of appropriate models to suit the diverse needs of various stakeholders. Stockman and Tesar, 1995; Girouard and Blondal, 2001 and Lane, 2001 describe housing price behaviour from a dynamic general equilibrium point of view. Using comparative studies of empirical housing prices across different markets, researchers have often theorized about what factors create shocks and maintain price regularities in the housing market. Driffill and Sola, 1998 developed a model in the context of market bubbles to go further and show how housing prices respond to the boom and bust scenarios that often characterize real estate markets. Attempts have also been made by Ortalo-Magne and Rady, 2003 to incorporate variables like transactions in the real estate sector, changes in the demography of participants and other macro factors into a model that responds to market bubbles, such as that developed by Driffill and Sola, 1998. Engle, 1982 estimates inflation based on time-varying volatility (ARCH, Autoregressive Conditional Heteroscedasticity). He argues that, prior to a tradeoff between allocation assets, an estimate of uncertainty has to be made. DiBartolomeo *et.al*, 2005 constructed a model which estimated risk and correlation at both the property and portfolio levels. This allowed a potential investor to analyze the potential risk corresponding to a possible investment as well as the particular risk prone components. Although modern portfolio theory (Markowitz, 1959) has become accepted as a tool in managing real estate portfolios, diversification has become a more established norm amongst real estate investors.

With regard to how investors will respond to diversification, one cannot make the assumption that investors are uniform in their behaviour and their perception of risk. Blume and Friend, 1974 have concluded the following about the behaviour of risk averse and risk seeker investors with regard to diversification: Risk-averse investors will always invest in portfolios and asset classes that will give them the highest long term risk adjusted rates of return. Risk-seeking investors will always invest in portfolios and asset classes that give them the chance to achieve abnormal rates of returns. Risk-averse investors will favour diversified assets if those assets have a high risk-adjusted rate of return. They may however refrain from investing in markets that demonstrate extreme price volatility. Risk-seeking investors will see diversification as a positive in the sense that they will be prone to different shocks in different markets, which will likely give them abnormal returns.

As investors are often more risk-averse than risk-neutral or risk-seeking in most scenarios, the general objective in developing a product portfolio is to maximize return with minimum risk. In order to develop a portfolio with a low standard deviation signifying lower levels of risk, one needs to be familiar with the past price movements of the same. Mueller and Mueller (2003) showed that public and private real estate investment, when invested simultaneously in a mixed-asset portfolio, improves efficient frontiers of the Markowitz mean-variance model more than the inclusion of just one or the other in a portfolio. Therefore the risk of an overall real estate portfolio can in fact be reduced.

The analysis of several sources in the literature (Ling and Naranjo, 1999; Quan and Titman, 1999; Ziering, Liang and McIntosh, 1999; Ciochetti, Craft and Shilling, 2002; and Feldman, 2003) suggests that real estate investments usually produce better risk adjusted returns when combined in a portfolio with stocks and bonds. Trippi, 1977 pointed out that the appropriate time to sell real estate units is highly uncertain. Following up on studies that investigate the correlation between residential property transaction cost and property selling time, Anglin *et.al*, 2003 and Capozza *et.al*, 2004 focus on the vibrant properties of the difference equations that occur when markets show mean reversion and serial correlation. This is specifically applied to the more illiquid U.S. single-family home markets. The aim is to discover reasons for dissimilarities in the vibrant response of the metro areas in order to alarm the local economy. The information is borrowed from the big panel Figures set, as described by Capozza *et.al*, 1995. The data covered 62 metro areas in the U.S. for a period of 17 years from 1979-1995. Capozza *et.al*, 2004 link experiential estimates of mean reversion with sequential correlation, which implies the difference equation. Their work further categorizes and analyses the dynamics of the difference equation and then establishes main attributes such as frequency and amplitude to correlation and reversion parameters. The study gives a thorough idea of "cycles" and "bubbles" in the guideline of the typical model which can be linked to the probable individually motivated patterns to the required reversion parameters. Pyhrr *et.al*, 1990 developed a verdict framework and ready model that projected investment returns under different inflation cycle circumstances. Their analysis explained that inflation cycles affect the risk of investments, although the extent to which investments are exposed to risk as a result of inflation is left open.

The vibrant properties of housing markets are subject to the specifics of location and time. Economic shocks can change potential outcomes of the real estate market, and one needs to consider the impact that they have on real estate assets before investors incorporate them into their portfolio. Risk is therefore always a factor in the purchase of real estate assets, even if this can be minimized by diversification strategies.

2.5 Real Estate Evaluation

The different sections and subsections of this Chapter have so far concentrated on different aspects of real estate investments including REITs. This section is going to review some common methodologies which can be used to determine when to buy and sell real estate units. Real estate investment return is a combination of income and capital growth. Clayton, 2009 discusses capitalization (cap) rate approach as a very common method of analyzing the value of a property that is producing income. Cap rate is defined as the ratio of property net operating income to current market value. Cap rates are generally derived from the discounted cash flow (DCF) methodology. Woychuk, 2011 emphasized that real estate return is difficult to calculate and requires the property to be appraised frequently.

As suggested by Green, 2012; although real estate prices fluctuate in the short term, in the long run they are very much driven by its rental income. Through rent, such property generates regular and steady income and its value has the propensity to fluctuate like stocks. The price to earnings ratio can help an investor learn about the true value of a property. According to Green, 2012; although the P/E ratio is a crude method it still manages to give an idea of the transaction i.e. when to buy and when to sell. If the ratio is lower than the market current nationwide fair value, then the deal is supposed to be good; in other words one is eligible to collect high rents from the property. On the other hand, if the ratio is twice as high as the current nationwide fair value then the real estate unit ought to be sold.

Discounted Cash Flows (DCF) require clear inputs if they are to yield accurate forecasts. If the inputs to DCF such as free cash flow forecasts, discount rates and growth rates are not accurate the final value of the asset will also be inaccurate. The thesis tries to improve the forecasting accuracy of the methodologies by improving the accuracy of forward P/E and DCF inputs through better scientific decisions with regard to real estate location and timing of buying and selling.

Many authors have used one or more methods of time series analysis for real estate prices depending on the nature of the times series (Baum and Key, 1999). The method that is found to be very effective in time series forecasting is the Box-Jenkins, 1970 method since it can handle (a) volatile or stationery fluctuations, (b) periodic or cyclical movements, and (c) trends (either increasing or decreasing). Green, 2012 stated that unlike stocks and bonds, in real estate, one does not have the option to sell at one's discretion; it is not that easy. One has to be a good forecaster because one needs to buy in a down market and sell in an up market.

2.6 Identifying Real Estate Investments Locations by Considering Socio-Economic Factors

There have been three major methods for determining the proper location for real estate investment (Fik, 2005; Fujiwara *et.al*, 2011); ring study, drive time analysis and the gravity model. GIS tools have been used within the above three methods. The new contribution of this thesis is to combine socio-economic databases with the GIS tool to create thematic maps that can help to identify optimal location for real estate investment. GIS helps to map relationships between demographics, household income of a particular location and investment in real estate (De Man, 1998). Real estate market analysts have most frequently relied upon ring study and drive time analysis (Fik, 2005). In ring study, circles are drawn around the real estate unit. The area of the circle goes on increasing until it covers specified attributes or a desired number of customers. Drive-time between areas considers customers' willingness-to-travel, multiple modes of transportation, and consumer preferences. A variety of issues such as migration routes, telephone traffic, passenger movements, commodity flow, etc. have been analyzed using the gravity model, whereby as distance increases, interaction between two objects or places decreases provided other factors remain constant. Conversely, as mass increases, so does the interaction between the two objects (Thrall, 1997).

Real estate located within the radius of a closely linked network of rails, roads, supermarkets, schools, and hospitals automatically attracts a sizable number of investors, as these investors feel safe with respect to the fact that the chances of losing their investments are low. The real estate industry does not consist of only risk averse investors. There are also risks seeking investors who seek undeveloped and virgin areas for investment. When the land in these locations develops over time, these investors become the first to reap the financial benefits. The question is how it is possible to identify locations that have the potential to achieve substantial rates of return. The answer to this may lie in the use of the Geographic Information System (GIS) that helps to identify potential locations but also allows for management across functions, such as site planning, asset management, and market analysis (Hernandez, 2007). With the emergence of geographic information systems (GIS), both spatial and non-spatial data can be managed simultaneously, resulting in data query, analysis, and data visualization (Li *et.al*, 2003). A thematic map is a combination of spatial and non-spatial data created with the help of GIS. It helps to solve location problems and presents the results in visualized forms when necessary.

It is desirable for researches to map surfaces (e.g., rent or price surfaces) (Clapp, 1997). These maps show the distribution of location price within a city. Since, real estate valuers use large databases the Geographic Information Systems (GIS) helps to analyse these data and create useful models in the valuation process.

GIS is designed to categorically manage and analyse spatial relationships. Over time, the use of GIS has changed too; where in the past it was predominantly used to prepare maps and associated databases, now it can be used to help in planning, managing and operating facilities. GIS is used throughout the life span of a facility. With respect to real estate its use can be defined as: site selection, design, construction, maintenance, investment and ultimately disposing and reclamation.

The importance of spatial and temporal relationships was first mapped by John Snow in 1854 when he represented the relationship between deaths from cholera and the locations of water pumps used by the residents of Soho, England. Goodchild, 1989, Densham, 1989 and Peterson, 1998 have outlined how GIS could be utilized for various applications including risk management. Thrall and Marks, 1993 described the ability of the tool to study the spatial aspect of real estate decisions. Marks, Stanley and Thrall, 1994 provided guidelines for evaluating GIS software for the analysis of real estate. Castle, 1993 argued that the real estate industry including brokers, home buyers and developers could use GIS to improve residential brokerage by merging it with Multiple Listing Service records. All of these studies have underlined the importance of using GIS in real estate but none have come up with a model that combines spatial and non-spatial data to choose optimal locations for real estate investments within identified competing areas.

Donlon, 2007 reported that precision of both spatial and non-spatial data improved the reliability of GIS analysis. With the increasing demand of GIS data in the planning profession, spatial data has become very important. This data can usually be acquired from the county planning office by the private sector for a nominal fee. This data is usually relatively current, though it is not uncommon for it to have small errors.

The spatial data selection function includes a set of variables for particular locations from a spatial database (Fik, 2005). Fik's, 2005 objective was to find the factors that demanded higher house prices. The objective of the current thesis is to identify the optimal location for house investment that would demand a higher price based on socio-economic factors.

Brown, 1997 argues that investment in real estate is a heterogeneous type of investment. Although market factors are important for the ups and downs in the market, returns are also dependent on many other factors such as type of tenant, age of property, location, socio-economic parameters, condition of the neighbourhood and so on. Paul and McNamara, 1997 mentioned that forecasts in rental growth are important in the real estate market. Breedon and Joyce, 1993 analysed the relationship between real house prices, earnings, disposable income, demographic factors, the rate of repossessions by lenders, and the stock of dwellings. They concluded that real house prices are directly linked to real personal disposable income, demography etc. They also concluded that investment in residential units is also linked to mortgage lending. Paul and McNamara concluded that forecasting and econometric modelings are important but not totally sufficient to guide the investment process.

Fujiwara and Campbell, 2011 discussed the link between an individual's wellbeing and the health of the adjoining locality. Location or neighbourhood preference is not constant over time. Location is about identifying the best space or site that matches clients' preferences. This location is one of the most important factors determining the value of real estate. Andrea Podor, 2010 highlighted the following factors as important from the aspect of getting peoples' preferences on the location: *historical review (neighbourhood established status), present population, demographical tendencies, employment rate, major work possibilities, income conditions (average salary) etc.*

Denzer, 2005 stated that GIS has proven to be useful for all sorts of business use. Its addition to a business decision making process enhances the efficiency of decision maker. However, its usage has been limited to the creation of thematic maps, joining spatial and non-spatial databases, running spatial query, distance measurement and buffering.

GIS is particularly well-suited for real estate practitioners where "location" is the main criteria (Amos, 2009). Market analysis involves studying geographic areas to identify large numbers of potential customers. Dangermond, 1988 discusses how GIS stores and analyses geographical data and points out that before the housing crash, people relied on about seven odd variables to determine site selection.

Sweet, 1995 noted the requirement for many specialists as an integrated part of a wise decision making process; in addition to the normal requirements that were needed for any successful land program. He mentioned that it was very much important for the investor to have a current knowledge about the constant change happening in the surrounding environment. Sweet, 1995 mentioned that GIS defined the geography which is important for measuring demand for the properties since that area determined future demand for the residents who would be living in those residential units.

Weber, 2001 discussed the basics of GIS and how it can be used for the valuation of real estate. German, 1999 noted that in real estate assessment models, parameters such as land sizes, building ages and other physical attributes had both been regularly calculated and calibrated, but that location was not. Location was an important feature to empirical research and quantification. German, 1999 mentioned that this location parameter could be well addressed by the use of GIS.

Cowley, 2007 discussed how most of the research centered on using GIS technology in relation to property markets is dedicated to mass appraisal for rating and taxing purposes. The importance of socio-economic factors when it comes to real estate location decisions can never be ignored. The present thesis utilizes the Geographic Information System (GIS) to improve on the existing methodologies of ring study, drive time and the gravity model. GIS has the advantage of linking analytical database with spatial components such as locality. This linking advantage is what distinguishes GIS from other alternatives such as ring study, drive time and the gravity model.

A real estate project can fail not because it does not provide value, but because it does not correspond to what buyers are looking for. The basic objective of qualitative factors about buyers' preferences is to ensure robustness of decision making. Brereton *et.al*, 2008 analyzed the relationship between an individual's well being, status quo and adjoining environment. They found that staying in public housing leads to a negative set up of mind and thereby decreases an individual's satisfaction. Living in a deprived area has a direct effect on life satisfaction (Dolan *et.al*, 2008 and Abraham *et.al*, 2010). Thrall, 2002 reports evidence that success in real estate investment is dependent on success in identifying the location of a successful real estate asset. Successful parameters which measure the success rate of real estate development are demographics and other associated spatial attributes.

Real estate market analysts are mostly dependent on ring study and drive time analysis (Fik, 2005). In ring study, rings are created around potential real estate asset(s) or proposed locations. Areas of rings are increased until the required number of potential clients has been included within the trade-able rings areas. Drive time analysis helps to calculate the time of travel between the source and the destination.

Consumer behavior is the study of individuals' choices when it comes to decision making. On a macro level, marketers are keen on understanding demographic changes which affect the investors' interaction with the market. On a micro level, investor behavior focuses on individual behavior and the reasons behind those actions (Gibler *et.al*, 1998).

Individuals may make decisions using compensatory or non-compensatory rules, or by using both. In the former, the individual chooses key features, ranks available real estate on each feature, and chooses that real estate which has the maximum score (Alba *et.al*, 1987). In the latter, all the essential features are not summed up and do not point to the fact of positives outmatching the negatives. Important features which in a true sense help to evaluate all available alternatives are called determinant features (Alpert, 1971).

The personality of an individual can be a criterion affecting their investment in real housing choices. A risk seeker is likely to buy a dilapidated property, would invest in the property and would reap the benefit when the neighbourhood starts to improve. Change in personal attitudes, change in personal lifestyle, and change in personal tastes tend to influence individual preferences for space. Market analysts should not only consider census information but must also take into consideration behavioral changes and motivations (Rabianski, 1995).

There are certain issues that need to be addressed before establishing a trend for behavioral patterns. Armstrong, 2009 highlighted the importance of understanding price behavior of real estate units since this will explain how the price will change under different market scenarios such as boom and bust. Podor, 2010 discussed the importance of tax attached to the unit. People tend to estimate the annual maintenance of the unit through this tax band. Also the tax band helps people to identify the status of the locality i.e. whether affluent people or poor people living on state benefit are residing in that locality. Fujiwara *et.al*, 2011 and Podor, 2010 have discussed how crime and safety affect neighbourhood status and have a great impact on real estate prices.

Dolan *et.al*, 2008 stated that a key parameter for life satisfaction is neighbourhood status. Rabianski, 1995 concluded that a major element for life satisfaction is access to green areas. Podor, 2010 has explained that basic education is one of the factors which influence people's decision in buying or selling real estate units. People want to live in a locality which matches their education because they will be able to find neighbours according that educational status. Brereton *et.al*, 2008 found out that there is a tendency for people with same ethnic race to live in the same locality. Brereton *et.al*, 2008, Thrall, 2002, Fujiwara and Campbell, 2011 and Podor, 2010 indicate the importance of the proximities of basic facilities from the units people they want to buy. These facilities are hospitals, schools, community centres, police stations, post offices, railway stations etc.

Thus, we can see that there exists an individual mindset towards real estate investment and real estate location; how human behaviour takes an important role in developing a causal relationship between real estate unit and real estate location. Using the database of the ONS, we will be able to group these causal factors into respective domains which are defined below. The Office for National Statistics (ONS), United Kingdom uses seven domains to paint a broad picture of an area to determine the surrounding conditions (the data on the seven domains is available at www.ons.gov.uk, and the database used in this Thesis is for 2011). The domains measure the general 'health' and determine the success localities across the city.

The seven domains are:

- Crime and Safety
- Economic Deprivation
- Health Care
- Housing
- Personal Consumer Debt
- Social Grade
- Environment

The inputs for the above domains are updated at least annually. These domains are used as guidelines for mapping real estate investment opportunities. The above domains are explained below:

- Crime and Safety

ONS Statistics hold data on vital offenses from the crime series record of the Home Office. This data is used to analyze crime trends at the local level. In the database of crime and safety, data of both crime (theft, burglaries, harassment, arson attacks etc.) and safety (fire related issues) have been clubbed together. The records under the crime component have been suppressed because it is considered a sensitive issue, and the ONS has not included records on this. However, information on the same component can be acquired through an official letter. This domain is not used as data is considered sensitive and can only be released on official request. However, the main rationale behind this domain is that a postcode investment will be more desirable if its crime rate is lower than another competitive postcode.

- Economic Deprivation

This is a count of housing benefit (council tax benefit) claimed by the residents. The ONS collects this data from the Department for Work and Pensions. This factor can help the real estate developer in locating his targeted clientele. For example, an area with a low count of council tax benefit may be suitable for an above average housing development project.

- Health Care

The ONS holds data on the general health of the residents living in different areas of the UK. As per the ONS data are categorized into: good health, fairly good health and not good health. This information can help the real estate developer in locating appropriate projects.

- Housing

The ONS holds data on the following housing relevant items:

1. Total number of dwellings, which gives the total number of domestic properties (known as dwelling stock as per ONS statistics) along-with the number and percentage of properties which have got respective standard council tax bands.
2. Average house price.
3. Housing stock, which includes households' spaces and occupied household spaces. This item is not used as we could not get data for it from the ONS website

This domain can provide vital statistics with regard to the target house price and its related council tax band.

- Personal Consumer Debt:

The ONS data shows in GBP the range of average personal consumer debts per person for different districts. Debts relating to individual businesses or commercial concerns have not been included. A district with higher average consumer debt and classified as dominated by high a percentage of unemployment will exclude high end real estate projects for this district. It could be more suitable for affordable real estate projects with options for land or real estate swaps.

- Social Grade:

The ONS data shows individuals aged 16 and over living in households as per their respective social economic classification: higher and intermediate managerial professionals, supervisory professionals, skilled manual workers, semi-skilled and unskilled workers, and on state benefit or unemployed. This domain is very important in the planning stage of a real estate investment location. As explained in the discussion of the personal debts domain, social grades combined with personal debts are very important drivers for real estate investments.

- Environment

The ONS estimates the square meters of green space available in the locality. Given the type of real estate project considered and given the target customers' preferences, a decision can be made on whether or not customers will be willing to pay a premium for access to larger green areas.

These domains contain the data of each of the explanatory variables as per respective MSOA. These attributes are then linked with the area maps (will be mentioned in 5.1) through the ArcGIS software to generate thematic maps of each Middle Super Output Area. A thematic map is a type of map especially created to portray a particular theme related with a specific geographic area. These maps can help to visualize physical, social, political, cultural, economic, sociological or any other aspects of a city. The symbols and legends for the thematic maps are created by the default symbols and legend-shades present in the software. The yellow colored polygon boundary on the maps in the appendix shows the particular area of the MSOA. The legends for each map reflect the respective domain according to the datasets of the ONS (crime and safety, economic deprivation, health care, housing, personal consumer debt, social grade and environment). Each of the thematic maps gives us an entire picture of the full MSOA, for example, the thematic map on housing represents three aspects: total dwelling stock, house prices over the entire area and the council tax bands of the houses. The graduated colour or the symbol in the thematic map helps to identify areas with similar values i.e. in the case of council tax band, which polygon has houses under tax band A and so on. The maps help to visualize the area remotely i.e. without actually going to the site one can see a picture of not only a particular locality but also localities surrounding the same. The legend values of the maps will be used to do the empirical analysis in section 5.2 of the causal relationship between real estate price on one hand and socio-economic explanatory variables on the other.

2.7 Real Estate Forecasting

Miller and Sklarz, 1986 posit that forecasting in housing market analysis is both an art and a science. The art consists of sound economic factors which capture the true behaviour of the market, while the science is dependent on adopting the right modeling approaches. Medium to longer term housing price drivers were employment, income, supply constraints and interest rates. They concluded that the accuracy of forecasting depends on choosing the right market indicators in addition to looking at house price.

Bryan and Colwell, 1982 represented a good example of the specific regression method. Gloudemans, 1990 and Jensen, 1991 used it to set standards for property sale prices through analysis of time series for residential markets. Clapp and Giaccotto, 1992 measured quarterly price changes by applying dummy variables. Schwann, 1998 used the same method but changed dummy variables with an autoregressive scheme.

Birch and Sudarman, 2003 stated that all models of regression are related to the chosen fixed time interval span. There is no guide book for this i.e. selecting the right or appropriate time interval. The two ways exponential smoothing through its process of selecting less data overcome this difficulty.

Mehmedović *et.al,* 2010 tried to identify what factors govern first time buyers' investment decisions. Their work was mostly related to the qualitative aspects of decision making. In their conclusion they highlighted the need for additional quantitative analysis methods such as correlation and regression analysis. Mehmedović *et.al,* 2010 stated that people have various reasons for buying property. Their individual requirements determine what they look for in a property and accordingly they decide the amount they want to pay out to match those requirements.

At present, the significant driving forces of the global economy are both real estate units and financial assets. After the worldwide financial crisis started by the Lehman Shock, there has been a great demand for scientific know how regarding the potential returns and risks of financial investments on real estate markets. According to Ishijima and Maeda, 2012 investors should have easy access to measure return on investment for appraising their financial investment opportunities.

Wilson and Okunev, 2001 studied the combination of forecasts from different property markets (the USA, UK and Australia). They found out that pooling the forecast resulted in better out of sample results. Similar results were obtained by Diebold, 1989 who found that pooling forecasts from different models tends to yield better outcomes for shorter term forecasts than for component models. Winkler, 1989 found a similarity between this forecast combination and diversification in portfolios, as diversification tends to reduce risk; similarly a combination of forecast methods tend to yield better results.

Case, Quigley, and Shiller, 2005 reported that price variations in real estate had a direct effect on the stock market. Reinhart and Rogoff, 2009 endorsed this fact as a common aspect of markets across a number of countries and over longer time periods.

Lynn and Wang, 2011 analyzed the importance of forecasting in real estate returns. They held that there is a difference between retrospective and predictive investment practices. They mentioned that forecasting acts best on present data rather than reactive practice based on previous data. The need for forecasting techniques depends on the desired outcome i.e. someone whose interest is in income return will look for a quantitative analysis, whereas a forecast for capital return will require understanding of the expected change in the net operating income. According to Feldstein, 1992 and Kearl, 1979; higher inflation leads to decreased demand for housing. However, Quan and Titman, 1999 argued that inflation may increase demand for housing because real estate is seen as a hedge against inflation.

Kochin and Parks, 1982 report that in spite of the importance attached to the real estate sector get due to its huge returns, few studies have analyzed the forecasting of house price using timer series models. This is alarming when so much importance is placed upon predicting mortgage defaults, property taxes etc. There are two important features regarding the movement of house prices (Case and Shiller, 1990; Glaeser and Gyourko, 2006): the first is that house price changes are highly persistent from one period to the next, and the second is that the housing market is prone to large return periods followed by periods of low returns (Muellbauer and Murphy, 1997; Glaeser and Gyourko, 2006).

The problems with real estate measurement have been well researched by Graff and Young, 1996 and Geltner, 1998. Fisher, 2000; Fisher and Geltner, 2000; mentioned that traditional real estate appraisals normally use: (a) replacement costs, (b) comparable sales and (c) capitalizing the expected outcome. In the past real estate investors have tried to forecast or evaluate the risk of individual properties through Monte Carlo simulations where the investors have changed the inputs across their expected range. However, they are not reliable over large asset classes. It is important to note that in an approach using a Monte Carlo process, the period to period outcomes are random. That means that if in period six, you forecast NOI to be 100, then in period seven NOI could be any value – zero, 105 or 1000! In real estate there is a high correlation over time for NOI or any market variable and, therefore, the Monte Carlo process is not correct for assessing future uncertainty.

In economics, the housing market occupies a special status. This is because housing construction boosts the economy through increasing expenditure, employment and house sales volume. Recent studies further justify the necessity of housing price analysis by indicating that the housing sector acts as a major indicator of the economy, and that its prices help to forecast both inflation and output (Forni *et.al*, 2003; Stock and Watson, 2003; Das, Gupta, and Kabundi, 2009). Thus, a forecast for house prices can provide important results for policy makers and help them better control inflation and design policies. Also, these forecasts can direct individual market participants to make wise investment decisions. During the economic recession started by the sub-mortgage crisis, analyzing the influence of the burst of the housing price bubble and predicting its future trends is thus invaluable.

Theory of DCF is based on the fact that Weighted Average Cost of Capital (WACC) is the rate of discount. The DCF method takes into account the value of money with respect to current time (Mun, 2002). However, method comes with drawbacks: it is performed under deterministic assumptions (Wofford, 1978; Mollart, 1988; French and Gabrielli, 2004). In other words, one does not take into account the uncertainty in the estimated cash flows (Kelliher and Mahoney, 2000; Weeks, 2003). Fama and French, 1989 and Ferson and Campbell, 1991 mentioned that there is a circularity problem when debt is financing a portion of portfolio assets.

According to Wheaton *et.al*, 2001, when investment returns are largely random, as in the traded-security markets, the ‘model’ which best fits the series is often some (complete or partial) form of Brownian motion, in which the movement in returns today is independent of previous returns and largely a function of random error, possibly with a trend. The way that such ‘intrinsic’ variability is forecast is by simulating the random movements using Monte Carlo methods.

Real estate forecasting has become an integral part of the larger process of business planning and strategic management. Investors in real estate such as pension funds, insurance companies and real estate funds will make allocations, which they may maintain for some time. They are therefore looking to forecast future asset performance across real estate sectors and locations. Investors need to know where rents, yields and prices will be within a predetermined time frame.

Real estate developers would like to have an idea of future demand, rents and prices. Investors with loan portfolios, including those secured on real estate, demand objectivity and transparency in the analysis of future trends. Scenario analysis concerning future rent and cash flows is also important (Brooks and Tsolacos, 2010). Real estate forecasting can be considered for model selection. Forecasting is the natural progression in real estate as more data become available for a larger number of markets.

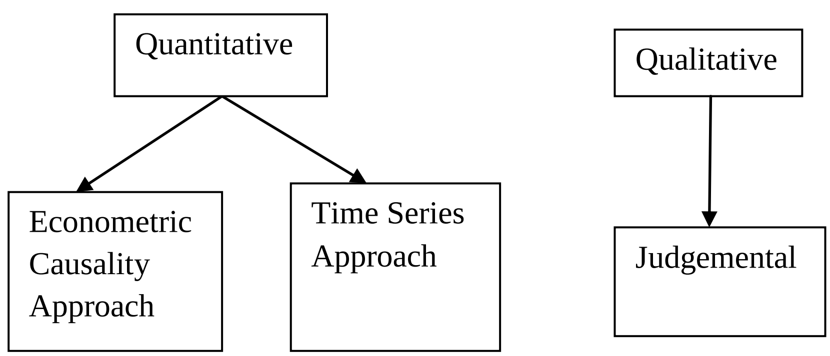

Figure 2.3: Towards an Integrated Approach in Real Estate Investment (Journal of Property Valuation and Investment)

The above Figure shows the current status of real estate investment. Gone are the days when only one technique was used in a forecasting study. Such studies are now a combination of both quantitative and qualitative approaches. Under quantitative approach, along-with the traditional use of time series there has been an increase in the use of the Causality Approach (for example, asking questions such as what if inflation causes rents to go down and vice versa). The qualitative approach is pure, and based on individual decisions which are normally governed by socio-economic factors and the wellbeing of the individual.

2.8 Cross-sectional Analysis in Real Estate Investment

According to Barker, 2004 and 2006; the period between 1980 and 1990 saw house prices in the UK rise by 83% and subsequently decline by 38%. This price change is higher than that of any single metro in the United States during the same period. According to Barker, 2004 and 2006; the planning system of the UK is widely viewed as inflexible. Historically, due to ignorance of market movements, the UK has failed to cope with changing socio-economic conditions in many ways. The planning rationale is straightforward: As long as there are no perfect substitutes for land plots, shifts in demand should create a larger impact on house prices and as a result prices should be highly volatile in places where space is entangled with strict urban planning laws. McGough and Tsolacos, 2000 stated that there is no single method to establish fair values in different markets and that is why the investor needs to look towards an integrated approach and apply these in real estate investments. Cross-sectional analysis is one of the methodologies that can be deployed for this purpose.

Hollies, 2007 conducted analysis for higher yields or lower yields across locations. She argued that, 'higher inflation is related to higher output, with investors asking for a higher premium to invest in higher inflation markets.' Saks, 2008 showed that a regulatory index which was created from various surveys between the 1970s and early 1980s in metros and which had fewer barriers to construction would experience an increase in residential construction and smaller increases in house prices in response to an increase in housing demand. Saiz, 2008 found that most metros which are widely regarded as supply inelastic are in fact severely affected by land use patterns. He concluded that elasticity of supply in the housing market is negatively affected by regulatory and physical constraints.

Hilber and Robert-Nicoud, 2009 through their work concluded that more desirable places would not only be physically developed but also would be more regulated. Empirical analysis would shed light on the question of the extent to which planning regulations differ across space and how this affects house price levels and corresponding volatility (Hilber and Vermeulen, 2009).

Hilber and Vermeulen, 2009 assumed that housing demand $Q^D_{j,t}$, with respect to place or location j and time span t is related to price and income. It is also related to time independent factors such as GDP, rates of interest in mortgages etc. Based upon the assumption that the above relationship is log-linear, the following demand equation is developed:

$$\log(Q^D_{j,t})=\epsilon^D\log(P_{j,t})+ \epsilon^\gamma\log(Y_{j,t})+\alpha^D_j+\beta^D_t+v^D_{j,t}$$

where $v^D_{j,t}$,reflects idiosyncratic unobserved heterogeneity. The parameters ϵ^D and ϵ^γ are the price and income elasticity of demand, where in small locality, price flexibility of house or real estate demand is large in comparison to regional or national area. This is due to the fact that neighbouring location plays a part in replacement of housing.

Miller and Sklarz, 1986 have discussed many indicators. They studied price changes for the period 1987–2007, using two indices, the Case-Shiller Index and the House Price Index, of the US to find evidence of price seasonality. On average house prices were higher in summer. Cheng and Han, 2013 used stochastic population forecasting technology in explaining price changes in real estate. They argued that the effect of long-term population development on housing prices on the deterministic population forecasting has a limited credibility. Their study was based on the city of Shanghai, China, and found that 'without the housing mortgage loan, the buyers take full risk of a decline in price. With the mortgage loan, the buyers regard the banks as the bottom line back-up.' They concluded that a positive association existed between long term population development and demand in real estate.

UK house prices rose 9% per annum from 1996 to 2002. By the last quarter of 2002, house price inflation was at 25%. Weeken, 2004 found that the ratio of house prices to net rentals was currently well above average. Barker, 2004 mentioned that the UK would have been 8 billion USD better off had real house prices risen in line with the European average since 1975.

Plazzi *et.al*, 2010 estimated the returns using cross sectional dispersion and growth in net operating income (NOI) of different real estate units such as residential units, factories, retails properties and properties which contain offices. They concluded that such real estate units exhibit cross-sectional time varying dispersions. Interestingly, they explained their time series fluctuation by macroeconomic variables.

Brooks and Tsolacos, 2010 stated that real rents move through time in response to the vacancy rate (defined as vacant stock over total stock, expressed as a percentage) and an economic output variable. Here, the indicator is vacancy i.e. it gives a picture of demand and supply conditions. With a fall in vacancy due to high demand and supply, the market becomes ideal for landlords. Landlords during this time push for higher rents in new leases or review the existing rent. Clapp, 1993 stated that when the actual amount of vacant space in a market is higher than the natural vacancy level, this can be taken as a sign of weak demand and surplus space in the market. As a result, rents would fall and new development would slow.

The mortgage crisis of 2007 is a classic example of what can happen in the real estate industry. Between 2002-2005 real estate prices were increasing at a very rapid rate. A point came where the prices were too high for potential buyers to take out mortgages. As a result, sellers were forced to drop prices back to a point which would attract potential customers. This increase in house price was a part of high consumer price index during that time, because the general public uses consumer statistics to analyse and understand the impact of inflation on their personal and household budgets and are aware of how inflation linked parameters such as taxes, house prices, interest rate, benefits, pensions and wages were likely to change. Inflation is defined as the rate (%) at which the general price of goods and services increase which causes purchasing power to fall (Holt, 2009).

Cameron *et.al*, 2006 discuss the strong relationship between housing units, regional incomes, real and nominal interest rates and demographics. Thus, it is evident that real estate requires not only house prices for forecasting but also needs to take account of inter-related factors such as demographics, regional income etc. which are required to test the forecasting model. Hasan, 2006 considered the following attributes in analysing the demand and supply of housing framework; income, rate of interest on mortgage, and demographic factors such as population growth.

It is clear that the real estate market is risk prone because of market volatility. The answer to this risk aversion is building of a forecasting approach covering the issue of price change, along with house demand, mortgage rates, inflation etc.

2.9 Real Estate Forecasting and Box Jenkins ARIMA Approach

Pagourtzi *et.al,* 2003 reviewed the valuation methods with respect to real estate appraisal. According to their work, valuation methods can be categorized into traditional and advanced. The traditional methods can be categorized into regression models, cost, income, profit and contractor's method. The advanced methods can be classified into the hedonic pricing method, spatial analysis or use of the Geographic Information System, and the fuzzy logic and Autoregressive Integrated Moving Average (ARIMA) models. They report that the choice of model is regarded as suitable as long as the results appear justified, reasonable and logical in agreement with accepted beliefs. Malpezzi, 1999 and Gallin, 2006 found a correlation between two sets: income and house prices, and house prices and rental.

There are two facts worth mentioning with respect to the movement of house prices, first is that house price changes may follow a trend over time (Case and Shiller, 1990; Meese and Wallace, 1991; Glaeser and Gyourko, 2006). Second, the house prices are prone to large swings over time (Muellbauer and Murphy, 1997; Glaeser and Gyourko, 2006).

Pagourtzi *et.al,* 2003 stated that since house prices show considerable growth rates, the AR(1) model is likely to forecast poorly. An AR(p) model accommodates many of the same features as the AR(1) model, but is better specified because the optimal number of lags are taken into consideration. This eliminates autocorrelation that is likely to exist in an AR(1) model and therefore will produce better estimates.

The success of the Box–Jenkins methodology is based on the fact that it can reflect the behaviour of diverse types of series – and it does this without requiring many parameters to be estimated in the final choice of the model. Gooijer and Hyndman, 2006 state that many techniques have been suggested for ARMA model, including Akaike's information criterion (AIC), Akaike's final prediction error (FPE), and the Bayian information criterion (BIC). Often these criteria lead to over-fitting of the in-sample model by reducing one step-ahead forecast errors.

Hepsen and Vatansever, 2010 use a standard Box-Jenkins ARIMA approach to forecasting house price trends in Dubai. Tse, 1997 examines forecasting of real estate prices in Hong Kong in a similar framework. He finds that the ARIMA model indeed is able to indicate short-term market direction. ARIMA models also do well when compared to other model classes. Nevertheless, not everyone is as enthusiastic about the forecasting ability of ARIMA models. Stevenson, 2007 warns that although ARIMA models are useful in predicting broad market trends, they differ substantially in their forecasts obtained from different model specifications. Thus, they are sensitive to model selection biases.

Sklarz *et.al,* 1987 demonstrated that a long lagged Auto Regressive (AR) process produces lower forecast error variance, unlike the Auto Regressive Integrated Moving Average (ARIMA) model when they applied the same to U.S. housing data. Due to its lower forecast error variance, the AR model is a better option when used to forecast the housing market in general, which features strong seasonality and slowly changing trends.

Vishwakarma, 2013 studied the Canadian real estate price index using the ARIMA family models. He held that all these models worked fine for short term forecasting. He also argued that in the past researchers have applied various models to explain the real estate market, from simple linear regression to advanced models such as the Vector Error Correction model (VECM), the Kalman filter, and so on. However, in the end simple models were found efficient compared to more complex ones. In his model, he used macro-economic variables such as GDP, inflation, long-term and short term bond rate and exchange rate of the Canadian dollar against the US dollar. He tried to test his models along with this econometrics. Crawford and Fratantoni, 2003 used ARIMA, GARCH, and regime switching univariate models to forecast the real estate market in various parts of the US. They used state-level repeat transactions data for California, Florida, Massachusetts, Ohio, and Texas. Annual basis growth rates at a quarterly time span are calculated from each of these indices for the period from quarter one 1979 to quarter four 2004. The study found that ARIMA models are generally more suitable for out-of-sample forecasting and point forecasts. Stevenson, 2007 applied the OLS, ARIMA, and VAR models to forecasting housing supply in the Irish market, using quarterly data from 1978 through 2003. He found that the ARIMA model had better forecasting ability than the others for the period 1998–2001, because the Irish market had a sustained housing boom beginning in the mid-1990s that ignored the fundamentals. In the absence of fundamentals, ARIMA models perform well in predicting trends. Improved forecasts not only allow for proper pricing of mortgage credit risk, thus promoting financial stability, but also help institutional investors to manage risk from mortgage-backed securities (Miles, 2008). The forecasts from the ARIMA models are adaptive to structural breaks (Clements and Hendry, 1996). However, Crawford and Fratantoni, 2003 indicate that a linear ARIMA model displays better out-of-sample forecasting of home prices than the Markov-switching and GARCH models, although the Markov-switching model is superior for the in-sample fit.

In order to estimate true movements in residential property prices, Birch and Suderman, 2003 introduced a two-way exponential smoothing system. Their method appears to still be in its infancy and seems somewhat experimental. They point out that their system seems to overcome some of the problems attached to the more rigid nature of regression modeling. However, they do not offer any conclusive evidence that their model is superior to more common hedonic price models.

2.10 Previous Research Critiques

The aim of the thesis is to provide an approach to real estate residential modeling and forecasting covering a property type’s correlation, time series attributes within a region or a city, and socio-economic attributes of preferred real estate locations. The thesis covers residential estate markets and concentrates on property types in contrast to previous studies that have considered country wide house price indices.

There have been studies covering parts of the US, Canada, China, the UK real estate market (Miller and Sklarz, 1986; Vishwakarma, 2013 and Fratantoni, 2003). The studies examined real estate price indices at the macro level as well as correlations between regions' real estate prices. We could not locate any study that pays attention to within city real estate price forecasts by property types. In other words, the studies reviewed help on a macro decision making level. Macro decisions are relevant to policy makers and real estate fund managers for predicting trends. However, no study has looked at the micro level, i.e. studying time series performance within a city and the correlation and performance by property type. Such a study can provide critical information to policy makers, developers, real estate fund managers and private investors when it comes to micro level management. This thesis attempts to fill the gap that exists in the field, as no previous studies have been found that cover modeling and forecasting real estate prices on a micro level within a city.

Brooks and Tsolacos, 2010 stated that real rents move through time in response to the vacancy rate (defined as vacant stock over total stock, expressed as a percentage) and an economic output variable. Here, the indicator is vacancy i.e. it gives a picture of demand and supply conditions. With a fall in vacancy due to high demand and supply, the market becomes ideal for landlords. Landlords during this time push for higher rents in new leases or review the existing rent. Their work implies seasonal pattern in real estate prices. As a matter of fact Brooks and Tsolacos, 1999 report a model that has been previously used by researchers including themselves to model seasonality in real estate price indices. There has not been a study that examined real estate prices within a city and by properly types using the proposed methodology of Brooks and Tsolacos, 1999.

The thesis also tries to incorporate recent developments in geographic information systems (GIS) to identify multivariate causal models that can help in real estate forecasting within a city by property types. Previous research such as that of Breedon and Joyce, 1993 has analyzed the relationship between real estate prices, earnings, disposable income, demographic factors, and the rate of repossessions by lenders. There is no study that uses geographic maps to identify causal relationships that can be used in forecasting real estate prices. Goodchild, 1989; Densham, 1989; Peterson, 1998; and Thrall and Marks, 1993 described the ability of the GIS to study the spatial aspect of real estate decisions. However, these studies are more descriptive of the spatial (geographic) attributes of a local address. They did not aim at studying the economics and social variables associated with these local addresses to drive which variable demands a higher price tag. Real estate is marked with a definite spatial boundary which is called locality. This thesis aims at using GIS to identify the potential investment locality by studying the socio-economic parameters (such as tax claimants, education of local residents, number of houses in different tax bands etc.) of locality(s). No study has used GIS to map causal relationships between real estate prices and explanatory variables like higher income, council tax claimants, good health and green space. GIS will help to identify the areas of investment and the causal study will help in determining and forecasting real estate prices. The thesis is the first to provide a framework that ties social and economic variables to the price of real estate property within a city.

2.11 Summary

A gap has clearly been identified in the literature review that is it is necessary to study price changes by property types within a city; seasonality and possible casual factors than can explain real estate price changes. Chapter four deals with research problem one which is “lack of studies of real estate price determination by residential property types within a city”. The Box-Jenkins (1970) methodology is adopted to identify a linear ARIMA (Autoregressive Integrated Moving Average) model that can help in estimating and later on forecasting real estate prices within a city and by property types. The choice of the ARIMA models was due to studies by Crawford and Fratantoni, 2003; Tse, 1997 and Stevenson, 2007 indicating that a linear ARIMA model displays better out-of-sample forecasting of home prices.

Brooks and Tsolacos, 1999 and 2010; proposed a methodology to study seasonality that was applied to real estate price indices. That led to the identification of research problem two which is “lack of studies of seasonality of real estate prices by residential property types within a city”. Brooks and Tsolacos methodology will be applied to real estate prices by property types (Flats, Terraced, Semi-Detached, and Detached) within the Manchester City.

There is no study that uses geographic maps to identify causal relationships that can be used in forecasting real estate prices. Goodchild, 1989; Densham, 1989; Peterson, 1998 and Thrall and Marks, 1993 described the ability of the GIS to study the spatial aspect of real estate decisions. Previous studies are more descriptive of the spatial (geographic) attributes of a local address. They did not aim at studying the economics and social variables associated with these local addresses to drive which variable demands a higher price tag. This led to research problem three which is “lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting”. Chapter five will handle this research problem by using multiple regression method.

Chapter 3: Methodology

3.0 Introduction

Chapter two reviewed the existing literature and identified three research problems, lack of studies of real estate price determination by residential property types within a city, lack of studies of seasonality of real estate prices by property types within a city, and lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting. A gap has clearly been identified in the literature review that is it is necessary to study price changes by property types within a city; seasonality and possible casual factors than can explain real estate price changes. Chapter three will discuss the methodology with regard to the models that are going to be used to fill the research gap as identified by the three research problems. The models used are time series models, cross sectional multiple regression models, GIS tools. The methodology is based on six steps which are statement of research problem, gathering relevant data, estimation methods, diagnostic tests for the residuals, model evaluation, and out of sample forecasting. The six steps are widely used in real estate literature and summarized by Brooks and Tsolacos, 2010. Time series analysis will be implemented in Chapter four on real estate data. Chapter five will build up the causal multiple regression models. Chapter six contains the conclusions and recommendations of the thesis.

3.1 Proposed Approach

The thesis' proposed approach for modeling and forecasting real estate prices uses the general framework suggested by Brooks, 2008 and Brooks and Tsolacos, 2010. Their suggested framework consists of six steps summarizing the common econometric steps in economics with particular emphasis on real estate sectors. The six steps will be applied to the three objectives as identified earlier in the thesis; the examination of the correlations between property types within a region or a city, property types' seasonality and time series attributes within a region or a city, and connecting GIS maps and socio economic parameters that can help in identifying the most important socio-economic variables driving house prices. The six steps are as follows:

3.1.1 Statement of research problem

The researcher formulates the model based on economic theory or expectations. The model will formulate the variable of interest (real estate prices in the present thesis) as a function of possible explanatory variables. The model will be at best an approximation as it is unlikely that it will capture every relevant piece of information that contributes to real estate prices. If the model is too complex, it will be difficult and practically irrelevant as it will become difficult to implement. On the other hand, if the model is too simple, it is unlikely to capture and explain variation in real estate prices. An acceptable model should present a good approximation that is practical to implement and at the same time does not ignore important contributing information.

3.1.2 Gathering relevant data

Once the model formulation is accomplished in stage one, the next step is gathering relevant data. Two types of data will be handled in the thesis, univariate time series and multivariate (causal) time series. The data types are interval (cardinal) meaning all types of calculations are permitted on them (Mills and Markellos, 2008). Real estate prices for example are real data which allows all types of mathematical manipulation. Other relevant variables that can explain variation in real estate prices such as household income and council tax are also real variables. Univariate time series data are data on a single variable that can be charted against time points, for example monthly real estate prices of a house over a year. As will be explained later, time series models focus on the history of the variable in attempting to form a model that can explain its future. Multivariate (causal) time series data are data on multiple variables over time, for example, annual data on real estate prices, household income and council taxes.

This thesis will focus on house price index and on property type within the real estate sector in a city. Other relevant variables that can explain variation in real estate prices such as household income and council taxes are also real variables. Univarate time series data are data on a single variable that can chart its history against time points. For example, the monthly real estate prices of a house over a year. As will be explained below, time series models focus on the history of the variable in an attempt to form a model that can explain its future. The data will represent the quarterly price changes of property types (for example, Detached; Semi-Detached; Flats and Terraced) within a city. The results of the data analysis are quarterly (in percentage), which is the mean return of the property types; equivalent yearly return (in percentage), which is calculated by multiplying the mean with the number of quarterly. i.e. four; quarterly risk (in percentage) which is the standard deviation and annual risk (in percentage) which is four multiplied by quarterly risk (Brooks, 2008). Coefficient of variation (CV) is a standard measure in finance that measures the percentage risk per a percentage return. It is a useful measure to compare different investments as they are ranked from highest CV indicating highest risk per unit of return to lowest CV indicating lowest risk per unit of return.

Sources of data for this thesis are databases which are available electronically as will be explained in detail in later chapters.

3.1.3 Estimation methods

Once the model is formulated in step one and the data required is gathered in step two, step three handles the model parameters estimation. Chapter four of the thesis deals with time series models. The Box-Jenkins, 1970 methodology is adopted to identify a linear ARIMA (Autoregressive Integrated Moving Average) model that can help in estimating and later on forecasting real estate prices. Crawford and Fratantoni, 2003; Tse, 1997 and Stevenson, 2007 indicate that a linear ARIMA model displays better out-of-sample forecasting of home prices. The choice of Box-Jenkins, 1970 over automatic model identification techniques such as Akaike's information criterion (AIC), Akaike's final prediction error (FPE), and Bayesian information criterion (BIC) is that the Box-Jenkins approach is more informative in that it helps the researcher to understand the data. The AIC, FPE and BIC treat the time series as a black box as they suggest the best fitting model without revealing much about the data generating process' characteristics. Also, these criteria often lead to over-fitting of the in-sample model (Brooks and Tsolacos, 2010). Multivariate time series analysis focuses on relationships between the dependent variable (real estate prices) and explanatory variables that can explain variation in the dependent variable such as household income and council tax.

A time series is a collation of observations of well-structured data which has been obtained through repeated measurements over time (Davis *et.al*, 1986). Vishwakarma, 2013 studied the Canadian real estate price index using the Autoregressive Integrated Moving Average (ARIMA) family of models. He compared his results with more econometric advanced models than ARIMA such as vector autoregression, co-integration and error corrections models. He found that simple models such as ARIMA are more efficient compared to more complex models. Crawford and Fratantoni, 2003 used ARIMA, GARCH, and regime switching univariate models to forecast the real estate market in various parts of the US. They used state-level repeat transactions data for California, Florida, Massachusetts, Ohio, and Texas. Annualized growth rates at a quarterly frequency were computed from each of these indices from 1979 (quarter 1) to 2001 (quarter 4). The study found that ARIMA models are generally more suitable for out-of-sample point forecasts. Stevenson, 2007 applies the Ordinary Least Squares (OLS), ARIMA, and Vector Autoregressive (VAR) models for forecasting housing supply in the Irish market, using quarterly data from 1978 through 2003. He found that the ARIMA model has better forecasting ability than the others for the period 1998–2001.

One of the main goals of time series analysis is to forecast future values of the series. Slow and regular rising pattern in the series level is called a trend. The cyclical patterns are important, especially in the context of real estate management. For example, Reed, 2002 concluded in his paper that there was distinct evidence of existence of short property cycles among 117 Brisbane Suburbs in Australia during 1957-99. Reed, 2002 first removed the trend in the data before analyzing for any cyclic price patterns using Spectral analysis on the data of each suburb. The average cycle year for all the 117 suburbs was found to be about 9.95 years with an associated standard deviation of 1.97 years. The number of short-term cycles was 56 out of a total of 117 suburbs, and the range of cycle years was found to be 7-14 years.

The simplest class of time series models that one could entertain is that of the moving average process as in Brooks and Tsolacos, 2010. Let u_t (t = 1, 2, 3, . . . , n) be a white noise process with Expected (u_t) = 0 and variance (u_t) = σ^2. Then

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} \qquad (3.1)$$

is a q order moving average model, denoted MA(q), and θ is parameter estimates for each lag up to q.

A moving average model is simply a linear combination of white noise processes, so that y_t depends on the current and previous values of a white noise disturbance term. An autoregressive model is one in which the current value of a variable, *y*, depends upon only the values that the variable took in previous periods plus an error term. An autoregressive model of order p taken from Box and Jenkins, 1970, denoted an AR(p), can be expressed as

$$y_t = \mu + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + u_t \qquad (3.2)$$

Where μ is the constant, $\emptyset_1$ is the parameter estimate for lag one of y_t and $\emptyset_2$ is the parameter estimate for lag two of y_t, u_t white noise disturbance term. Combining the AR(p) and MA(q) models, an ARMA(p,q) model is obtained. Such a model states that the current value of some series y depends linearly on its own past values plus a combination of the present and past values of the white noise error terms.

Identifying the order p and q of the ARMA(p,q) can be done by one of two methods; automatic order selection or Box-Jenkins, 1970 methodology. The automatic order selections are dominated in the literature by the Akaike Information criterion (Hirotugu, 1974, AIC) and the Bayesian information criterion (Schwarz, 1974, BIC). AIC and BIC criteria are based on information entropy, i.e. they offer a relative estimate of the information lost when a given model is used to represent the process that generates the data. Since it is possible to increase the likelihood by adding parameters, both BIC and AIC resolve this problem by introducing a penalty term for the number of parameters in the model. However, the penalty term is larger for BIC than for AIC. The problem with the AIC and BIC criteria is that they handle the data as black box and do not offer any insight about the data generating process.

The alternative to model selection criteria is the Box-Jenkins, 1970, which is based on model identification by examining the autocorrelations functions (ACF) and partial autocorrelations functions (PACF). The autocorrelation of a random process describes the correlation between values of the process at different times, as a function of the time lag. Let X represent the real estate price at time t. Then X_t is the value (or realization) produced by a given run of the process at time t. Suppose that the process is further known to have defined values for mean μ and variance σ^2. Then the definition of the autocorrelation, between times t and t-s where s is the time lag, is as follows (Box and Jenkins, 1970):

$$R(t,\ t-s) = \frac{E[(X_t-\mu)(X_{t-s}-\mu)]}{\sigma^2} \qquad (3.3)$$

Where "E" is the expected value operator. If the function R is well-defined, its value must lie in the range [−1, 1], with 1 indicating perfect positive correlation and −1 indicating perfect negative correlation.

Fuller, 1976 reported that the partial autocorrelation function (PACF denoted τ_{kk}), measures the correlation between an observation k periods ago and the current observation, after regulating for observations at intermediate lags (that is, all lags$<$k). For example, the PACF for lag 3 would measure the correlation between y_t and y_{t-3} after controlling for the effects of y_{t-1} and y_{t-2}.

Box-Jenkins, 1970 methodology shows how different patterns of ACFs and PACFs can help in identifying an initial ARMA(p,q) model. The model parameters are then estimated using maximum likelihood using the BHHH algorithm of Berndt *et. al*,1974.

The overall significance of the model is tested using Wald Chi-squared (Harrell, 2001). The residuals from the model are then examined for no patterns in their ACFs and PACFs. Also, the portmanteau Q-test of Ljung and Box, (1978) is applied to test for the significance of the accumulation of the ACFs lags up to a certain cut-out point. Once the model has been passed for no ACFs and PACFs patterns, normal probability tests such as W-test of Shapiro–Wilk (1965) and Q-Q test of Seber (1977) are applied on the residuals. If the model shows serious signs of not conforming to the assumptions such as no ACFs and PACFs in the residuals as well as normality of the residuals, the whole model identification process is repeated again till a satisfactory model has been reached.

Since this thesis handles quarterly data, there is the potential that the ARMA(p,q) did not fully catch seasonality in the data. As a matter of fact seasonal ARMA models can be fit. However, in this thesis the approach of Brooks and Tsolacos, 1999 is adopted where a seasonal model is used as a competing forecasting model to ARMA. Brooks and Tsolacos, 1999 suggested a very simple method for coping with seasonality. In the context of our quarterly data, the model of Brooks and Tsolacos, 2010 is as follows.

$$y_t=\beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + u_t \quad (3.4)$$

where,

y_t= quarterly real estate price;

β= parameter estimates;

D_1=dummy variable 1 in quarter 1, and 0 in the other three quarters;

D_2=dummy variable 1 in quarter 2, and 0 in the other three quarters;

D_3=dummy variable 1 in quarter 3, and 0 in the other three quarters;

D_4=dummy variable 1 in quarter 4, and 0 in the other three quarters;

u_t= residual term with normal distribution

The above model will be estimated for quarterly data using OLS as well as the identified ARMA(p,q) model. The forecasting abilities of those models will also be compared to the cross sectional causal models that will be built in chapter five after identifying their relationships.

Another approach to forecasting real estate prices is to establish causal relationships between the dependent variable which is real estate prices and the independent explanatory variables that can explain changes in real estate prices such as income, inflation, GDP etc. The main model of interest here is the simple and multiple regression models. This section provides the basis for the causal models that will be built in chapter five.

The population regression function (PRF) is a function which generates the actual data and it shows the actual relationship among the variables. Another name for PRF is date generating process (DGP), which is shown in equation 3.5 (McGough *et.al*, 2000).

$$y_t = \alpha + \beta x_t + u_t \tag{3.5}$$

In the above equation there is a noise term so that, even if one had in one's inventory the entire population of observations on x and y, it would still be not possible to obtain a perfect fit of the line to the data. α and β are the population true parameters that need to be estimated using sample data. Their equivalent in sample data will be called a and b. McGough *et.al,* 2000 used simple regression equation to model the relationship between the growth in real rents at time *t* as the dependent variable (RRg_t) and the growth in employment in financial and business services at time *t* ($EFBSg_t$). The regression was applied on real office rent and employment in UK financial and business service annual data starting 1979 and ending 2005 giving twenty seven annual observations. The model estimated is in equation 3.6.

$$RRg_t = \alpha + \beta\, EFBSg_t + \mu_t \tag{3.6}$$

Where RRg_t is the growth in real rents at time t and $EFBSg_t$ is the growth in employment in financial and business services at time t. Their estimate for β is +3.27 indicating that when employment growth is positive, real rental growth would be positive too.

Another application of simple regression was carried out on the Helsinki market by McGough *et.al,* 2000 using annual data for the period from 1970 to 1998. The model is represented in equation 3.7.

$$\Delta OFFRET_t = \alpha + \beta \, \Delta GDP_t + \mu_t \qquad (3.7)$$

Where ΔOFFRET is the dependent variable and denotes the change in office rents from one period to another. OFFRET is an index of office returns adjusted for inflation. The explanatory variable is Gross Domestic Product at constant prices adjusted after inflation (GDP).

The empirical estimation of equation 3.7 resulted in the following estimates as shown in equation 3.8.

$$\Delta \widehat{OFFRET}_t = -24.5 + 20.0\Delta(\widehat{GDP_t}) \qquad (3.8)$$

The positive sign for the parameter estimate for ΔGDP indicates that a positive change in ΔGDP results in positive change in office rent returns and vice versa. On average, over the sample period, a change in the GDP index by one unit would result in a change of twenty units in the index of real returns.

The causal model's main statistical tests that are used to determine whether the relationship is significant and not due to random chances can be summarized in the following tests.

The simple regression model introduced can be expanded to include extra explanatory variables (x's). Amy *et.al,* 2000 studied the Singapore office market. Their empirical analysis used quarterly data and included the estimation of different specifications for rents. The model estimated is given by the following equation:

$$\%\Delta R_t = \beta_0 + \beta_1 \%\Delta E_t - \beta_2 V_{t-1} \qquad (3.9)$$

where %Δ denotes a percentage change (over the previous quarter), R_t is the nominal rent (hence $\%\Delta R_t$ is the percentage change in nominal rent this quarter over the preceding one), E_t is the operating costs (as for example consumer price index, CPI) and V_{t-1} is the vacancy rate (in percent) in the previous quarter. Adjusted R-squared is 23% meaning that the changes in explanatory variables have explained only 23% of the variation in the dependent variable. The relationship between the previous quarter vacancy and the current quarter change in rent is negative meaning if previous vacancy is high the current rent is low. The relationship between the changes in operating costs is positively related to the change in rent meaning a higher operating cost leads to higher rents.

<u>3.1.4 Diagnostic tests for the residuals</u>

As explained in step one, the model formulated is at best an approximation of the data generating process. There are assumptions behind the model that need to be met. The model assumptions are tested and if they were not met, a new model specification is tried out till assumptions are met. The time series models' assumptions are:

1. The residuals (unexpected error) that the model could not capture are linearly independent from its own values over previous points in time,

2. The residuals squared do not show signs of linear dependence on their past values,

3. The residuals are normally distributed.

The logic behind assumption one is that if there is any pattern left in residual, this will be an indication that a different model specification should be used to remove the predictable pattern from the residuals. The logic behind assumption two is to check if there is any linear dependence in the squared residuals of the process. The logic behind assumption three is to ensure that statistical tests carried out are valid as the residuals are assumed to be normally distributed and therefore all statistical tests are built on that assumption. It is more critical here to examine the extent to which the residuals distribution departs from normality. This will determine how much credibility we can associate with the model parameter's statistical tests.

Geographic Information System (GIS) analyses the data of domains such as the environment, social grades, housing, health care, economic deprivation etc. The data are found in the website of the Office for National Statistics (ONS). These domains contain the data of each of the explanatory variables as per respective MSOA (middle super output area, explained in Chapter five). These attributes are then linked with the area maps (mentioned in 5.1) through the ArcGIS software to generate thematic maps of each Middle Super Output Areas. A thematic map is a type of map especially created to portray a particular theme related with a specific geographic area. These maps will help to visualize physical, social, political, cultural, economic, sociological or any other aspects of a city. The symbols and the legends for the thematic maps will be created by the default symbols and legend-shades present in the software. Each of the thematic maps will give us an entire picture of the full MSOA. The maps help to visualize the area remotely i.e. without actually going to the site one can see the picture of not only a particular locality but also localities surrounding the same.

<u>3.1.5 Model evaluation</u>

Once the model adequacy has been accepted in step four, the next step is to evaluate the model parameters' estimates in line with the economic theory or expectations as formulated in step one. If the model does not pass this step, steps one to three are repeated. If the model passes, then we move to the last step which is forecasting of out of sample data that is available (Box and Jenkins, 1970). Adjusted R^2 which explains how much of the variation in the dependent variable y can be explained by the variation in the independent variable x. Its boundaries are 0 and 1 where a value close to one highlights a more significant relationship while a value close to zero identifies no relationship. In the case of simple regression Adjusted R-squared can be replaced with the R-squared. Adjusted R-squared takes into account that adding more explanatory variables can increase R-squared due to random chances and therefore gives a penalty for including more explanatory variables. Adjusted R-squared (Lomax, 2007) is defined in the following equation:

$$\bar{R}^2 = 1 - \left[\frac{T-1}{T-k}\,(1 - R^2)\right] \qquad (3.10)$$

Where k is the number of parameters to be estimated in the model and T is the sample size. If an extra regressor (variable) is added to the model, k increases and unless R^2 increases by a more than offsetting amount, $\bar{R}^2$ will actually fall. Hence $\bar{R}^2$ can be used as a decision making tool for determining whether a given variable should be included in a regression model or not, with the rule being: include the variable if $\bar{R}^2$ rises and do not include it if $\bar{R}^2$ falls.

F-test is used to determine the overall significance of the Adjusted R-squared, hence the overall significance of the model. The F-test statistic is the ratio of the mean squared of explanatory and mean squared error of residuals, that is, F = (Mean Squared (x_i))/(Mean Squared (residuals)). Each F-statistic has an F distribution, with the numerator degrees of freedom (DF) value for the corresponding term, and the denominator DF are $n – p$ where n is the number of observations, and p is the number of coefficients in the model (Lomax, 2007).

The overall model can be significant according to F-test but some of the explanatory variables can be insignificant and need not to be included in the model. T-tests are used to test the significance of the individual parameters. The t-test (Lomax, 2007) is shown in equation 3.11.

$$\text{Test statistics} = \frac{\hat{b}-\beta}{SE(\hat{b})} \qquad (3.11)$$

The calculated t-test is compared with the critical t-statistic at say the 5% level of significance and for particular (n-2) degrees of freedom when n is the sample size. If the calculated t is lower in absolute terms than the critical, the null hypothesis of no relationship is not rejected. If the calculated t is larger in absolute terms than the critical, the null is rejected in favour of the alternative hypothesis that there is a significant relationship.

3.1.6 Out of sample forecasting

Once the model has passed step five, the model is used for out of sample forecasting if the data is available. In the univariate time series section of the thesis, out of sample forecasting is implemented since data is available for long time periods. However, in the multivariate section, out of sample was not feasible since data is only available on an annual basis. The naïve model (Stevenson, 2007) is in equation 3.12.

$$\hat{y}_t = y_{t-1} \qquad (3.12)$$

The naïve model assumes that best forecast value for the next quarter is equal to the previous observed quarter value.

The forecasting accuracy over the naïve model (Stevenson, 2007) is judged by the root mean squared error which is shown in equation 3.13:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \tag{3.13}$$

The ARMA estimated model will be accepted if it beats out the naïve model out of sample.

The above steps are shown below in a flow diagram:

Statement of Research Problem

↓

Gathering Relevant Data

↓

Estimation Methods

↓

Diagnostic Test for the Residuals

↓

Model Evaluation

↓

Out of Sample Forecasting

Figure 3.1: Six Steps Methodology

3.2 Summary

Chapter three presented the six step approach applied in the thesis to answer the three research problems identified in chapter two. The six steps are statement of research problem, gathering relevant data, estimation methods, diagnostic tests for the residuals, model evaluation, and out of sample forecasting. The six steps will be used in each of the three research problems identified in chapter two, lack of studies of real estate price determination by property types within a city, lack of studies of seasonality of real estate prices by residential property types within a city, and lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting. The main estimation models that will be used in the empirical chapters of the thesis are Box-Jenkins time series models and causal models including geographic information systems

(GIS). Time series models and Box-Jenkins methodology has been presented and discussed. The methodology will be applied in chapter four on data for the UK all index in addition to Manchester. The choice of Manchester is based on its ranking as the second city in the UK after London according to the methodology used by the office for national statistics ONS. This thesis is the first to examine property types within a city while all other research on the UK market that was reviewed in chapter two concentrated on the correlation and time series properties between different regions in the UK. Seasonal time series models will also be estimated in chapter four. Causal models that emphasize relationships between dependent and explanatory variables were presented and discussed. The causal relationships for Manchester will be examined using Geographic Information Systems (GIS) and socio-economic variables of the ONS. Chapter six will conclude the thesis providing insight for future research.

Chapter 4: Real Estate Time Series Model Building and Testing

4.0 Introduction

Chapter three examined the six step methodology to be adopted in this thesis. The steps are statement of research problem, gathering relevant data, estimation methods, diagnostic tests for the residuals, model evaluation, and out of sample forecasting. Chapter four deals with thesis research problems one and two which are lack of studies of real estate price determination by property types within a city and lack of studies of seasonality of real estate prices by property types within a city. The chapter deals with objective one which is to *investigate the time series properties of different property types (Flats, Terraced, Semi-Detached, and Detached) and compare their time series characteristics with each other as well as the UK all house price index.* This objective aims at building appropriate statistical models based on time varying parameters that can help in improving real estate forecasts beyond that of simple naïve or seasonal models. The contribution of this chapter is in applying the methodology and models by property types and within a city in contrast with previous research studies that examined market indices.

Step one in addressing the problem statement is to find an approximate model for real estate prices that beats naïve model forecasting based on averaging. Step two in the methodology is based on gathering relevant data. The data are for the UK all real estate index in addition to Manchester. The choice of Manchester is based on its ranking as the second city in the UK after London according to the methodology used by the office for national statistics ONS (2011). The data were obtained from the UK land registry from 1995 (Quarter 1) to 2011 (Quarter 1). Data for the UK is the all index real estate price. The data for Manchester are categorized into four property types: Semi-Detached, Detached, Flat and Terraced.

Step three estimation method is based on the Box-Jenkins (1974) methodology to model real estate prices in order to generate price forecasts (step six). This Chapter will also use different diagnostic checks (steps four and five) for the residuals from the built models to assess for model adequacy and evaluate the parameter estimates according to economic theory.

All the analysis in this Chapter will be on the percentage change in the real estate house prices since they are non-stationary as it well known in the literature (Brooks and Tsolacos, 2010). Equation 4.1 shows the price change Y_t as used in the analysis for respective series.

$$Y_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{4.1}$$

If a time series house price percentage change follows an ARMA(1,1) process, the equivalent process for its price is ARIMA(1,1,1). This means that there is a one differencing in price needed to obtain a covariance stationary process.

All Figures produced in this Chapter for different house types in comparison with the UK house price index are scaled on the basis of 100 given to the first quarter of 1995. This is to standardize the units of measurements between the data.

4.1 Data

The data used in the thesis have come from four sources (Nationwide Building, UK Land Registry, Office for National Statistics, Official Labour Market Statistics). These are explained below.

4.1.1 UK All Price Index

Quarterly prices for the UK all price index from first quarter 1952 to first quarter 2011 were obtained from Nationwide Building Society (www.nationwide.co.uk). Nationwide Building Society has a huge data repository that goes back to the year 1952. This data is used in the time series analysis in chapter four.

4.1.2 Property Types Price Data

The quarterly price data by property types (Flats, Terraced, Semi-Detached, and Detached) for Greater Manchester (UK) during the period from quarter 1 of 1995 to quarter 1 of 2011 were obtained from Land Registry of the UK (www.landregistry.gov.uk). This data is used in the time series analysis in chapter four. The Manchester City was chosen on the basis that it is the second most populous conurbation as per office for national statistics ONS (2011). A conurbation is defined and formed by ONS when cities and towns expand sufficiently that their urban areas join up with each other. The first according to ONS (2011) is Greater London. Manchester was chosen rather than London since there are more foreign influences on London than Manchester which increases the randomness of the residual term and makes forecasting more difficult as more international factors have to be studied.

4.1.3 Neighbourhood Data

Using the database of the ONS, we will be able to group these causal factors into respective domains which are mentioned below. The Office for National Statistics (ONS), United Kingdom uses seven domains to paint a broad picture of an area to determine the surrounding conditions (the data on the seven domains is available at www.ons.gov.uk, and the database used in this Thesis is for 2011). The domains measure the general 'health' and determine the success localities across the city.

The seven domains are:

- Crime and Safety
- Economic Deprivation
- Health Care
- Housing
- Personal Consumer Debt
- Social Grade
- Environment

The inputs for the above domains are updated at least annually.

4.1.4 Data for Explanatory (independent) Variables

The data for the causal model to explain real estate price changes were collected from NOMIS (Official Labour Market Statistics), UK. They were only available on an annual basis. The variables are change in employment, change in income, inflation and change in council tax. These data were collected from 1998 to 2013.

4.2 Model Building for the UK All House Price Index

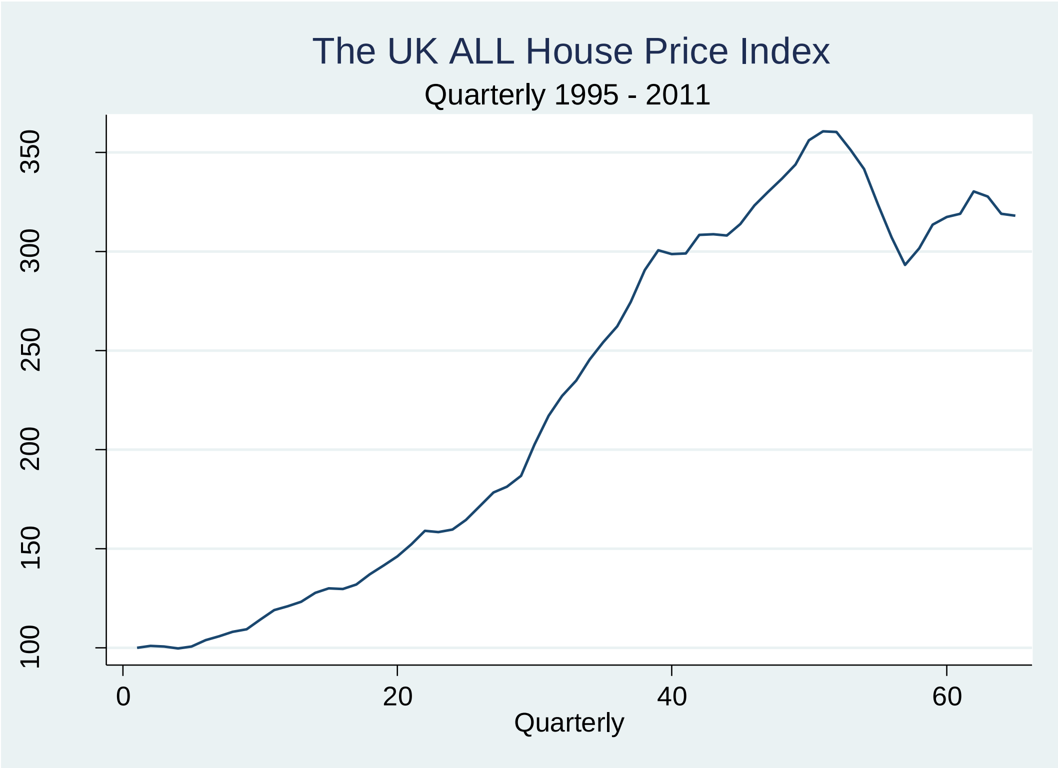

Figure 4.1: The UK ALL House Price Index

Figure 4.1 shows that the UK all house price index is non-stationary as expected and that it needs a differencing of order one. This differencing is what we call percentage price change and is plotted in Figure 4.2.

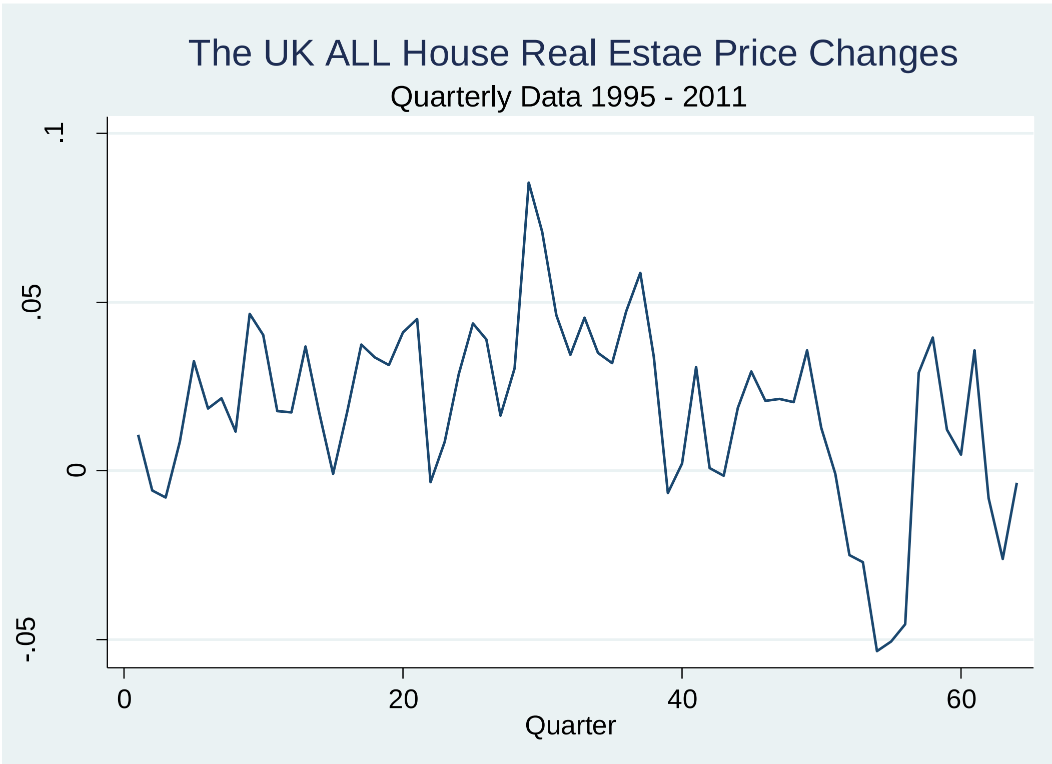

Figure 4.2: The UK ALL House Real Estate Price Changes

Figure 4.2 shows that the highest peak of 0.085 happened in quarter two, 2002, and lowest trough of -0.053 happened in quarter three, 2008. The autocorrelations functions are shown in Figure 4.3.

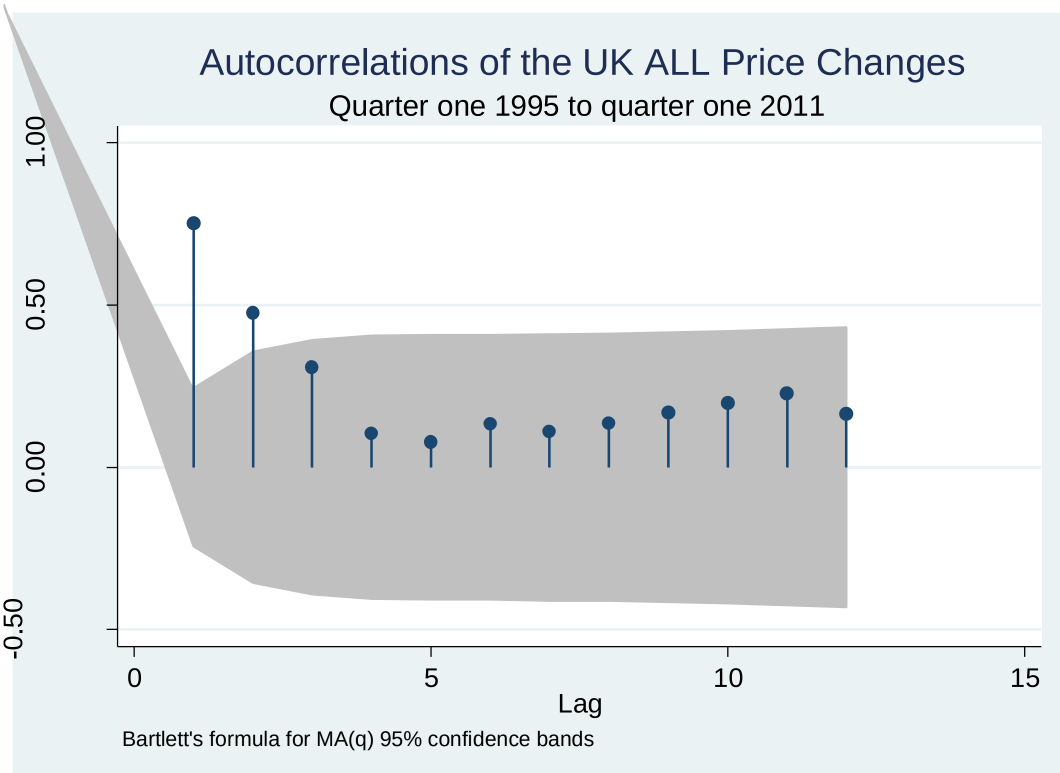

Figure 4.3: Autocorrelations of the UK ALL Price Changes

The autocorrelations are significant at lags one and two at the 5% level.

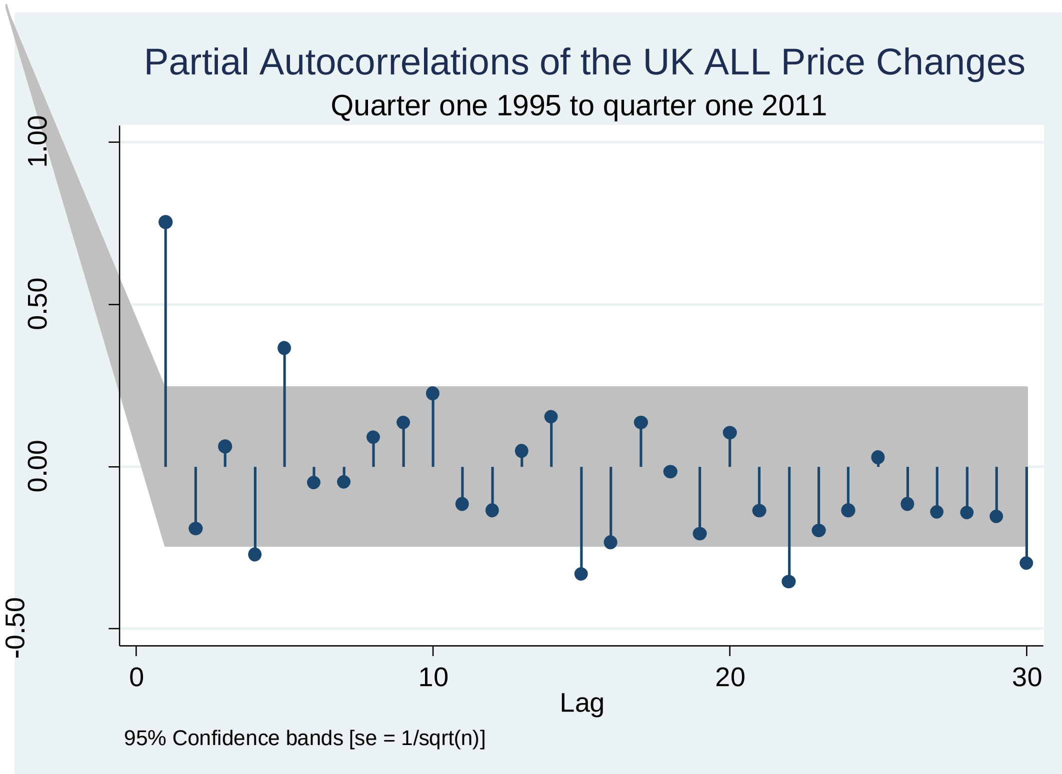

Figure 4.4: Partial Autocorrelations of the UK ALL Price Changes

The partial autocorrelations are significant at different lags at the 5% level. The candidate model is an ARMA(1,1). The model is estimated and reported in the table below.

Table 4.1: Parameter estimates for the ARMA(1,1) model for the UK ALL Real Estate Price Changes

Sample:1-65			Number of Observations	65
			Wald chi2(2)	93.77
Log likelihood = 178.1827			Prob > chi2	0
	Co-efficient	**Std.Err.**	**z**	**P>|z|**
μ	0.0175297***	0.0060233	2.91	0.004
$ø_1$	0.4943079***	0.1386969	3.56	0
θ_1	0.5713402***	0.1289325	4.43	0

*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%.

The parameter estimates for the ARMA(1,1) are highly significant at the 1% level. The overall model is significant at the 1% level as judged by the Wald Chi-squared. The residuals from the ARMA(1,1) are used as dependent variables in the multiple regression for seasonality where the independent variables are four dummy variables representing quarters one, two, three and four. The results are in Table 4.2.

Table 4.2: Results of the Regression of the residuals from ARMA (1,1) of the UK ALL House Price Change on four dummy variables representing the four quarters.

Source	**SS**	**df**	**MS**	**No.Of Obs.**	65
Model	0.00008295	4	0.000020738	**F(4,61)**	0.08
Residual	0.015942222	61	0.000261348	**Prob>F**	0.9884
Total	0.016025172	65	0.000246541	**R-Squarred**	0.0052
				Adj R-Squarred	-0.0601
				Root MSE	0.01617
	Coefficient	**Std. Err.**	**T**	**P> I t I**	
Dummy1	-0.0008674	0.0039209	-0.22		0.826
Dummy2	0.0018035	0.0040416	0.45		0.657
Dummy3	0.0001326	0.0040416	0.03		0.974
Dummy4	-0.0010559	0.0040416	-0.26		0.795

The results in Table 4.2 results indicate that there is no quarter seasonality in UK ALL House Prices as the overall model is not significant as judged by the F-tests. Also, none of the parameters are significant according to the t-tests. Figures 4.5 and 4.6 have the autocorrelations and partial autocorrelations estimates for the residuals of the ARMA(1,1) model.

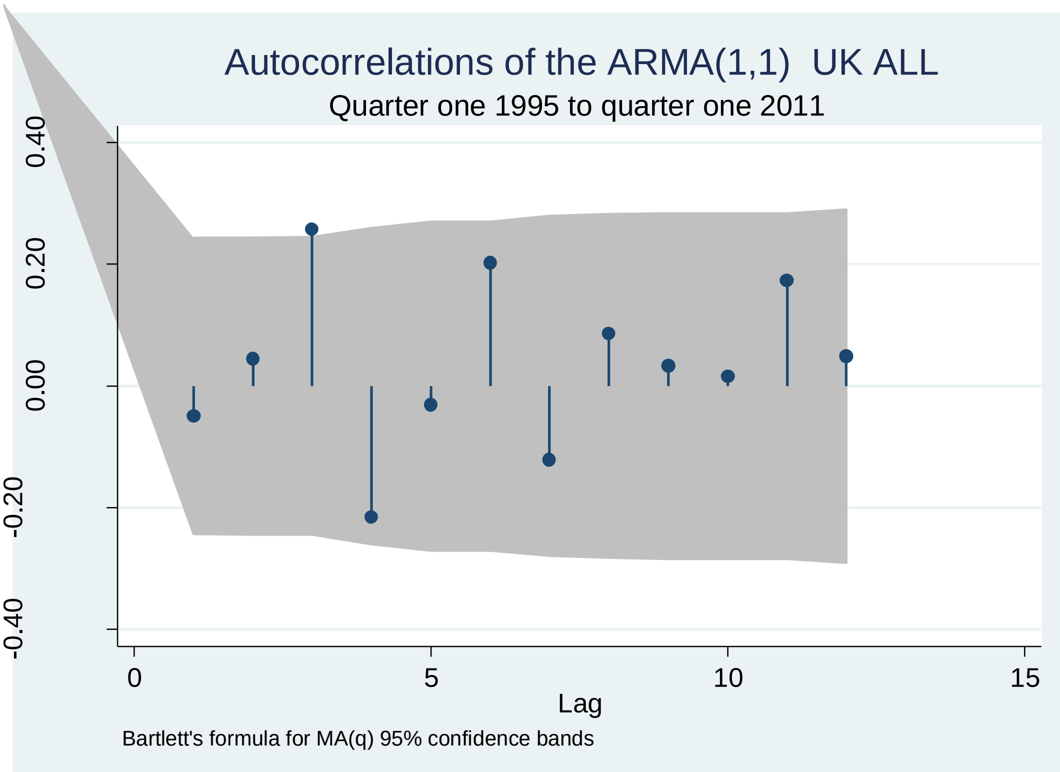

Figure 4.5: Autocorrelations of the Residuals from ARMA (1,1) UK ALL House

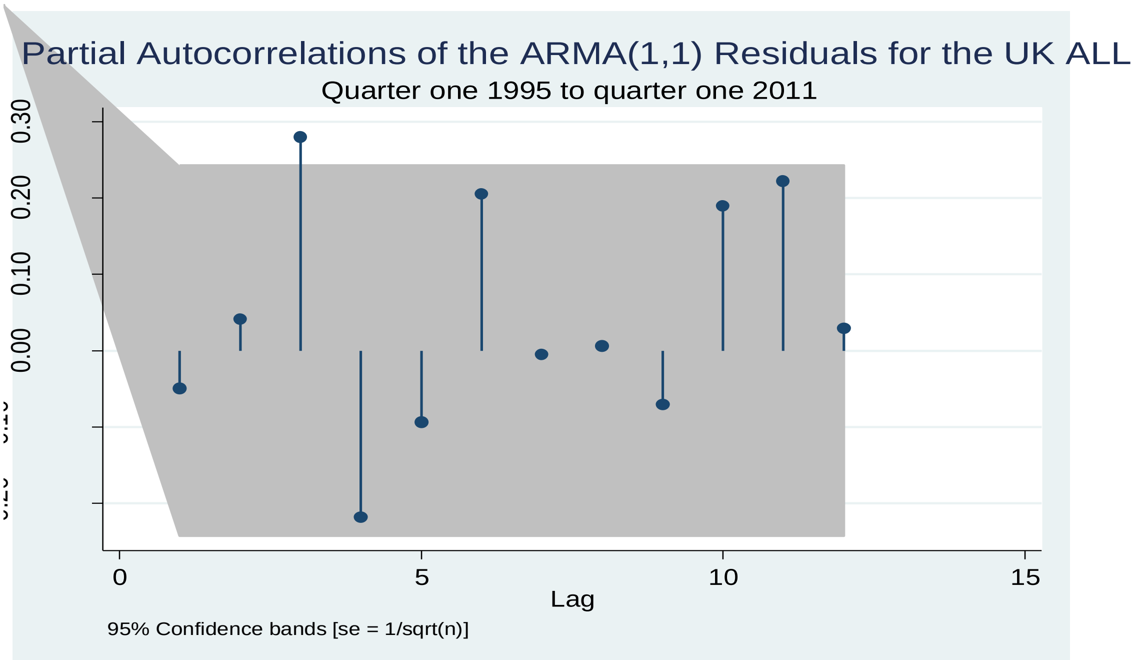

Figure 4.6: Partial Autocorrelations of the ARMA (1,1) Residuals for the UK ALL House

The ACFs in Figure 4.5 and the PACFs in Figure 4.6 do not show any significant pattern indicating the adequacy of the ARMA(1,1) model. The ARMA(1,1) model has passed the seasonality test as well as the autocorrelations and partial autocorrelations tests. Figure 4.7 has the Q-Q normal plot of the residuals from the ARMA(1,1) model.

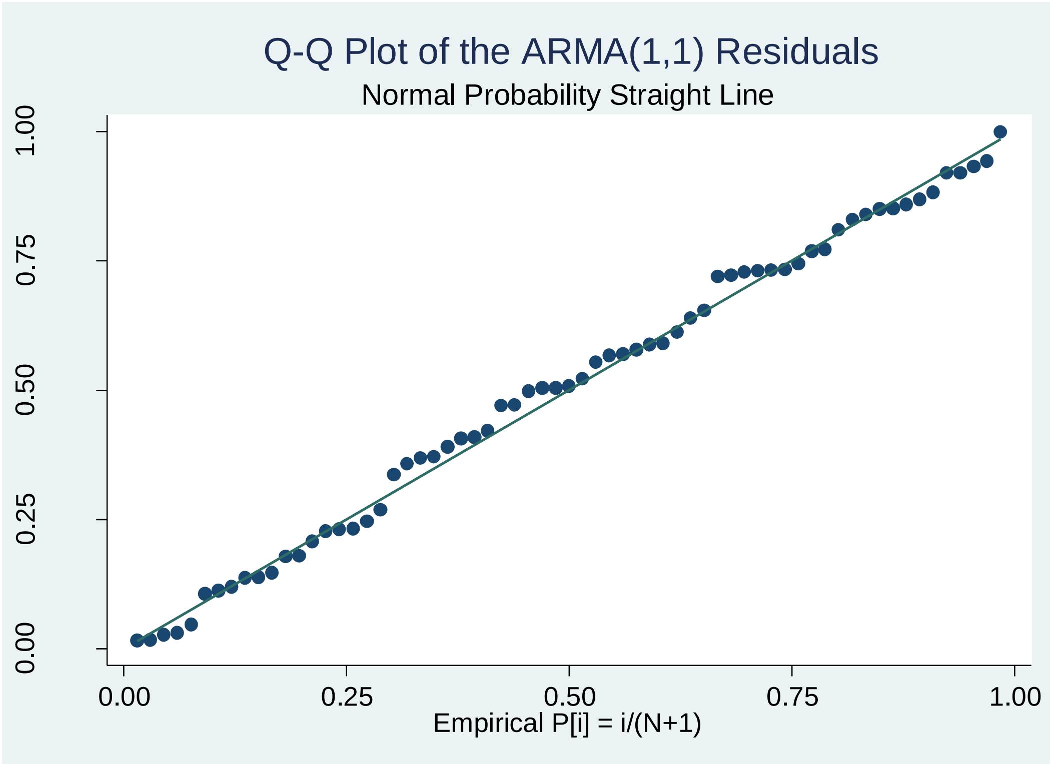

Figure 4.7: Normal Probability Plot of ARMA(1,1) Residuals

The residuals also pass the normality tests as indicated by the QQ plot in Figure 4.7, and also pass the formal tests of the Shapiro-Wilk and the Jarque Bera tests for normality. Tables 4.3 and 4.4 have the parameter estimates for the model along with their significance levels.

Table 4.3: Shapiro Wilk W Test for Normal Data

Variable	**Obs**	**W**	**V**	**Z**	**Prob>z**
Residuals	65	.98181	1.054	.114	.45446

Table 4.4: Skewness Kurtosis Test for Normality

Variable	**Obs**	**Pr (Skewness)**	**Pr (Kurtosis)**	**adj chi2 (2)**	**Prob>chi 2**
Residuals	65	.8431	.3189	1.07	.5869

The ARMA (1,1) model takes the following form as in equation 4.2.

$$y_t = \mu + \emptyset_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t \quad (4.2)$$

where,

y_t= quarterly real estate price; μ is the constant, $\emptyset_1$ is the autoregressive estimate, θ_1 is the moving average estimate and ε_t is the white noise.

4.3 Model Building For Semi-Detached

Figure (4.8) has the price pattern for the Semi-Detached versus the UK all house prices quarterly from quarter one 1995 to quarter one 2011.

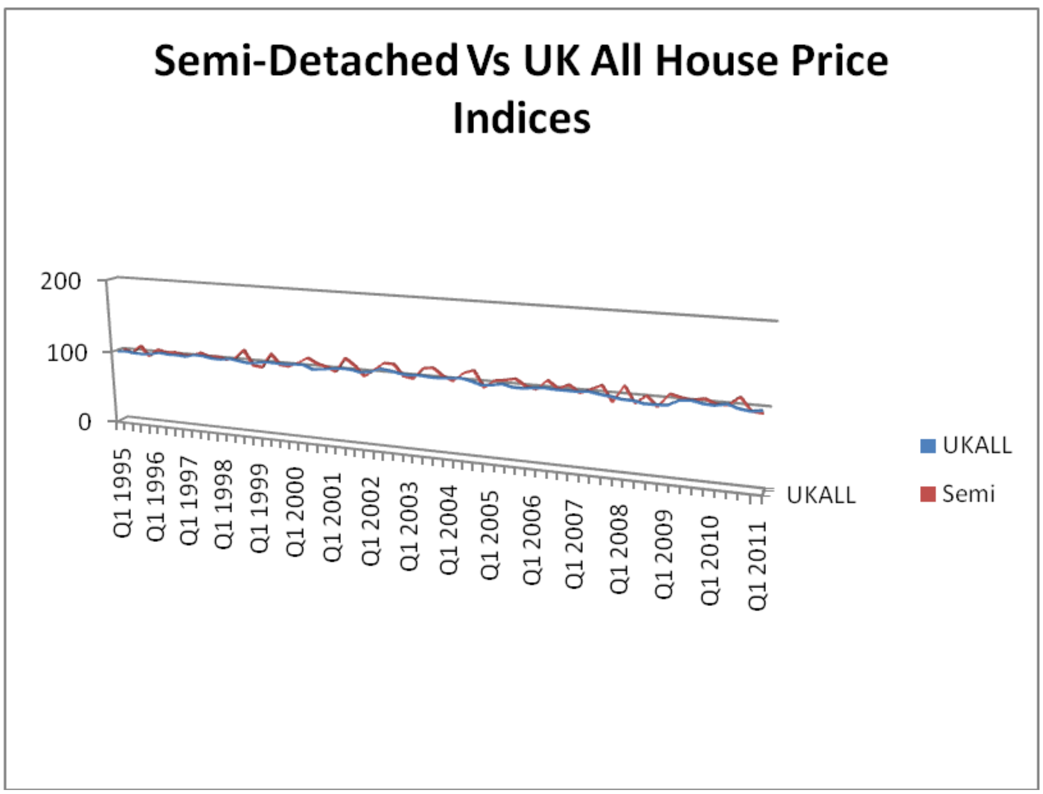

Figure 4.8: Semi-Detached Vs UK All House Price Indices

The blue line represents the UK All house price index while the red line represents the Semi-Detached index over time. The Semi-Detached price for Manchester is more volatile than the UK All house price. There has been a sharper fall for the Semi-Detached index compared with the UK All house price index during the credit crunch of 2008.

Figure 4.9 plots the Semi-Detached price changes against time from quarter two 1995 to quarter one 2011.

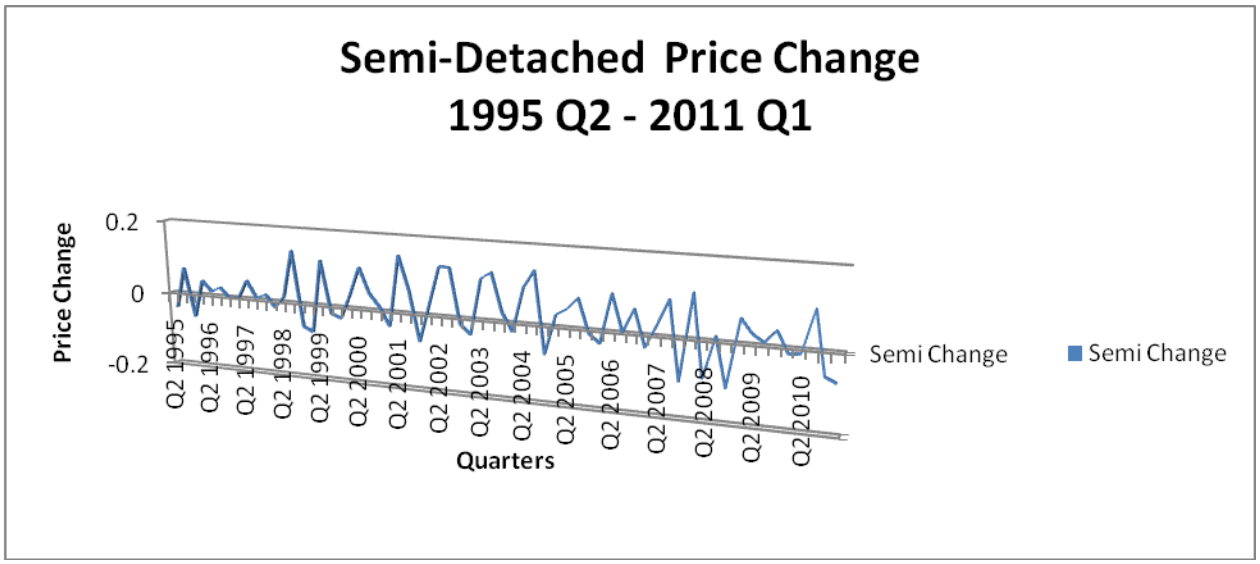

Figure 4.9: Semi-Detached Price Change

The above Figure represents the price change for Semi-Detached from the period 1995 Q2 – 2011 Q1. From the figure, it is clearly seen that the price change is volatile in the above mentioned time period. This price change can be categorized into three scales: low volatility, medium volatility and high volatility. In the above figure, low volatility is marked between the period 1995 Q2 – 2001 Q2; the medium volatility is marked between 2002 Q2 – 2007 Q2 and the high volatility is marked between 2008 Q1-2011 Q1.

Figure 4.10 represents the histogram for Semi-Detached prices along with the theoretical normal distribution imposed.

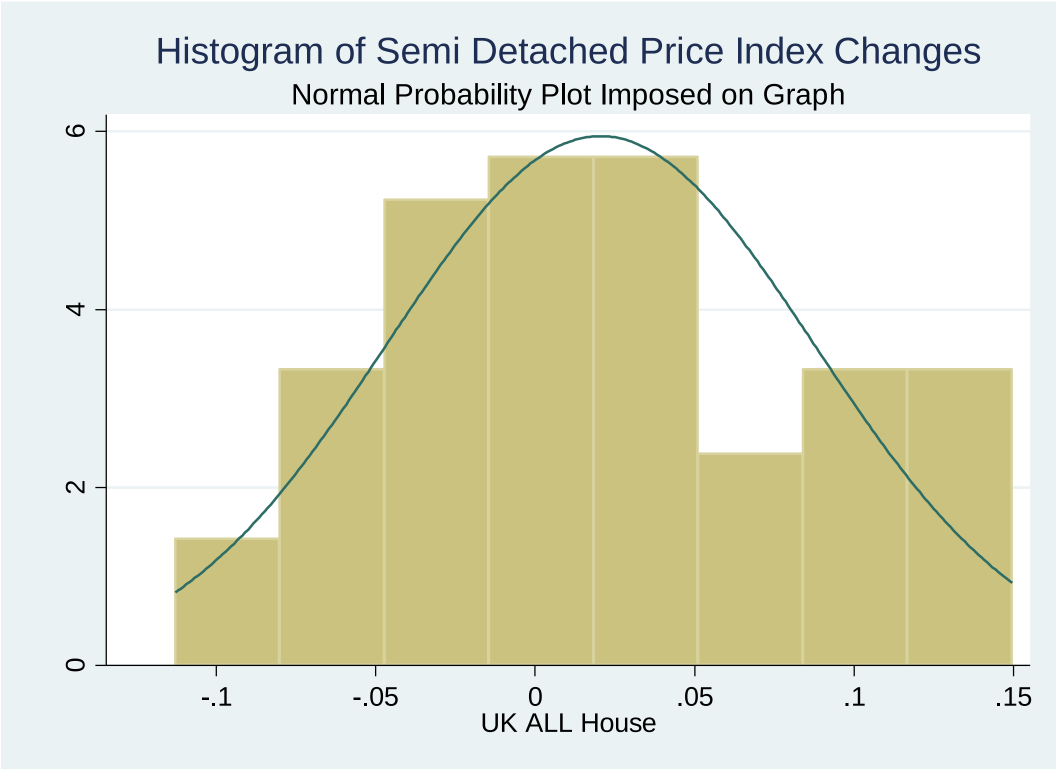

Figure 4.10: Histogram of Semi-Detached Price Index Changes

The above Figure shows the distribution to be bimodal where there are two peaks, one in the middle of the distribution between -.05 and 0.05 and the other in the right tail of the distribution. The distribution is right skewed with a skewness coefficient of 0.087 and has excess kurtosis compared with the normal distribution as evidenced by a kurtosis coefficient of 2.24. The mean is 0.020, which is greater than the median of 0.015. The standard deviation is 0.07. The distribution is not normally distributed. We next examine the Autocorrelations (ACF) and Partial Autocorrelations (PACF) Functions for Semi-Detached price changes.

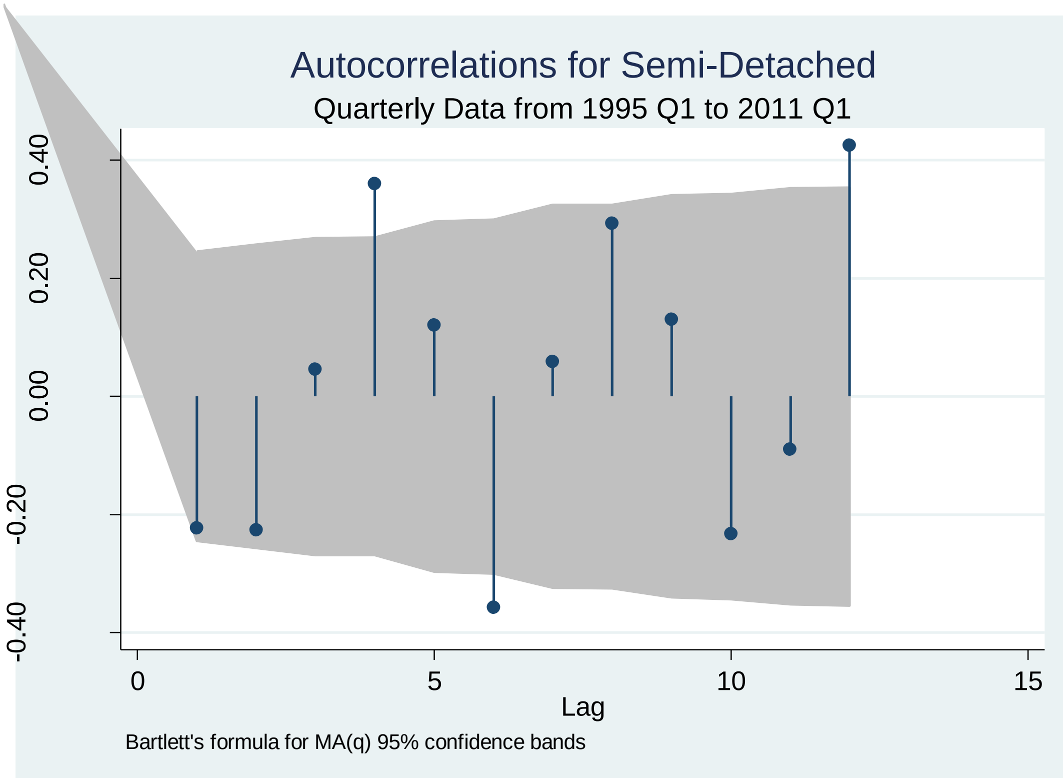

Figure 4.11:Autocorrelations for Semi-Detached

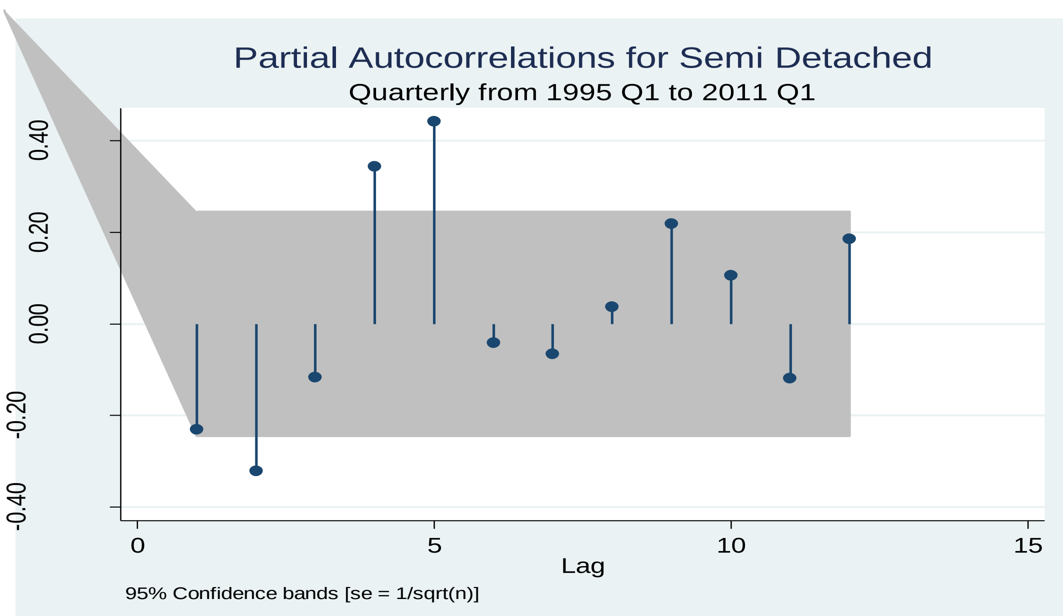

Figure 4.12: Partial Autocorrelations for Semi-Detached

Examining the ACFs and PACFs in Figures 4.11 and 4.12 for the Semi-Detached using Box-Jenkins methodology, ACFs in Figure 4.11 identify significant lags at 4, 6 and 12. The confidence intervals shown in gray in the Figure corresponds to the 95% level. The PACFs in Figure 4.12 identify significant lags at the 95% level at 2, 4, and 5. An ARMA(1,1) is fitted using the BHHH (1974) log likelihood maximization with the assumption of normally distributed residuals with a mean of zero and variance of 1. The results are in Table 4.5. The number of observations are sixty four.

Table 4.5: ARMA(1,1)Model for Semi-Detached

Sample:1-64			Number of Observations=64	
			Wald chi2(1)=5.77	
Log likelihood = 85.71225			Prob > chi2 = 0.0560	
Parameters	**Co-efficient**	**Std.Err.**	**z-test**	**P>|z|**
μ	0.021	0.005	4.1	0
ϕ_1	0.077	0.424	0.18	0.856
θ_1	-0.405	0.421	-0.96	0.337

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero. *** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%

The Wald Chi-square statistic is not significant at the 5% level indicating inadequate model. The parameter estiamtes for the ARMA(1,1) are not signifcantly different from zero at the 5% level.

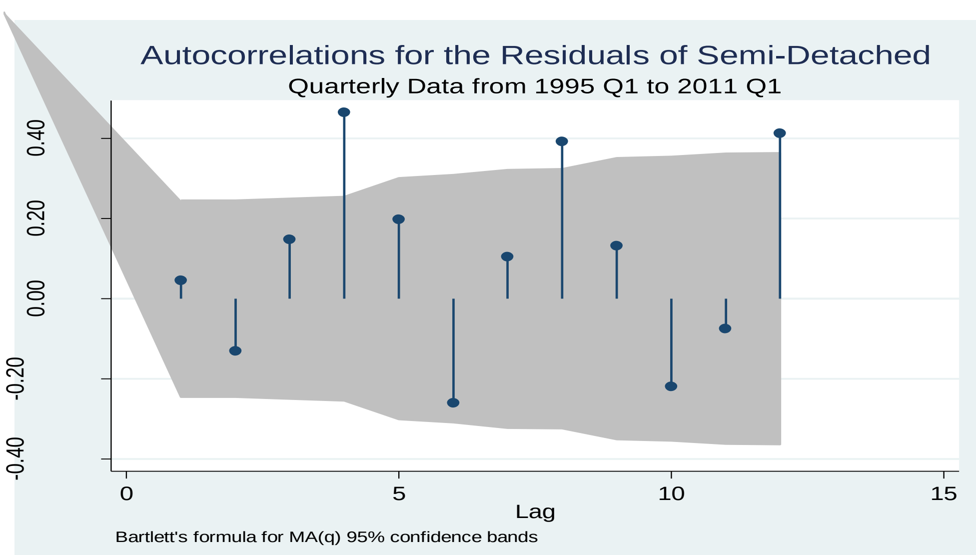

Figure 4.13: Autocorrelations for the Residuals of Semi-Detached

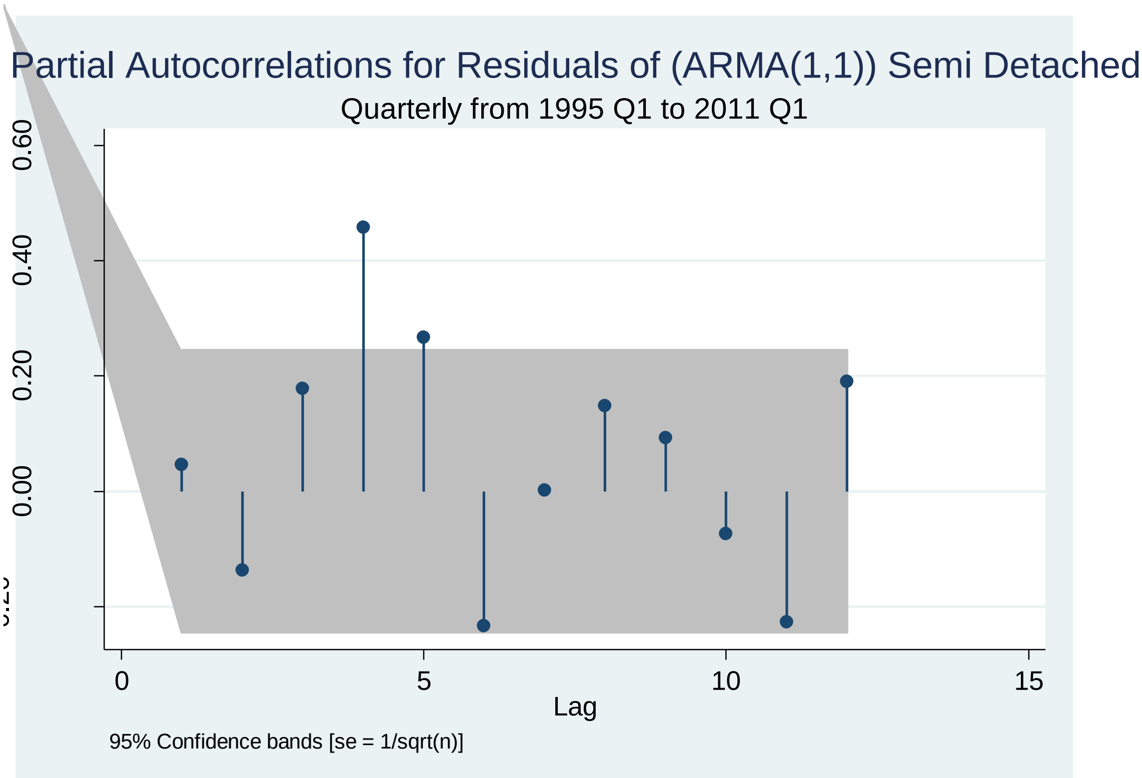

Figure 4.14: Partial Autocorrelations for the Residuals of ARMA (1,1) of Semi-Detached

A different model specification is required as the ACFs and PACFs of the residuals from the ARMA(1,1) as shown in Fiqures 4.13 and 4.14, which still indicate significant lags. An ARMA (1,2) is fitted next. The results are in Table 4.6. Also, Figure 4.15 plots the ACFs for the residuals from the ARMA (1,2).

Table 4.6: ARMA(1,2) Model for the Semi-Detached

Sample:1-64 Log likelihood = 89.46317			Number of Observations=64 Wald chi2(2)=227.31 Prob > chi2 = 0.0000	
Parameters	**Co-efficient**	**Std.Err.**	**z**	**P>|z|**
μ	0.017	0.011	1.51	0.131
$\o_1$	0.916***	0.092	9.89	0
θ_1	-1.51***	0.119	-12.69	0
θ_2	0.640***	0.116	5.49	0

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero.
*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%.

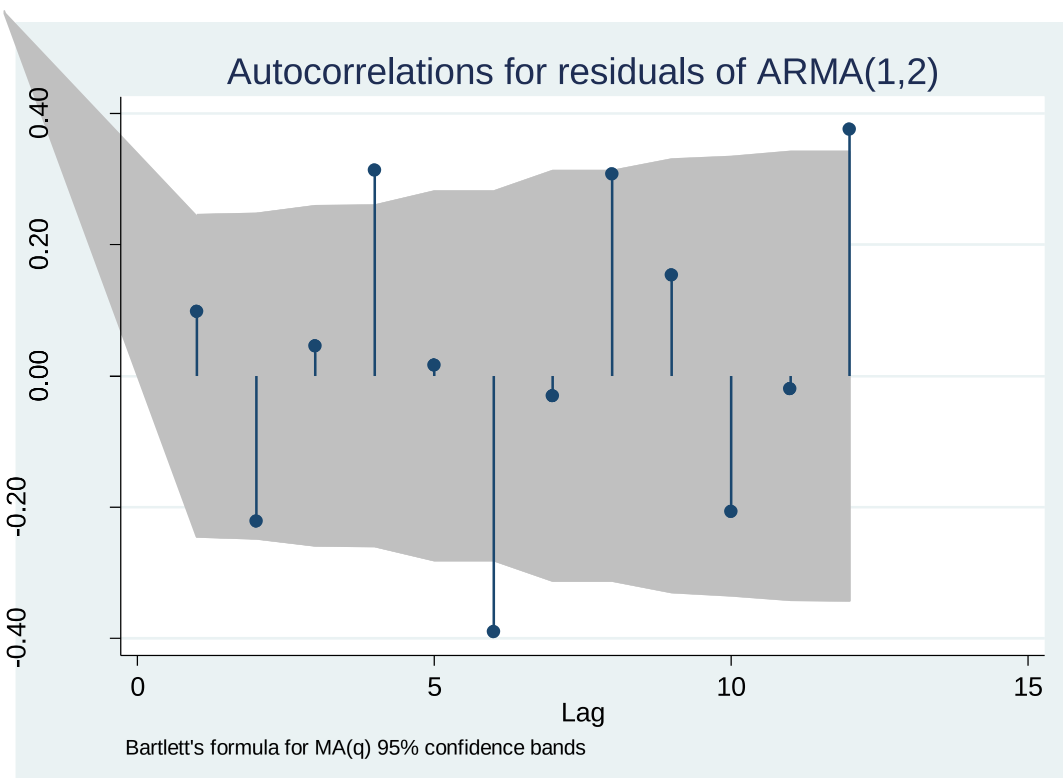

Figure 4.15: Autocorrelations for the Residuals of ARMA (1,2) of Semi-Detached

Although the parameter estimates in Table 4.5 indicate that the parameter estimates are significantly different from zero at the 1% level, the ACFs of the residuals indicate model inadequacy as there is a significant ACF at lags 4 and 6. A different model specification is required. An ARMA(5,0) is estimated and the results are in Table 4.7.

Table 4.7: ARMA (5,0) Model for Semi-Detached

Sample:1-64			Number of Observations=64 Wald*** chi2(5)=31.43	
Log likelihood = 97.74598			Prob > chi2 = 0.0000	
Parameters	**Co-efficient**	**Std.Err.**	**z**	**P>|z|**
μ	0.0180177*	0.0101906	1.77	0.077
$ø_1$	-0.4460045***	0.1150392	-3.88	0
$ø_2$	-0.2200073*	0.1185008	-1.86	0.063
$ø_3$	0.1097384	0.1583252	0.69	0.488
$ø_4$	0.470387***	0.1261588	3.73	0
$ø_5$	0.4328695***	0.1174463	3.69	0

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero.*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%.

The Wald Chi-Square is significant at the 1%. We next check for the adequacy of the residuals by examining their ACFs and PACFs in Figures 4.16 and 4.17.

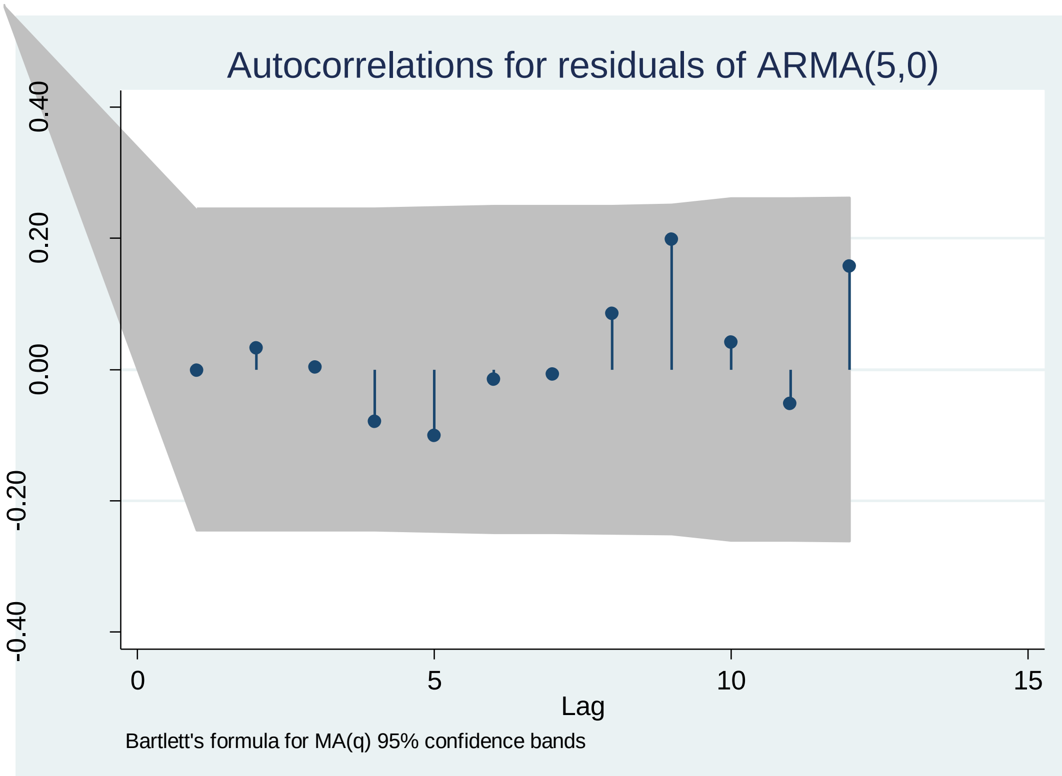

Figure 4.16: Autocorrelations for ARMA(5,0) for the Residuals of Semi-Detached

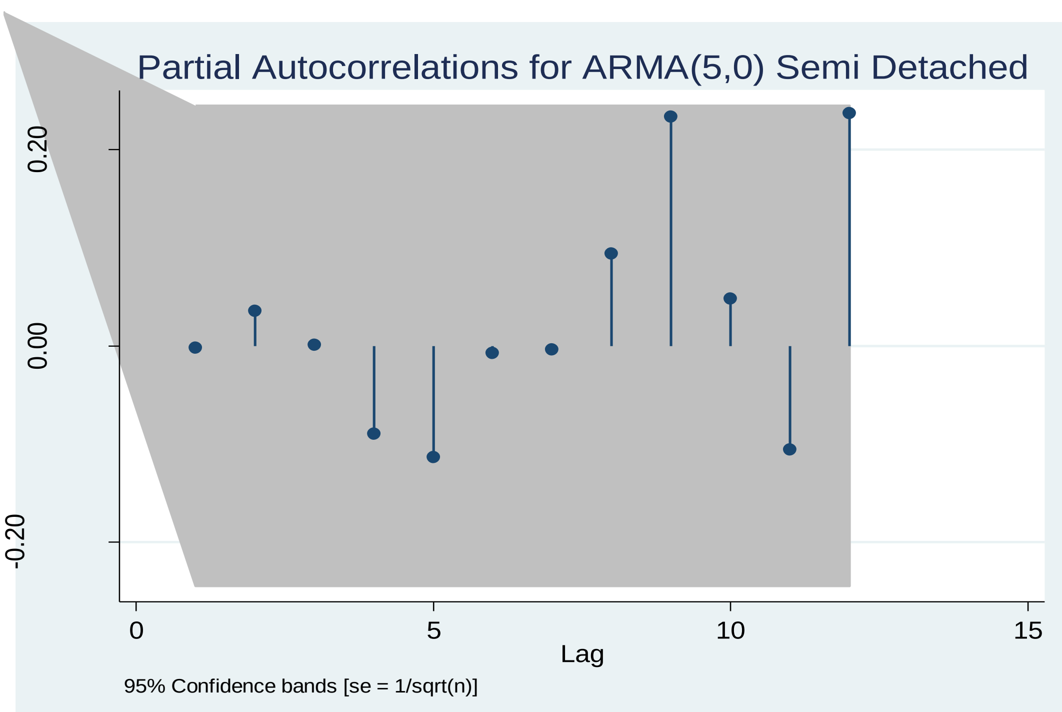

Figure: 4.17: Partial Autocorrelations for ARMA (5,0) Semi-Detached

None of the ACFs and PACFs are outside the two standard errors and therefore we ‘accept’ the hypothesis of no significant ACFs or PACFs at the 5% level. Accordingly, the Semi-Detached price changes follow the ARMA (5,0).

$$y_{t}=\mu + \emptyset_1 \, y_{t-1} + \emptyset_2 \, y_{t-2} + \emptyset_3 y_{t-3} + \emptyset_4 y_{t-4} + \emptyset_5 y_{t-5} + \epsilon_t \qquad (4.3)$$

Where,

y_{t}= price change in time t,

μ = constant term,

$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$= price change at lag one, two, three, four and five,

$\emptyset_1, \emptyset_2, \emptyset_3, \emptyset_4, \emptyset_5$ = coefficients estimates at lag one, two, three, four and five,

ϵ_t= residuals which are normally distributed with a mean of zero and variance of one.

We next examine whether or not the residuals follow the normal distribution. Tables 4.8 and 4.9 have the results of Shapiro-Wilk (1965) and Jarque-Bera (1980).

Table 4.8: Shapiro Wilk W Test For Normal Data

Variables	**Obs**	**W**	**V**	**Z**	**Prob>z**
Residuals	64	0.98198	1.032	0.067	0.47315

Table 4.9: Skewness Kurtosis Test For Normality

Variable	**Obs**	**Pr (Skewness)**	**Pr (Kurtosis)**	**adj chi2(2)**	**Prob>chi2**
Residuals	64	0.2204	0.7776	1.64	0.44

The Shapiro-Wilk (1965) and Jarque-Bera (1980) tests indicate adequacy of the model assumption of normally distributed residuals.

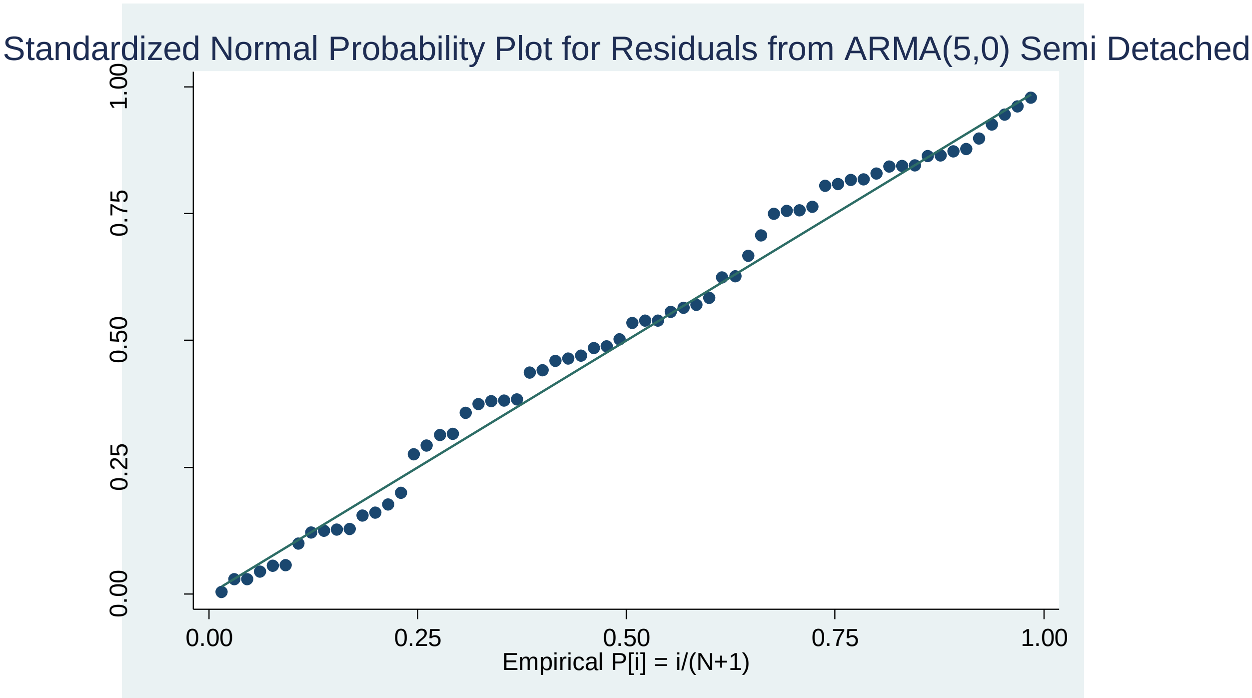

Figure: 4.18: Standardized Normal Probability Plot for Residuals from ARMA (5,0) Semi-Detachedrty

The results of Shapiro-Wilk (1965) and Jarque-Bera (1980) are confirmed by examining the normal QQ plot in Figure 4.18. The residuals approximately fall on the theoretical normal line. The ARMA(5,0) specification is 'accepted' for the Semi-Detached as none of the ACFs are significant at the 5% level. This indicates a relatively long memory as the price for the Semi-Detached for the current quarter is determined by its previous prices up to the five past quarters.

We also test for any subtle seasonality in the quarter data and whether the ARMA(5,0) has taken care of it. Four dummy variables of the type (0,1) have been created for quarters 1,2, 3, and 4 respectively. A regression is run where Semi-Detached quarterly price changes are the dependent variable and the quarter's one, two, three and four are the independent variables. There is no constant in the model. The residuals terms are assumed to be normally distributed with a mean of zero and a variance of one.

Table 4.10: Price Changes of Semi-Detached for ARMA (5,0) with Dummy Variables

Source	**SS**	**Df**	**MS**	**No.Of Obs.**	64
Model	0.118501471	4	0.029625368	**F(4,60)**	9.27
Residual	0.191785394	60	0.003196423	**Prob>F**	0
Total	0.310286865	64	0.004848232	**R-Squarred**	0.3819
				Adj R-Squarred	0.3407
				Root MSE	0.05654
Semi-Detached	**Coef.**	**Std.Err.**	**t**	**P> I t I**	
Quarter 1	-0.0028379	0.0141342	-0.2	0.842	
Quarter 2	0.0515962	0.0141342	3.65	0.001	
Quarter 3	0.0622961	0.0141342	4.41	0	
Quarter 4	-0.0292457	0.0141342	-2.07	0.043	

Quarters two and three are significantly different from zero at the 1% level. The model is highly significant at the 1% level as indicated by the F-statistic in the above Table. The hypothesis of no seasonality in the Semi-Detached market is rejected. The parameter estimates are positive indicating positive changes from the equivalent quarters in the previous year. The parameter estimate for quarter 4 is significantly different from zero at the 5% level. It is negative indicating a negative decline in the fourth quarter compared with the previous year. Seasonality explains 34% of the variation in price changes for the Semi-Detached as indicated by the significant adjusted R-squared.

Table 4.11: Results of ARMA (5,0) Residuals

Source	**SS**	**df**	**MS**	**No.Of Obs.**	64
Model	0.018135993	4	0.004533998	**F(4,60)**	1.73
Residual	0.157509992	60	0.002625167	**Prob>F**	0.1558
Total	0.175645984	64	0.002744469	**R-Squarred**	0.1033
				Adj R-Squarred	0.0435
				Root MSE	0.05124
Semi-Detached	**Coef.**	**Std.Err.**	**t**	**P> I t I**	
Quarter 1	-0.0102968	0.0128091	-0.8	0.425	
Quarter 2	0.0018513	0.0128091	0.14	0.886	
Quarter 3	0.0275237	0.0128091	2.15	0.036	
Quarter 4	-0.0163246	0.0128091	-1.27	0.207	

The residuals from the identified ARMA(5,0) were run as dependent variables on the four dummies corresponding to the four quarters. The F-statistic is not significant at the 1%, 5% or even 10% indicating that the suggested model ARMA(5,0) has successfully handled the observed seasonality in the Semi-Detached market.

4.4 Model Building For Detached

Figure 4.19 plots the quarterly price change for the Detached during the period from first quarter 1995 to first quarter 2011.

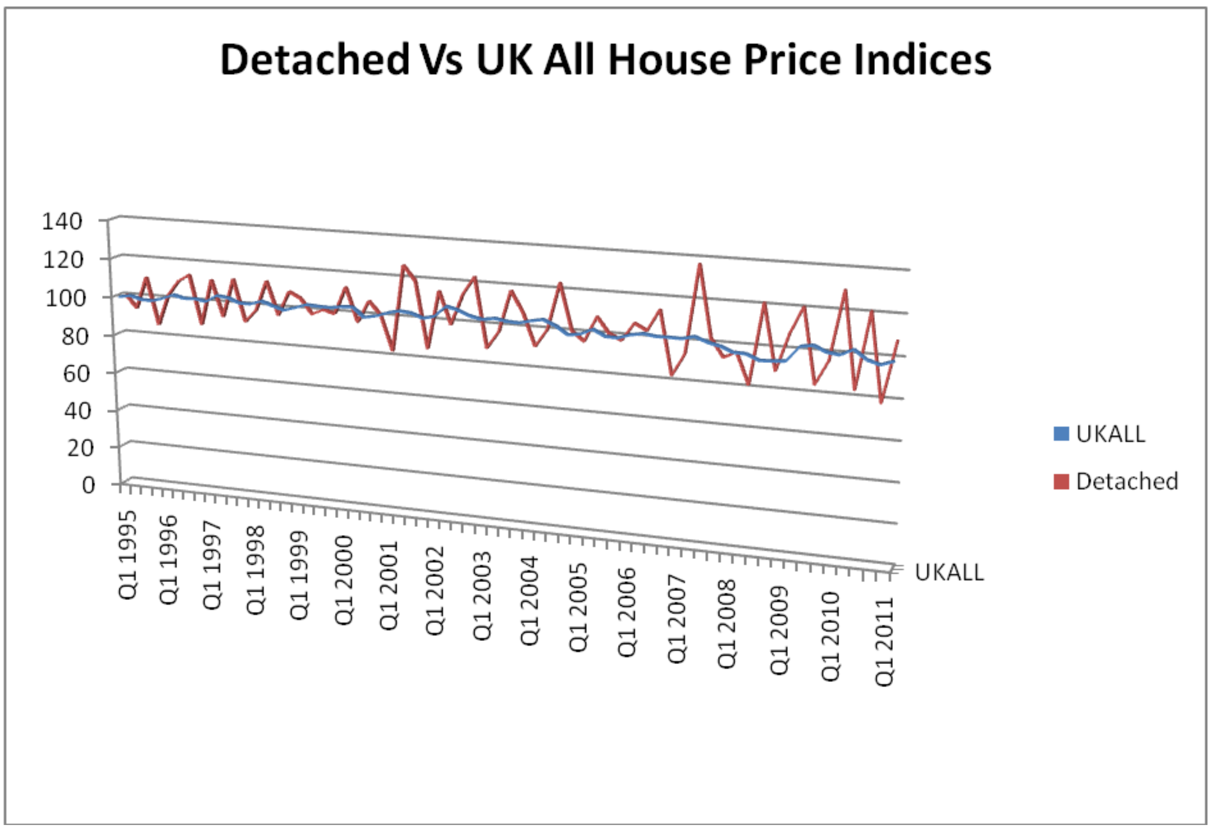

Figure 4.19: Detached Vs UK All House Price Indices

The Figure shows that the Detached price changes are far more volatile than price changes for the UK all house price index. There has been a sharp fall in the Detached price in September 2001 and in the credit crunch of 2008. Figure 4.20 plots the price changes for the Detached for the period form quarter two 1995 to quarter one 2011.

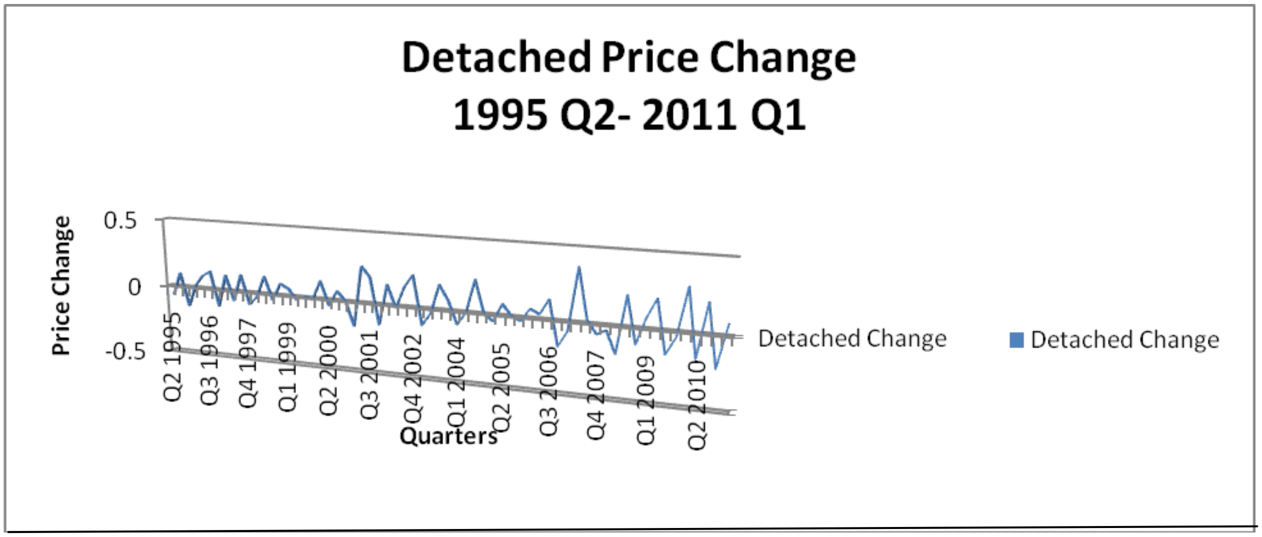

Figure 4.20: Detached Price Change

The above Figure represents the price change for Detached from the period 1995 Q2 to 2011 Q1. From the figure, it is clearly seen that the price change is volatile. Detached price change can be categorized into three periods: low volatility, medium volatility and high volatility. In the above figure, low volatility is marked between the period 1995 Q2 to 2001 Q2; the medium volatility is marked between 2002 Q2 to 2007 Q2 and the high volatility is marked between 2008 Q1 to 2011 Q1. Figure 4.21 has the histogram for the Detached price change along with the theoretical normal distribution imposed.

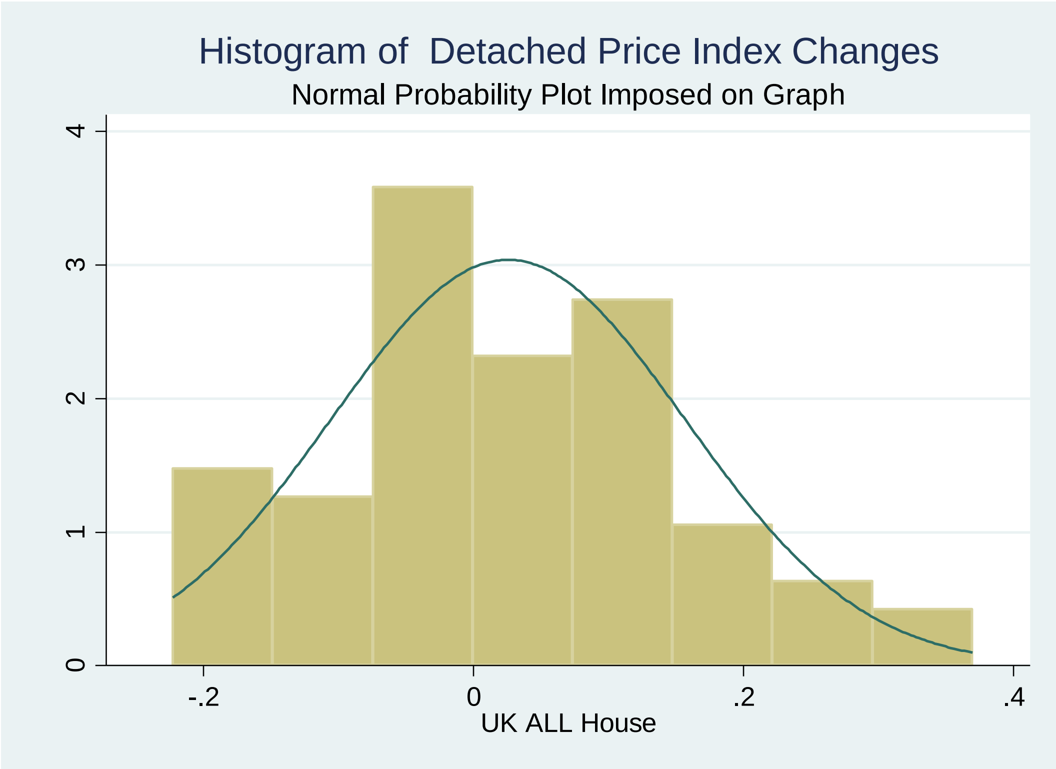

Figure 4.21: Histogram of Detached Price Index Changes

The above Figure shows the distribution to be skewed to the right. The skewness coefficient is 0.289 and has excess kurtosis compared with the normal distribution as evidenced by a kurtosis coefficient of 2.58. The mean is 0.03 which is greater than the median of 0.003. The standard deviation is 0.131. The distribution is not normally distributed. We next examine the Autocorrelations (ACF) and Partial Autocorrelations (PACF) Functions for Semi-Detached price changes. The ACFs and PACFs are shown in Figures 4.22 and 4.23**.**

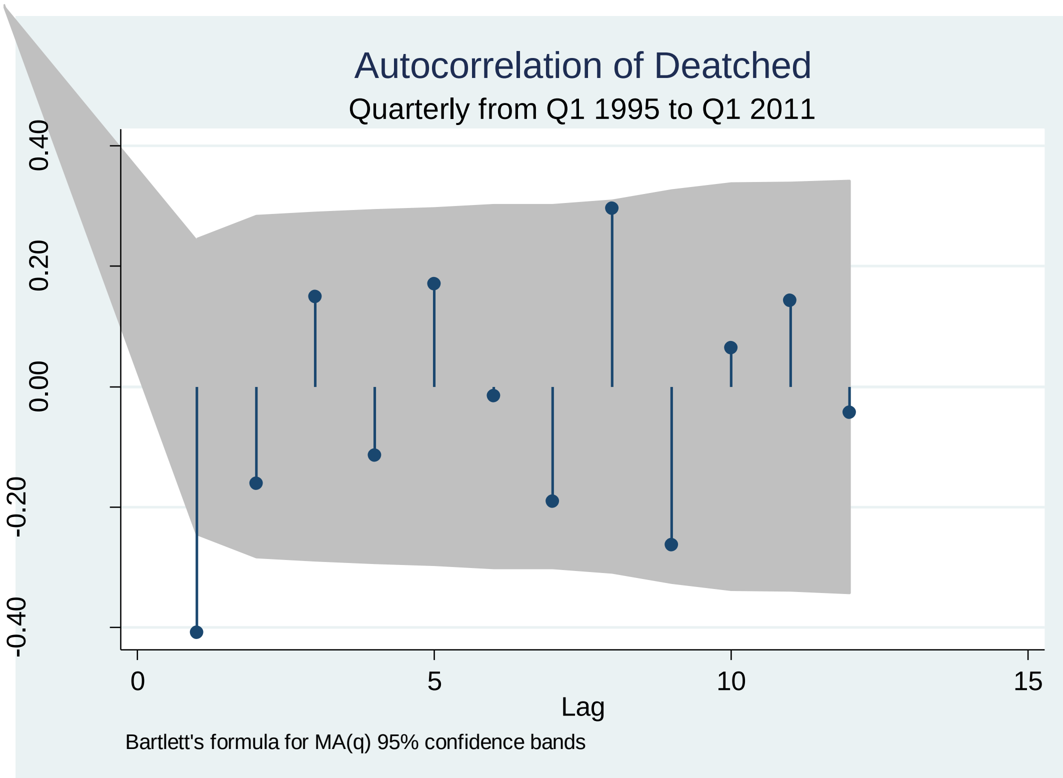

Figure 4.22: Autocorrelation of Detached

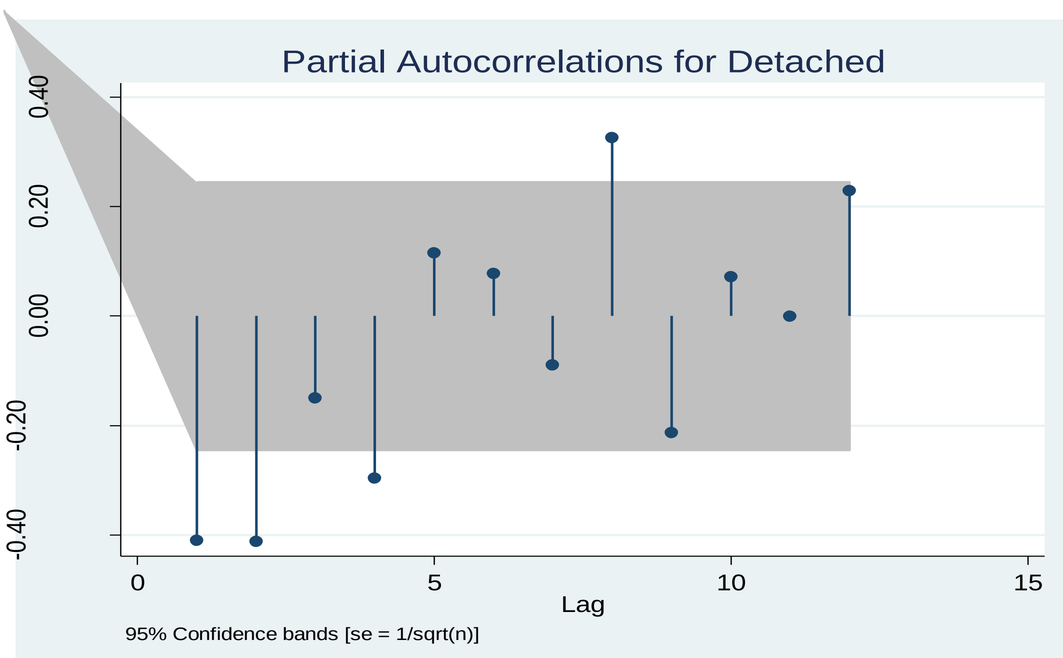

Figure 4.23: Partial Autocorrelations for Detached

The ACFs show a lag at 1 in Figure 4.22. The PACFs have significant lags at 1, 2, 4 and 8 in Figure 4.23. An ARMA(4,0) is estimated. The results are in Table 4.12.

Table 4.12: ARMA (4,0)Model for Detached

Sample:1-64			Number of Observations=64 Wald chi2(4)=27.24***	
Log likelihood = 54.14194			Prob > chi2 = 0.0000	
Parameters	**Co-efficient**	**Std.Err.**	**z**	**P>|z|**
μ	0.274857***	0.0048981	5.61	0
ϕ_1	-0.679***	0.1343554	-5.05	0
ϕ_2	-0.645***	0.1584498	-4.07	0
ϕ_3	-0.329**	0.1514085	-2.17	0.03
ϕ_4	-0.295**	0.1176434	-2.51	0.012

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero.*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%.

Table 4.12 shows that the Wald Chi-Square is significant at the 1%. All parameter estimates of the ARMA(4,0) are statistically significant at the 5%. We next check for the adequacy of the residuals by examining their ACFs and PACFs in Figures 4.24 and 4.25.

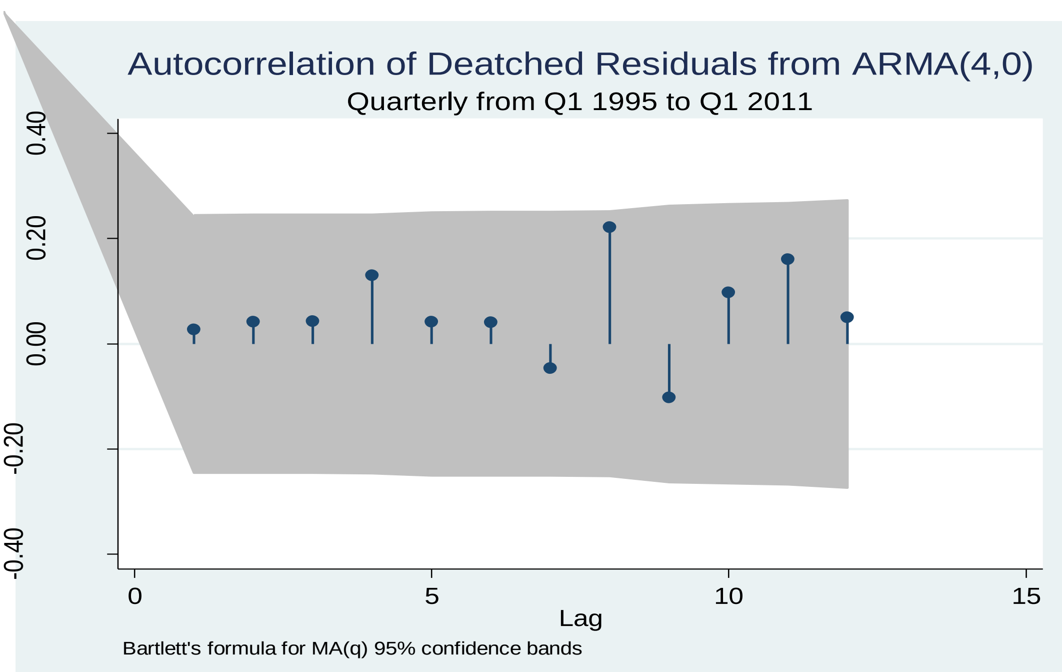

Figure 4.24: Autocorrelations of the Residuals of ARMA (4,0) of Detached

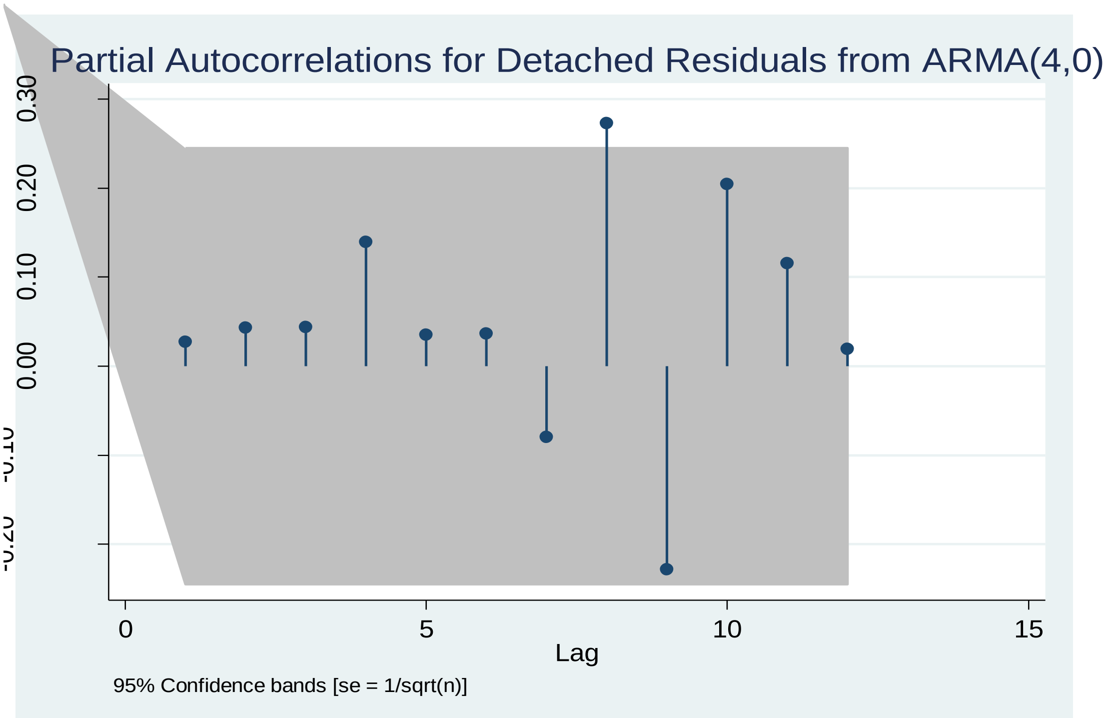

Figure 4.25:Partial Autocorrelations of the Residuals of ARMA (4,0) of Detached

None of the ACFs is significant at the 5%. However, there is a significant PACF at lag 8 as shown in Figure 4.25. Although it is significant, it is not far away from the 95% confidence intervals and not significant at the 1% level.

We next examine for whether or not the residuals follow the assumption of the normal distribution. Tables 4.13 and 4.14 show the Shapiro-Wilk (1965) and Jarque-Bera (1980) calculations.

Table 4.13: Shapiro Wilk W Test for Normal Data

Variable	**Obs**	**W**	**V**	**Z**	**Prob>z**
Residuals	64	0.99059	0.538	-1.339	0.90975

Table 4.14: Skewness Kurtosis Test for Normality

Variable	**Obs**	**Pr (Skewness)**	**Pr (Kurtosis)**	**adj chi2(2)**	**Prob>chi2**
Residuals	64	0.9891	0.4588	0.56	0.7555

The Shapiro-Wilk (1965) and Jarque-Bera (1980) tests indicate that the assumption of normally distributed residuals is not rejected. The results of the Shapiro-Wilk (1965) and Jarque-Bera (1980) tests are confirmed by examining the normal QQ plot in Figure 4.26. The residuals approximately fall on the theoretical normal line.

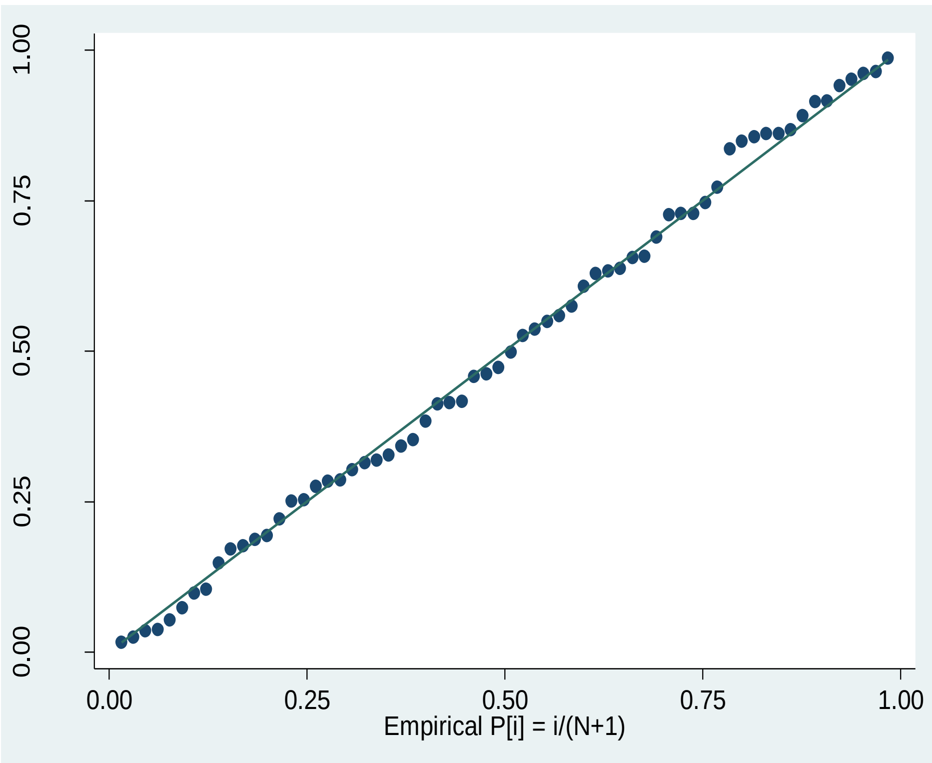

Figure 4.26: Standardized Normal QQ Test for Detached Residuals from ARMA (4,0)

The ARMA(4,0) specification is 'accepted' for the Detached. This indicates a relatively long memory as the price for the Detached for the current quarter is determined by its previous prices up to four past quarters.

The following equation is for ARMA (4,0)

$$y_{t=}\mu + \emptyset_1\, y_{t-1} + \emptyset_2\, y_{t-2} + \emptyset_3 y_{t-3} + \emptyset_4 y_{t-4} + \epsilon_t \qquad (4.4)$$

Where,

$y_{t=}$= price change in time t,

μ = constant,

$y_{t-1,} y_{t-2}, y_{t-3}, y_{t-4}$ = price changes at lag one, two, three and four quarter,

$\emptyset_1, \emptyset_2, \emptyset_{3,}, \emptyset_4$ = coefficient estimates at lag one, two, three and four.

ϵ_t = residual term.

The residuals from the above model were used as dependent variables in the following model to test for seasonality.

$$y_t = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + u_t \qquad (4.5)$$

where,

y_t= residuals from the fitted model;

β= parameter estimates;

D_1=dummy variable 1 in quarter 1, and 0 in the other three quarters;

D_2=dummy variable 1 in quarter 2, and 0 in the other three quarters;

D_3=dummy variable 1 in quarter 3, and 0 in the other three quarters;

D_4=dummy variable 1 in quarter 4, and 0 in the other three quarters;

u_t= residual term with normal distribution

Table 4.15: Results of ARMA (4,0) Residuals

Source	**SS**	**df**	**MS**	**No.Of Obs.**	64
Model	0.109177933	4	0.027294483	**F(4,60)**	2.83
Residual	0.578056029	60	0.009634267	**Prob>F**	0.0321
Total	0.687233962	64	0.010738031	**R-Squarred**	0.1589
				Adj R-Squarred	0.1028
				Root MSE	0.09815
Detached	**Coef.**	**Std.Err.**	**t**	**P> I t I**	
Quarter 1	-0.0116541	0.0245386	-0.47	0.637	
Quarter 2	-0.0149722	0.0245386	-0.61	0.544	
Quarter 3	0.0647492	0.0245386	2.64	0.011	
Quarter 4	-0.0476569	0.0245386	-1.94	0.057	

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero.
*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%.

The above table shows the results of running the residuals from the ARMA(4,0) as dependent variables on the four quarters. The model fails to take care of the seasonality observed at quarter three.

An ARMA(4,0) regression with the dummy for quarter three is estimated.

Table 4.16: ARMA(4,0)Regression With the Quarter Three Dummy

Source	**SS**	**df**	**MS**	**No.Of Obs.**	60
Model	0.496992723	5	0.099398545	**F(5,54)**	9.86
Residual	0.544512577	54	0.010083566	**Prob>F**	0
Total	1.0415053	59	0.017652632	**R-Squarred**	0.4772
				Adj R-Squarred	0.4288
				Root MSE	0.10042
Detached	**Coef.**	**Std.Err.**	**T**	**P> I t I**	
AR (1)	-0.6563309	0.1204346	-5.45		0
AR (2)	-0.6075435	0.1442362	-4.21		0
AR (3)	-0.2258545	0.1463273	-1.54		0.129
AR (4)	-0.3347502	0.1253411	-2.67		0.01
Quarter 3	0.0973883	0.0322194	3.02		0.004
Constant	0.0563749	0.0198881	2.83		0.006

The ARMA(4,0) for the Detached is highly significant at the 1% level as indicated by the F statistic. The adjusted R-squared is 43% indicating that 43% of the variation in percentage price changes for the Detached can be explained by the previous four lags in addition to the seasonality observed in the third quarter. The AR(1), AR(2), AR(4) parameter estimates are negative and each is significantly different from zero at the 1% level. The negative parameter estimates indicate that there is a reversal pattern with regard to percentage price changes, i.e. positive percentage price change will have a higher chance of being followed by negative percentage price change and vice versa. Table 4.16 check for the adequacy of the residuals from the ARMA(4,0) with a seasonal third quarter.

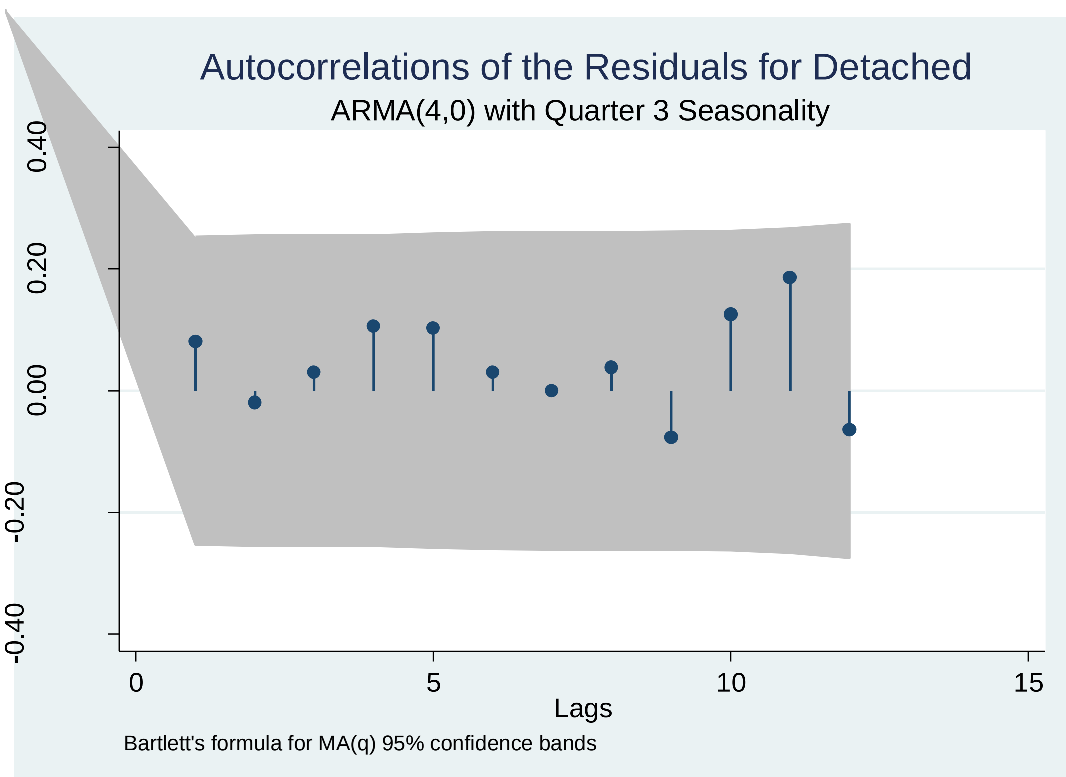

Figure 4.27: Autocorrelations of the Residuals for Detached ARMA (4,0) with Quarter 3 Seasonality

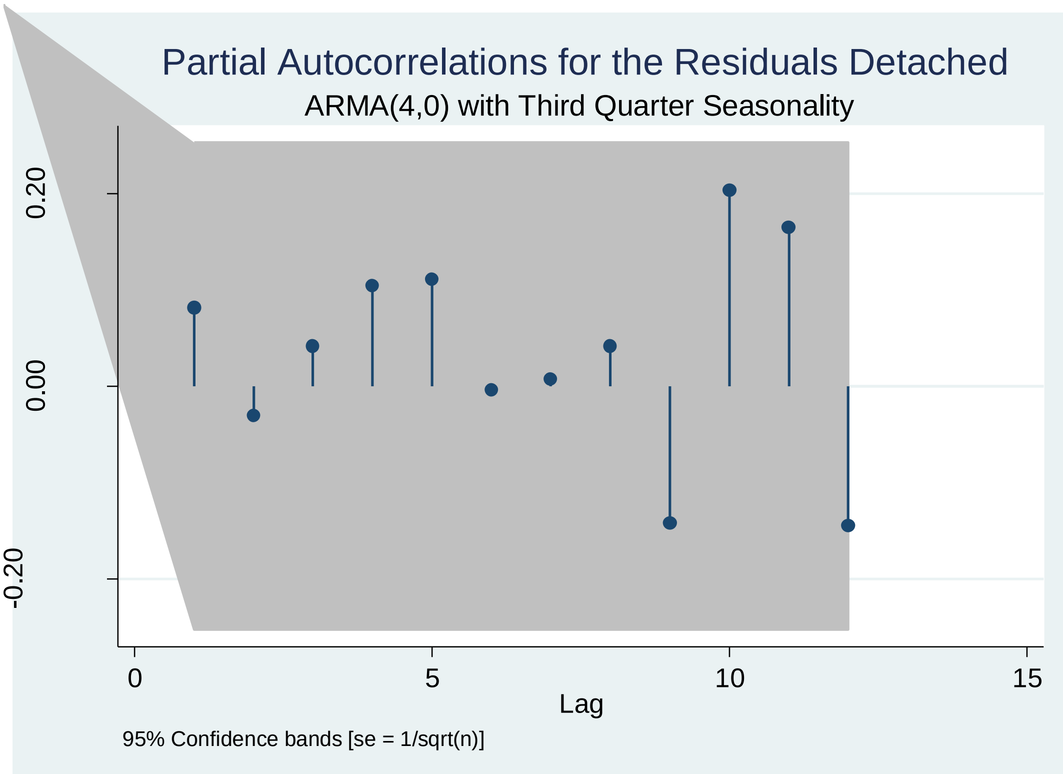

Figure 4.28:Partial autocorrelations of the Residuals for Detached ARMA (4,0) with Quarter 3 Seasonality

The ACFs and PACFs are within two standard errors from zero indicating no significance at the 5% level. The ARMA(4,0) with the third quarter seasonality seems to fit the data well for the Detached.

4.5 Model Building for Flats

Figure 4.29 plots the prices for Flats against the UK All price index during the period form quarter one 1995 to quarter one 2011.

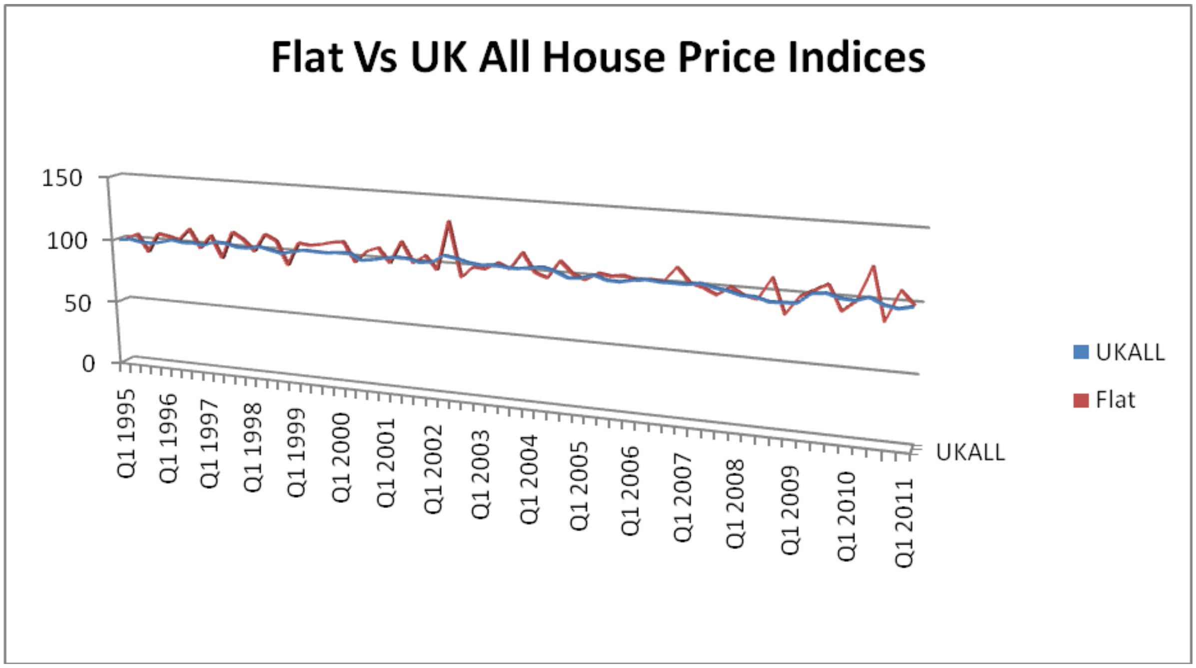

Figure 4.29: Flat Vs UK All House Price Indices

Flat prices in Manchester are more volatile than the UK All index. There has been a sharp rise in Flats reaching an index of 130 (a percentage price increase of 30) during 2002 Q1. The Flats price fell during the credit crunch of 2008 and was followed by a period of higher volatility. Figure 4.30 has the price changes for Flats in Manchester from quarter two 1995 to quarter one 2011.

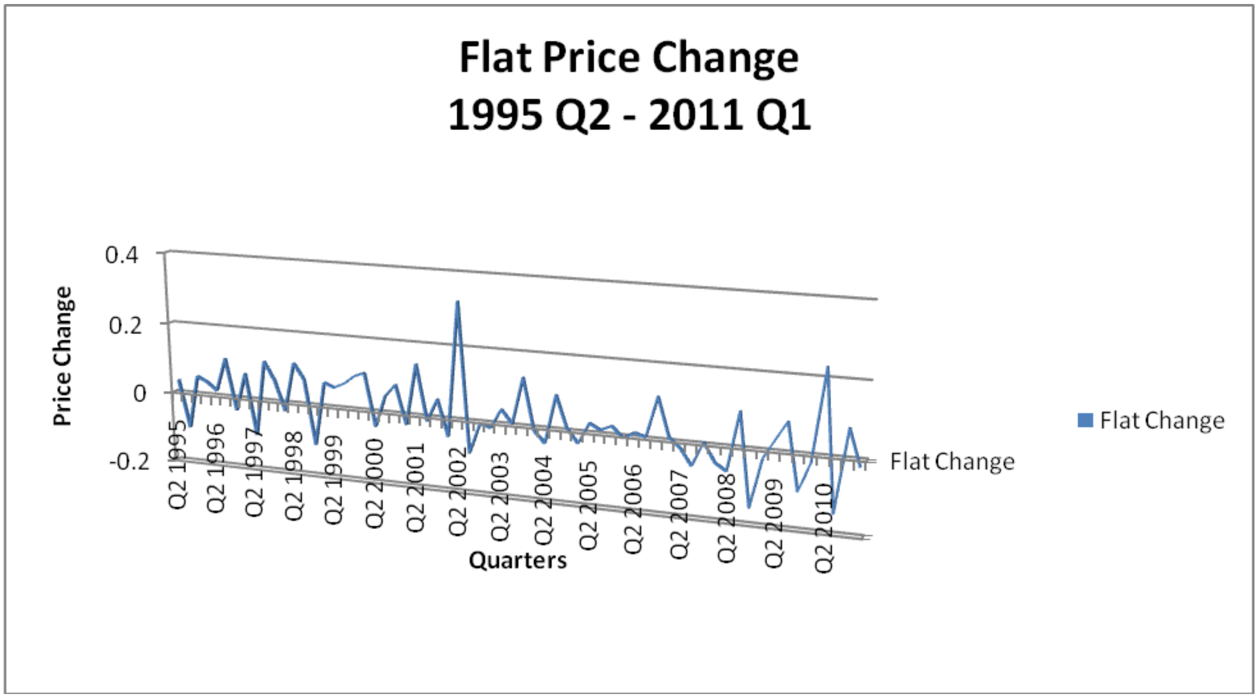

Figure 4.30: Price Change for Flats in Manchester

From this figure, it can be clearly seen that price changes for Flats were volatile in the above mentioned time period. This price change can be categorized into three scales: low volatility, medium volatility and high volatility. In the above figure, low volatility is marked between the period 2002 Q2 to 2007 Q2; medium volatility is marked between 1995 Q2 to 2001 Q2; and high volatility is marked between 2008 Q1 to 2011 Q1. Figure 4.31 shows the histogram of the price changes for Flats along with the normal distribution imposed.

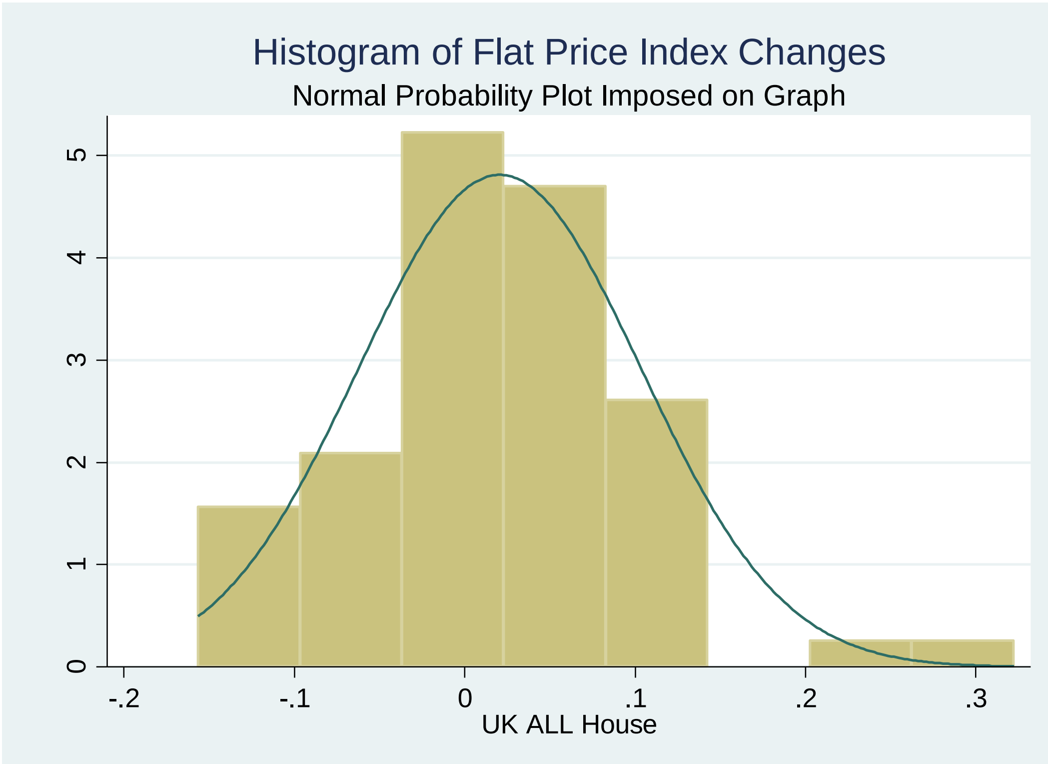

Figure 4.31: Histogram of Flat Price Index Changes

The above Figure shows the distribution to be skewed to the right. The skewness coefficient is 0.67 and has excess kurtosis compared with the normal distribution as evidenced by a kurtosis coefficient of 4.82. It is important to note that the excess kurtosis of the Flats is far more than that for Semi-DetachedSemi-Detached (2.34) and Detached (2.58). The right tails reach the 0.30 which gives a chance for 30% price appreciation. The mean is 0.021 which is greater than the median of 0.011. The standard deviation is 0.083. The distribution is not normally distributed. We next examine the Autocorrelations (ACF) and Partial Autocorrelations (PACF) Functions for Flat price changes. The ACFs and PACFs are shown in Figures 4.32 and 4.33.

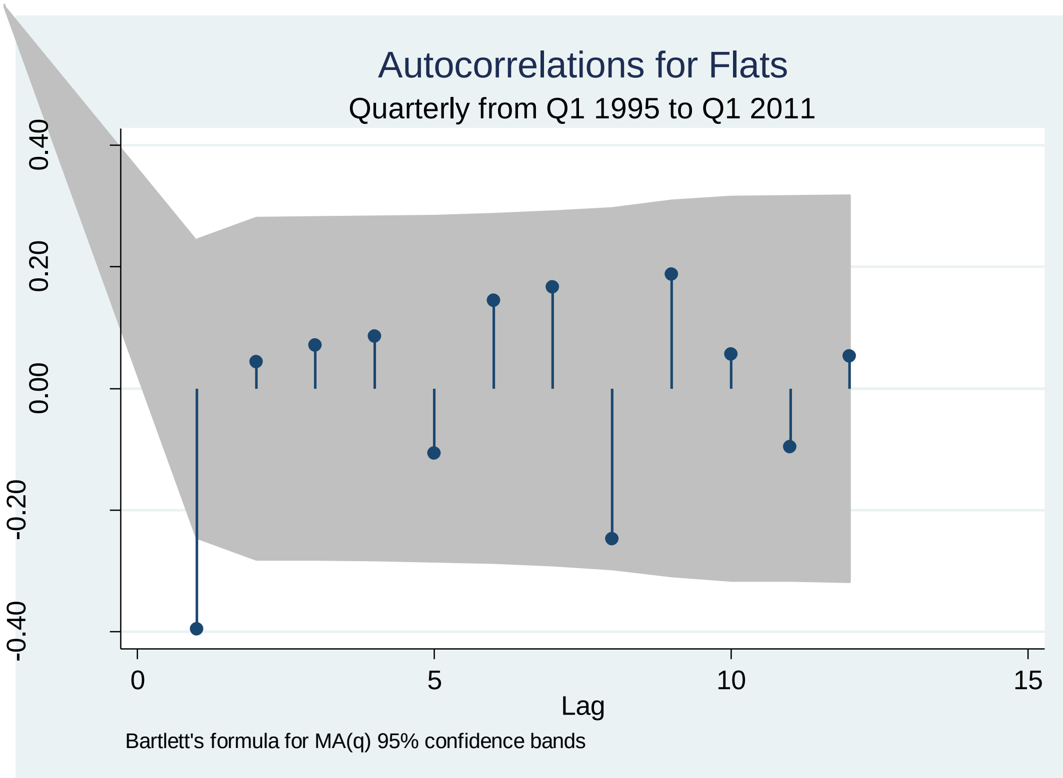

Figure: 4.32: Autocorrelation for Flats

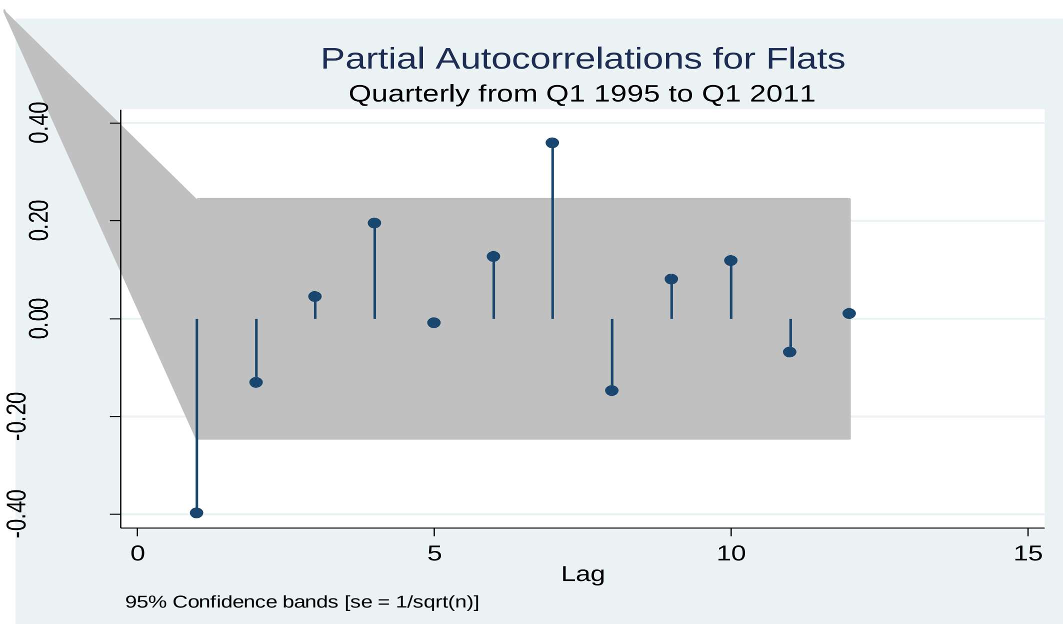

Figure 4.33: Partial Autocorrelations for Flats

The ACFs show a significant lag at 1 in Figure 4.32. The PACFs show significant lags at 1 and 7 in Figure 4.33. An ARMA(1,0) is estimated. The results are in Table 4.17.

Table 4.17: ARMA (1,0) Model for Flats

Sample:1-64 Log likelihood = 74.41271			Number of Observations=64 Wald chi2(1)=7.67*** Prob > chi2 = .0056	
Parameters	**Co-efficient**	**Std.Err.**	**z**	**P>|z|**
µ	0.0207035***	0.0073961	2.8	0.005
$ø_1$	-0.391424***	0.1413469	-2.77	0.006

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero.*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%

The Wald Chi-Square is significant at the 1%. We next check for the adequacy of the residuals by examining their ACFs and PACFs in Figures 4.34 and 4.35.

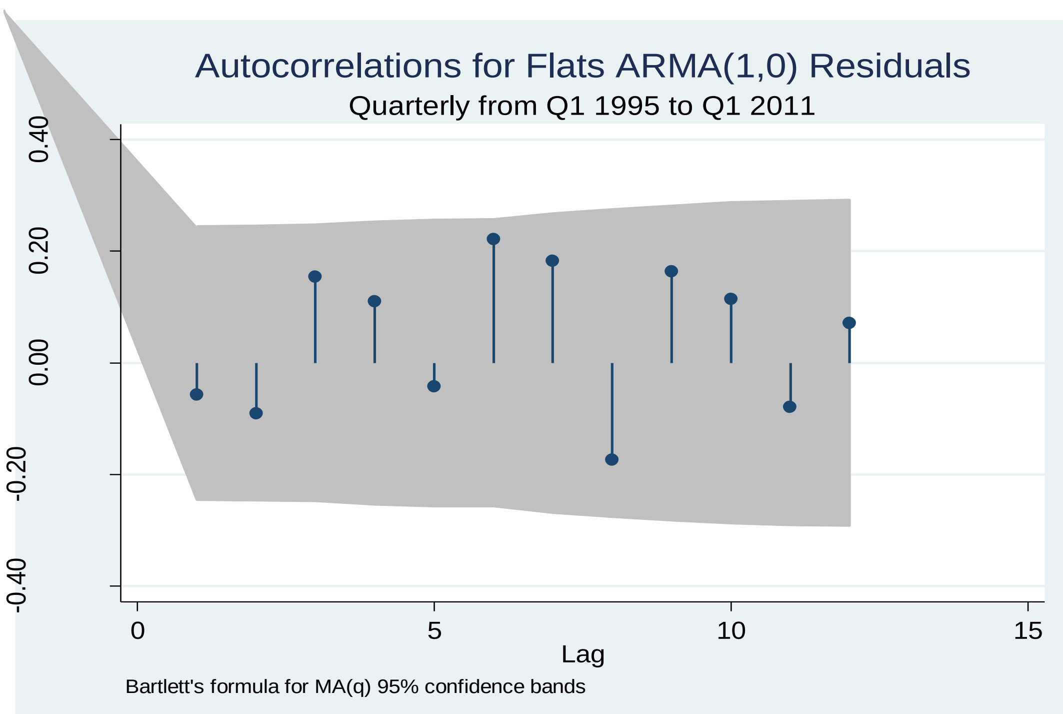

Figure: 4.34: Autocorrelations for Flats ARMA(1,0) Residuals

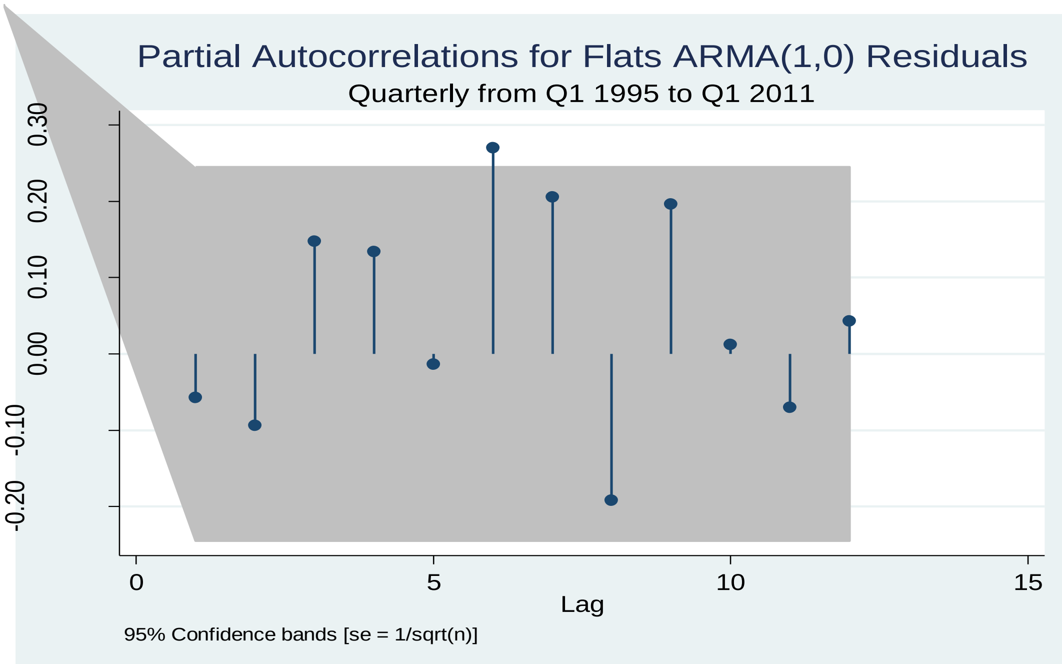

Figure: 4.35: Partial Autocorrelations for Flats ARMA(1,0) Residuals

None of the ACFs is significant at the 5%. However, there is a PACF at lag 6 as shown in Figure 4.35. Since it is not far away from the 95% confidence intervals and not significant at the 1% level, we applied the Q-test of Ljung and Box (1978) for up to 12 lags. The results are in Table 4.18

Table 4.18: Q-tests of Ljung and Box (1978) at lags up to twelve.

LAC	AC	PAC	Q	Prob>Q
1	-0.0572	-0.0572	0.21928	0.6396
2	-0.0904	-0.0904	0.77592	0.6784
3	0.1544	0.1473	2.4278	0.4885
4	0.1096	0.1341	3.2735	0.5131
5	-0.042	-0.0135	3.3999	0.6386
6	0.2218	0.2703	6.9838	0.3223
7	0.1829	0.2056	9.4617	0.2212
8	-0.1739	-0.1918	11.743	0.163
9	0.1642	0.1966	13.813	0.1291
10	0.1142	0.0122	14.833	0.1383
11	-0.0787	-0.0702	15.326	0.168
12	0.0711	0.0434	15.737	0.2036

Table 4.18 reports the ACFs, PACFs, Q-statistic and the probability of the significance of the Q-Statistic. As indicated in the Table none of the Qs are significant at 1, 5, or 10% level and therefore the hypothesis of serially uncorrelated residuals is not rejected.

We next examine for whether or not the residuals follow the assumption of the normal distribution. Tables 4.19 and 4.20 have the Shapiro-Wilk (1965) and Jarque-Bera (1980) test data.

Table 4.19: Shapiro Wilk W Test for Normal Data

Variables	**Obs**	**W**	**V**	**Z**	**Prob>z**
Residuals	64	0.95943	2.323	1.823	0.03412

Table 4.20: Jarque-Bera (1980) *Skewness Kurtosis Test for Normality*

Variables	**Obs**	**Pr (Skewness)**	**Pr (Kurtosis)**	**adj chi2(2)**	**Prob>chi2**
Residuals	64	0.0155	0.0305	9.05	0.0108

The results of Shapiro-Wilk (1965) and Jarque-Bera (1980) tests for normality indicate the adequacy of the hypothesis of normally distributed residuals. This is also confirmed by examining the normal QQ plot in Figure 4.36. The residuals approximately fall on the theoretical normal line.

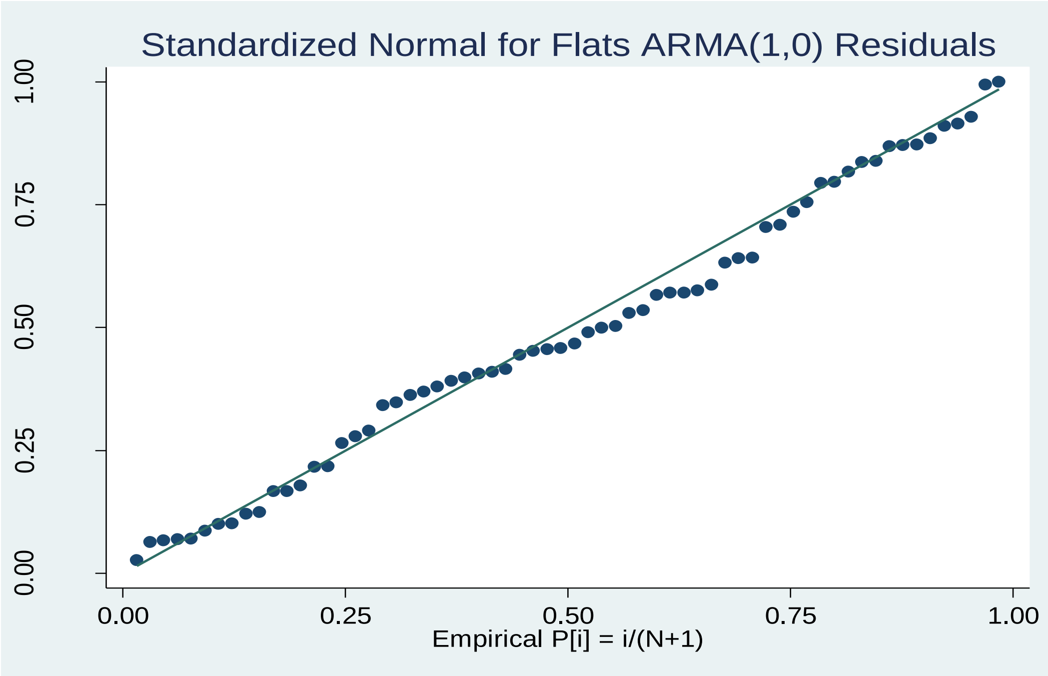

Figure 4.36: Standardized Normal for Flats from ARMA (1,0) Residuals

The ARMA(1,0) specification is 'accepted'. This indicates a relatively short memory as the price for the Flats for the current quarter is determined by its previous quarter only. This is in sharp contrast to the Detached when four lags were required.

Following equation is for ARMA (1,0)

$$y_{t=}\mu + \emptyset_1 \, y_{t-1} + \epsilon_t \qquad (4.6)$$

Where,

y_t= price change in time t

μ = constant

y_{t-1}= price change at lag one

$\emptyset_1$ = auto regression coefficient estimate at lag one

ϵ_t = residual term

The residuals from the ARMA(1,0) is used as dependent variable in the following seasonal dummy model.

$$y_t = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + u_t \qquad (4.7)$$

where,

y_t= the residuals from the fitted model;

β= parameter estimates;

D_1=dummy variable 1 in quarter 1, and 0 in the other three quarters;

D_2=dummy variable 1 in quarter 2, and 0 in the other three quarters;

D_3=dummy variable 1 in quarter 3, and 0 in the other three quarters;

D_4=dummy variable 1 in quarter 4, and 0 in the other three quarters;

u_t= residual term with normal distribution

Table 4.21: Results of ARMA (1,0) Seasonality

Source	**SS**	**df**	**MS**	**No.Of Obs.**	63
Model	0.013878865	4	0.003469716	**F(4,59)**	0.58
Residual	0.35263732	59	0.005976904	**Prob>F**	0.6779
Total	0.366516185	63	0.005817717	**R-Squarred**	0.0379
				Adj R-Squarred	-0.0274
				Root MSE	0.7731
ARMA 1:Flat Seasonality	**Coef.**	**Std.Err.**	**t**	**P> I t I**	
Quarter 1	-0.0216932	0.0193276	-1.12	0.266	
Quarter 2	0.0188439	0.0199615	0.94	0.349	
Quarter 3	0.0061162	0.0193276	0.32	0.753	
Quarter 4	-0.0051504	0.0193276	-0.27	0.791	

The results in the Table indicate no significant seasonality in the residuals as indicated by insignificant F-test and insignificant t-tests. The ARMA(1,0) for Flats passes all diagnostic tests including that of seasonality.

4.6 Model Building For Terraced

Figure 4.37 plots the Terraced price index for Manchester against the UK All price index from quarter one 1995 to quarter one 2011.

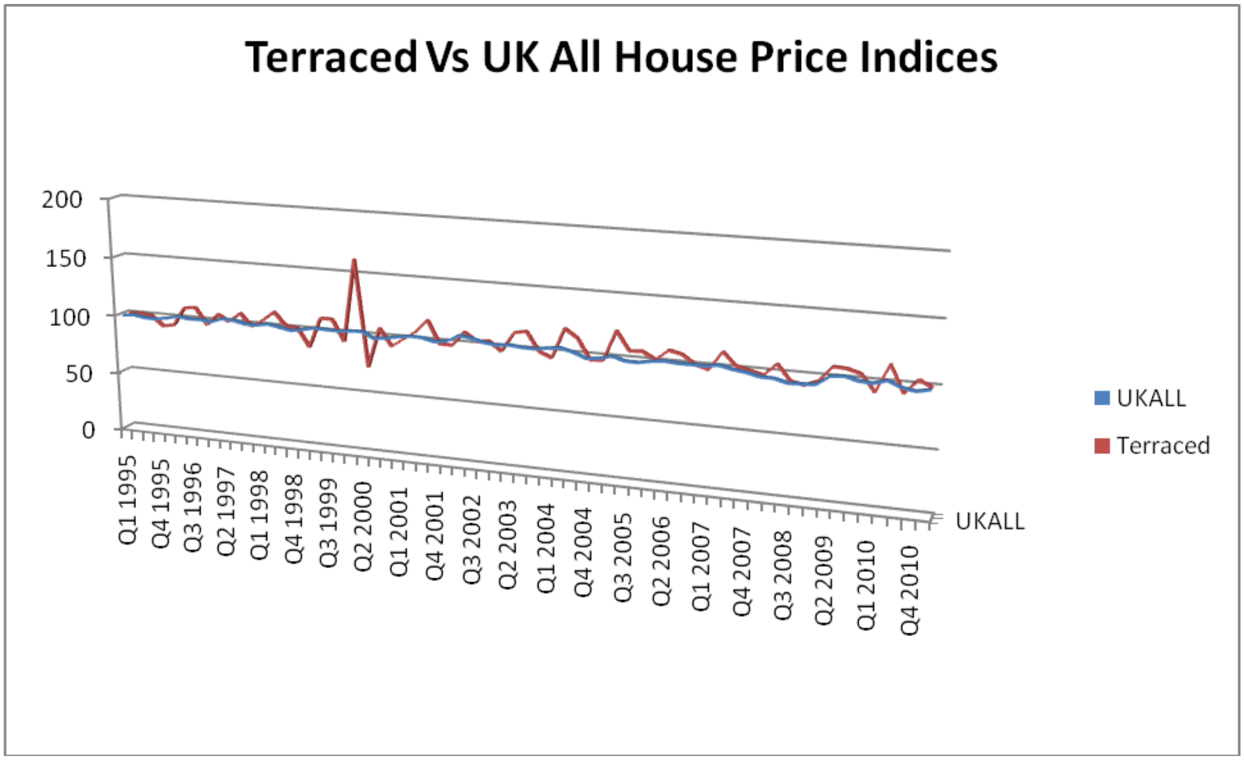

Figure 4.37: Terraced Vs UK All House Price Indices

As was the case with Semi-Detached properties, Detached properties and Flats indices for Manchester, price changes for Terraced homes seem to be more volatile than the UK All index. The main reason for this may be due to the fact that the UK all price index is smoother since it has far more coverage than Manchester. The behavior of the process of price change seems to be constant over time with the exception of the sharp rise in prices in 2000 quarter four when prices went up by more than fifty percent. Figure 4.38 indicates price changes in Terraced houses for the period 1995 quarter 2 to 2011 quarter one.

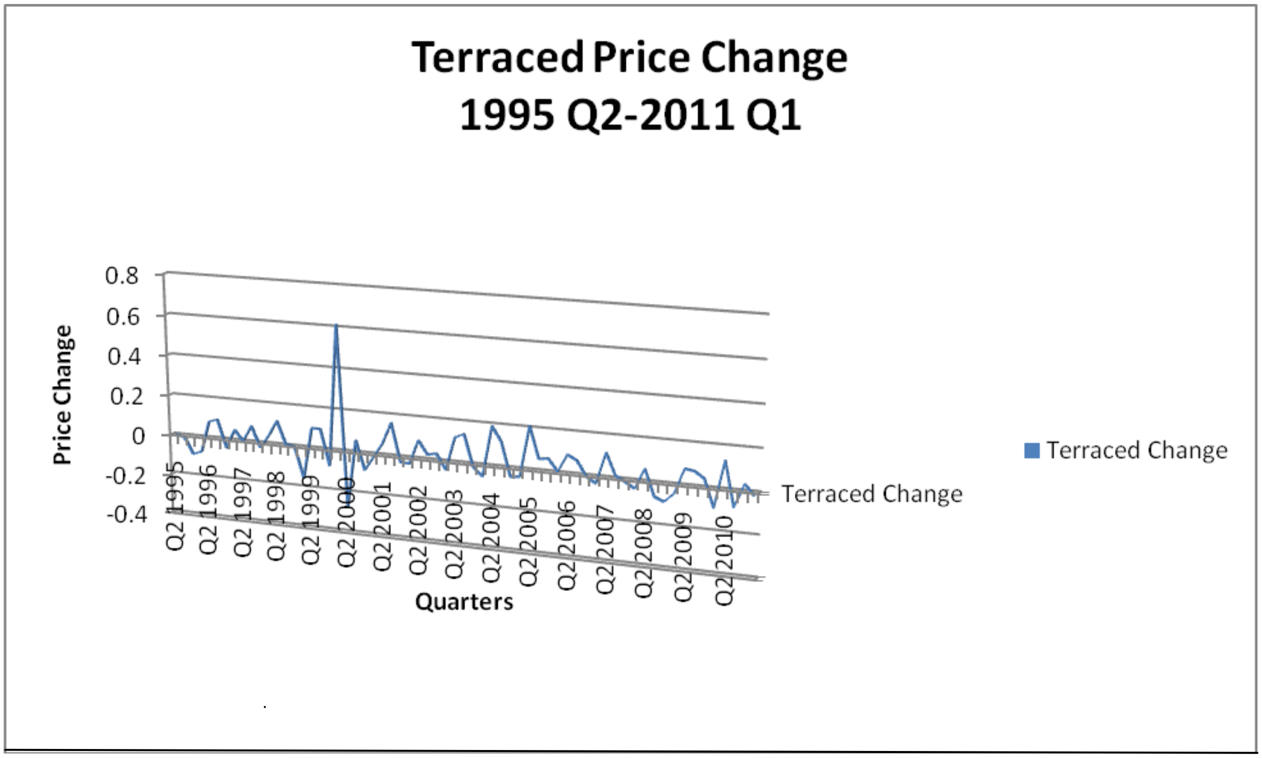

Figure 4.38: Terraced Price Change

This figure indicates that there was a large jump of 60% in price in the first quarter of 2000. The volatility of the series seems to be similar apart from the first quarter of 2000. Figure 4.39 has the histogram of the price changes for Terraced houses along with the normal probability imposed.

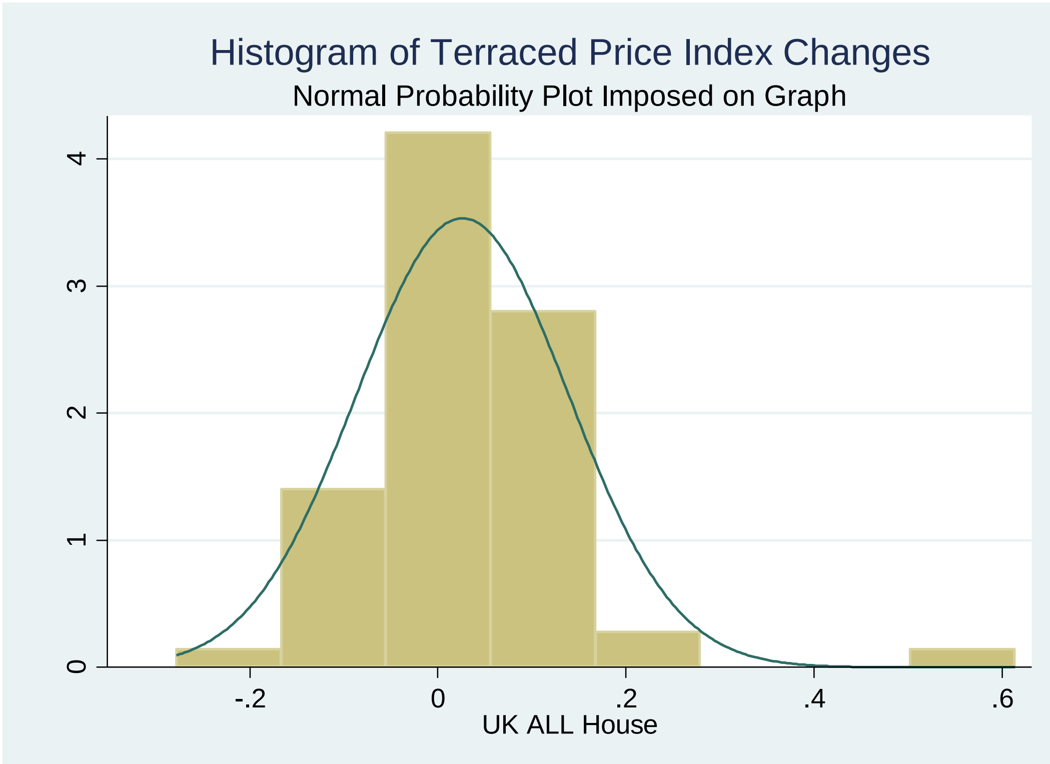

Figure 4.39: Histogram of Terraced Price Index Changes

The above figure shows the distribution to be skewed to the right as a result of one outlier corresponding to the first quarter 2000 that reached 0.60 (60% change in price). The mean is 0.026 which is greater than the median of 0.009. The standard deviation is 0.113. The distribution is not normally distributed as a result of one huge outlier which will largely influence measures such as skewness and kurtosis.

We next examine the Autocorrelations (ACF) and Partial Autocorrelations (PACF) Functions for Terraced price changes. The ACFs and PACFs are shown in Figures 4.40 and 4.41.

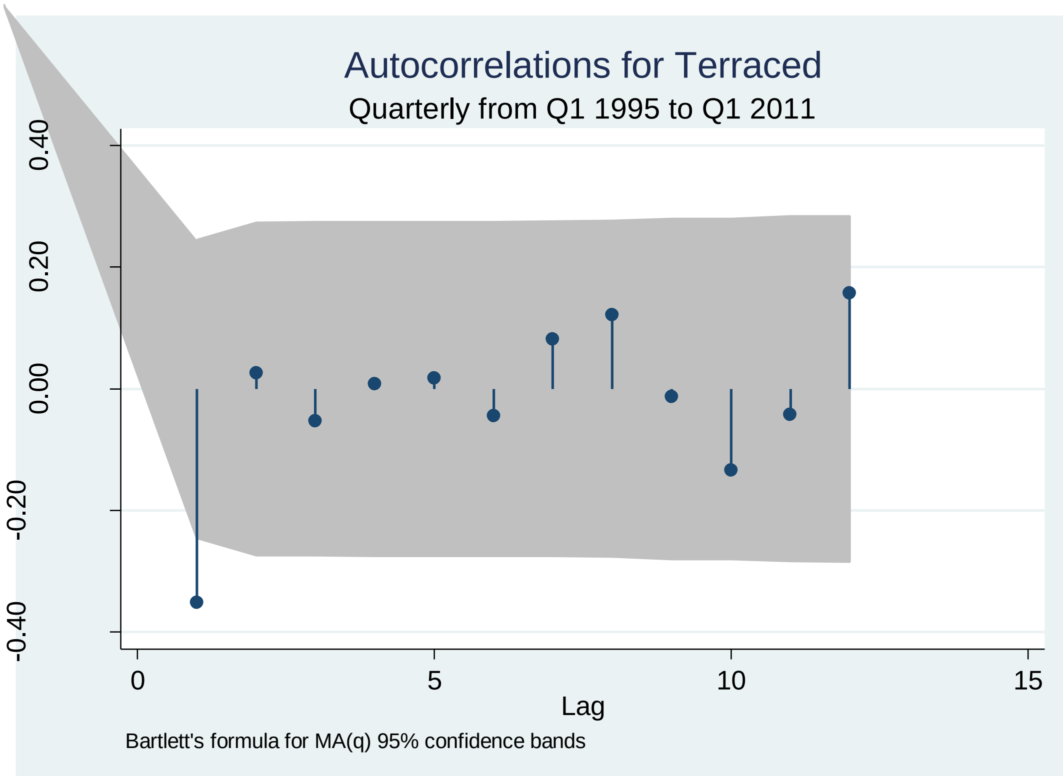

Figure 4.40: Autocorrelations for Terraced

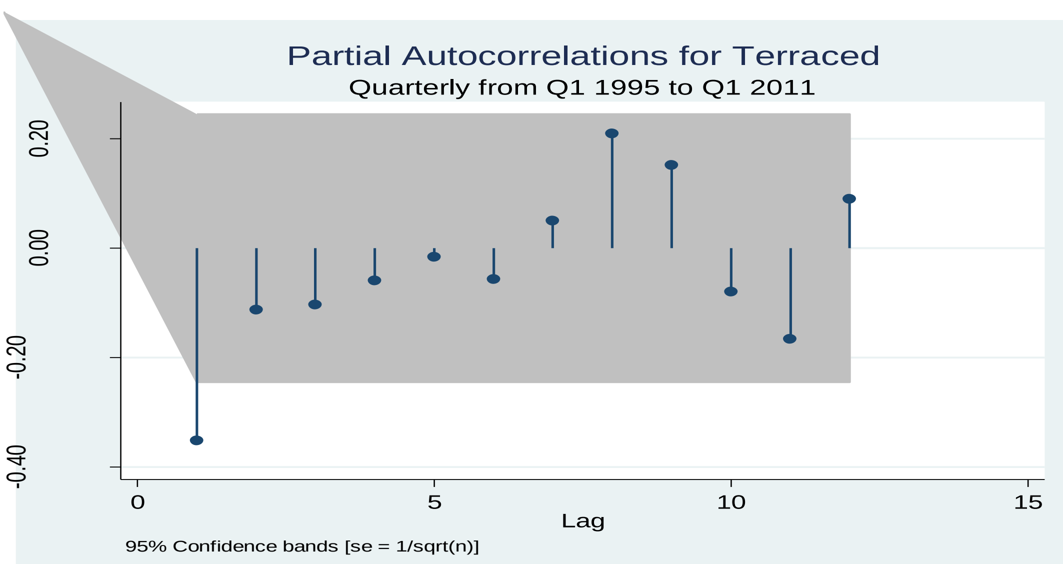

Figure 4.41: Partial Autocorrelations for Terraced

The ACFs show a significant lag at 1 in Figure 4.40. The PACFs show a significant lag at 1 in Figure 4.41. An ARMA(1,0) is estimated. The results are in Table 4.22.

Table 4.22: ARMA (1,0) Model for Terraced

Sample:1-64			Number of Observations=64	
			Wald chi2(1)=7.01***	
Log likelihood = 53.46057			Prob > chi2 = .0081	
Parameters	**Co-efficient**	**Std.Err.**	**z**	**P>|z|**
µ	0.0266118**	0.0126622	2.1	0.036
$ø_1$	-0.3469***	0.1310345	-2.65	0.008

Note: The test of the variance against zero is one sided, and the two-side confidence interval is truncated at zero.*** indicates significance at the 1% level, ** indicates significance at the 5% level, *indicates significance at the 10%

The Wald Chi-Square is significant at the 1%. The ARMA(1,0) constant and autoregressive parameter estimates are both significant at the 1% level. We next check for the adequacy of the residuals by examining their ACFs and PACFs in Figures 4.42 and 4.43.

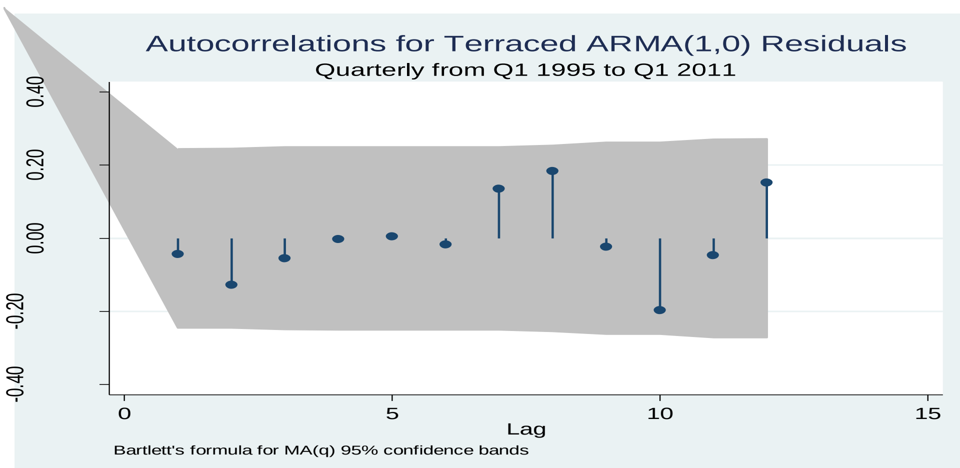

Figure 4.42: Autocorrelations for Terraced ARMA (1,0) for Terraced

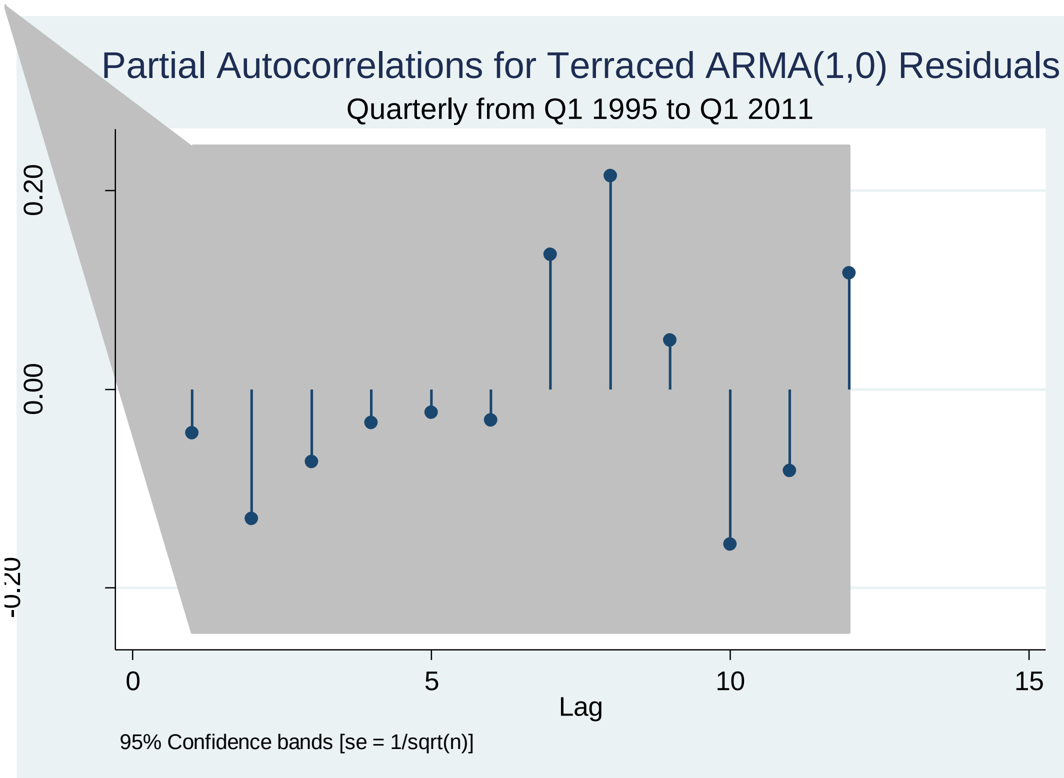

Figure 4.43: Partial Autocorrelations for Terraced ARMA (1,0) for Terraced

None of the ACFs and PACFs is significant at the 5%. We next examine for whether or not the residuals follow the assumption of the normal distribution. Shapiro-Wilk (1965) and Jarque-Bera (1980) tests for normality are reported in Tables 4.23 and 4.24.

Table 4.23: Shapiro-Wilk W Test for Normal Data

Variable	**Obs**	**W**	**V**	**Z**	**Prob>z**
Residuals	64	0.8448	8.885	4.726	0

Table 4.24: Jarque-Bera (1980) Skewness Kurtosis Test for Normality

Variable	**Obs**	**Pr (Skewness)**	**Pr (Kurtosis)**	**adj chi2(2)**	**Prob>chi2**
Residuals	64	0	0	36.78	0

The Shapiro-Wilk (1965) and Jarque-Bera (1980) test results indicate that the hypothesis of normally distributed residuals is not rejected. The results of Shapiro-Wilk (1965) and Jarque-Bera (1980) are confirmed by examining the normal QQ plot in Figure 4.44.

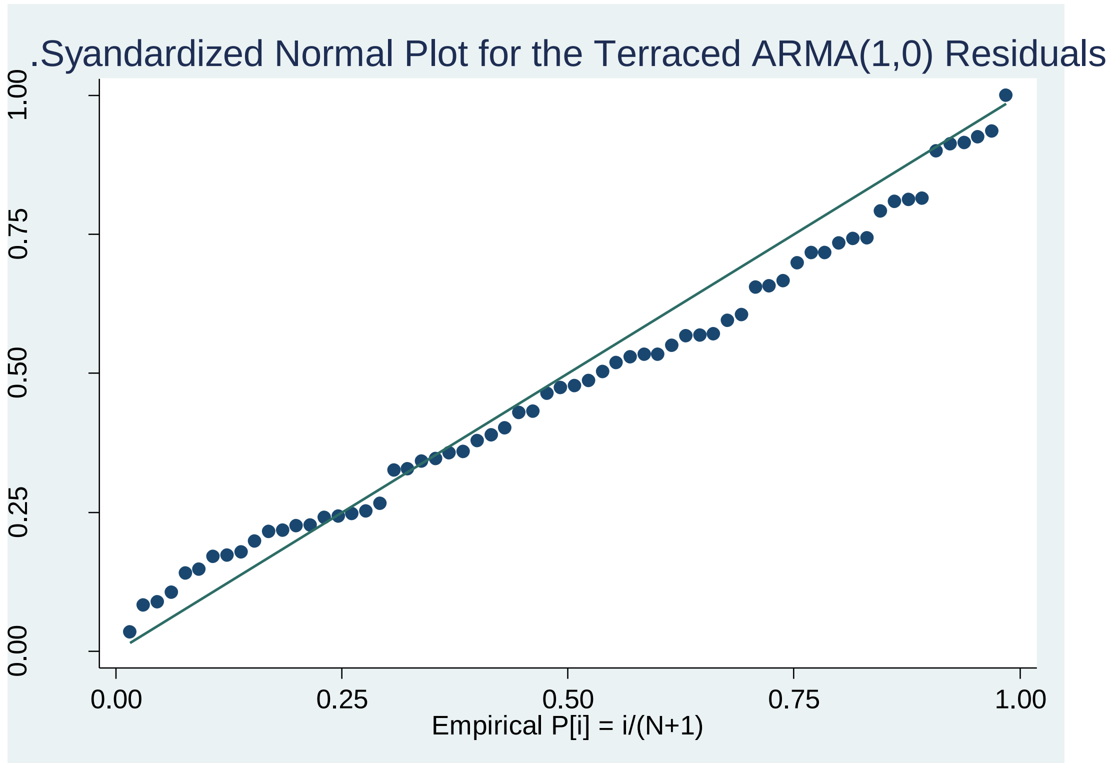

Figure 4.44: Standardized Normal Plot for Terraced ARMA (1,0) Residuals

The residuals approximately fall on the theoretical normal line. The ARMA(1,0) specification is 'accepted' for the Terraced houses as none of the ACFs are significant at the 5% level. This indicates a relatively a short memory as the price for the Terraced properties for the current quarter is determined by its previous quarter.

Following equation is for ARMA (1,0)

$$y_{t=}\mu + \emptyset_1 \, y_{t-1} + \epsilon_t \tag{4.8}$$

Where,

$y_{t=}$= price change in time t

μ = constant

y_{t-1}= price change at lag one

$\emptyset_1$ = auto regression coefficient estimate at lag one

ϵ_t = residual term

The residuals form the ARMA(1,0) model are then regressed using the following seasonal model.

$$y_t = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + u_t \quad (4.9)$$

where,

y_t= the residuals from the fitted model;

β= parameter estimates;

D_1=dummy variable 1 in quarter 1, and 0 in the other three quarters;

D_2=dummy variable 1 in quarter 2, and 0 in the other three quarters;

D_3=dummy variable 1 in quarter 3, and 0 in the other three quarters;

D_4=dummy variable 1 in quarter 4, and 0 in the other three quarters;

u_t= residual term with normal distribution.

Source	**SS**	**Df**	**MS**	**No.Of Obs.**	63
Model	0.104982153	4	0.026245538	**F(4,59)**	2.59
Residual	0.597091352	59	0.010120192	**Prob>F**	0.0455
Total	0.702073505	63	0.011144024	**R-Squarred**	0.1495
				Adj R-Squarred	0.0919
				Root MSE	0.1006
ARMA 1:Terraced Seasonality	**Coef.**	**Std.Err.**	**t**	**P> I t I**	
Quarter 1	-0.0400415	0.0251498	-1.59	0.117	
Quarter 2	0.0421847	0.0259746	1.62	0.11	
Quarter 3	0.0403098	0.0251498	1.6	0.114	
Quarter 4	-0.0408026	0.0251498	-1.62	0.11	

Table 4.25: Results of ARMA (1,0) Seasonality

The F-test is significant at 5% level but not significant at the 1% level. None of the t-tests is significant for any of the quarters. Accordingly, the ARMA(1,0) is accepted as a model for Terraced.

4.7 Testing for GARCH effects

We have carried out tests for Generalized Autoregressive Conditional Heteroscedasticity (GARCH) effects on the squared residuals from all ARMA models fitted for Semi-Detached, DetachedDetached, Flats and Terraced properties. The objective is that a series may not show patterns in its ACFs and PACFs of the first moment while there are significant patterns of correlations in its second moments. There were no significant GARCH effects found. To save space we only report the ACFs and PACFs for the squared price changes for Terraced. They are in Figures 4.45 and 4.46.

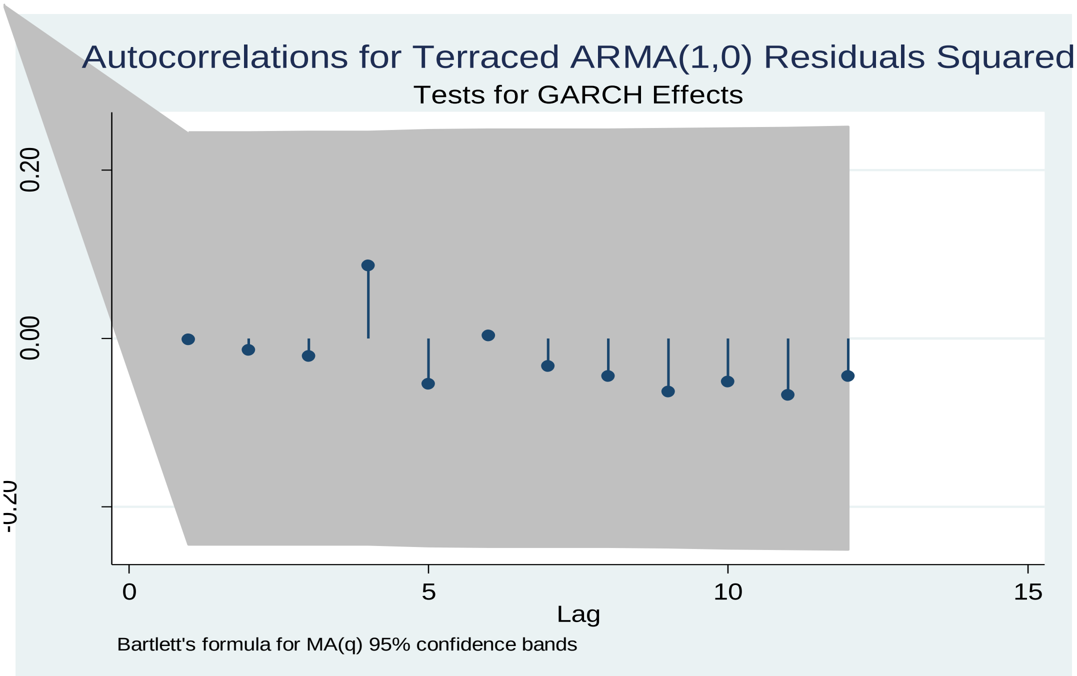

Figure 4.45: Autocorrelations for Terraced ARMA (1,0) Residuals Squared

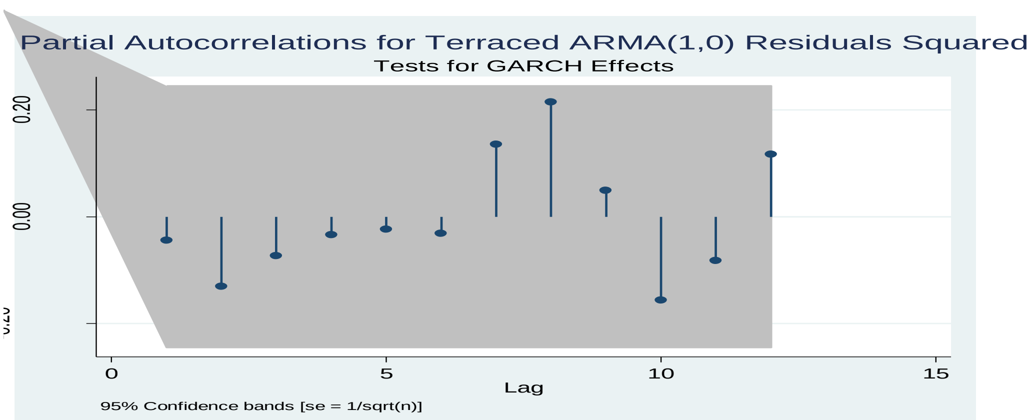

Figure 4.46: Partial Autocorrelations for Terraced ARMA (1,0) Residuals Squared

According to Figures 4.45 and 4.46 none of the ACFs and PACFs is significant, indicating the absence of volatility clustering (GARCH effects). The absence of GARCH effects were also found for all other time series including the UK All house price index.

4.8 Within City Correlation Analysis

The Chapter up till now has considered the univariate attributes of different property types within Manchester. Multivariate time series analysis requires far more observations than can be carried out within the current study. Fifty observations are usually regarded as the minimum needed for univariate time series since estimation methods such as BHHH of Brendt *et al.* (1974) are data hungry in their nature. Far more observations are required for multivariate analysis. There are sixty five quarter observations available from quarter one 1965 to quarter one 2011. However, bivariate correlation analysis between different property types within Manchester can provide some interesting facts about the co-movement of these properties. The correlations between property types are in Figure 4.47.

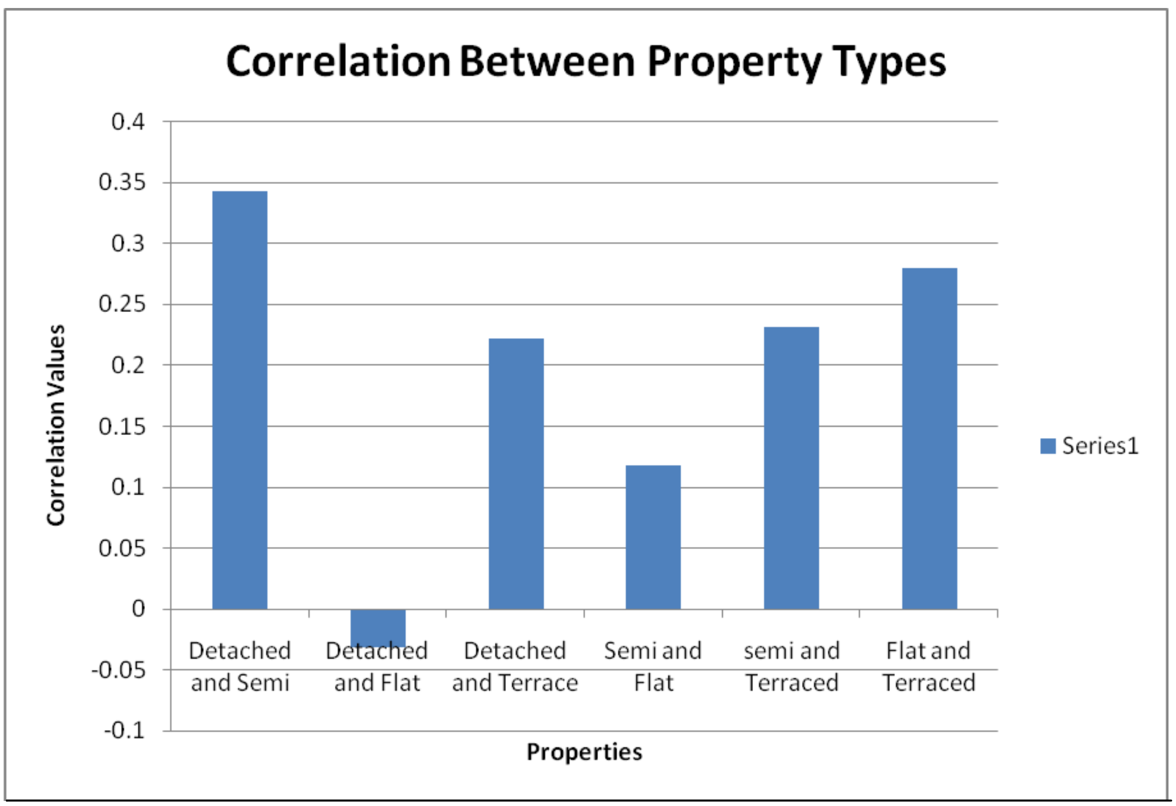

Figure 4.47: Correlation between Property Types

The above figure represents correlation between different property types, and shows the following correlations (highest to lowest): Detached and Semi-Detached, Flat and Terraced, Semi-Detached and Terraced, Detached and Terraced, Semi-Detached and Flat, Detached and Flat. It is interesting that price changes are positively correlated for Detached and Semi-Detached which are the close substitute property types for buyers as their first preference is for Detached houses and if this is not feasible they consider semi Detached properties. Also, for Flat and Terraced buyers there is a high correlation, as buyers prefer Terraced properties to buying Flats. The following Table (4.26) shows the quarterly and annual returns by property type for the data from quarter one 1995 to quarter 1 2011 (65 observations).

Table 4.26: Quarterly Return, Quarterly Risk and Quarterly Co-efficient of Variation of the Property Types

Variables	Quarterly Average Return	Quarterly Risk	Annual Average Return	Annual Risk	Coefficient of Variation
Semi	0.0205	0.0671	0.082	0.2684	3.28
Detached	0.0255	0.1312	0.1020	0.5248	5.13
Flat	0.0206	0.0823	0.0824	0.3292	4.04
Terraced	0.0263	0.1129	0.1052	0.4516	4.34

Quarterly return is the mean return of the property types. Equivalent yearly return is calculated by multiplying mean with the number of quarters i.e. four. Quarterly risk is the standard deviation and annual risk is four multiplied by quarterly risk. Coefficient of variation is calculated by dividing quarterly risk (standard deviation) by quarterly return (mean). From the Table, it can be said that highest annual return is from Terraced and Detached properties (10.52% and 10%) followed by Flats and Semi-Detached properties (8.2% and 8%) respectively. The annual risk is highest for Detached properties followed by Terraced, Flat and Semi-Detached properties. The highest risk per one percent of return as indicated by the coefficient of variation is for Detached properties followed by Terraced, Flats and Semi-Detached properties. The only surprise is that we would have expected Semi-Detached properties to be second in terms of risk per unit of return instead of last.

4.9 Forecasting out of Sample

The purpose of this section is to test the ability of the identified Seasonal ARMAmodels in beating naïve forecasting models using out of sample data for the UK All house price index and each of the property types within Manchester (Semi-Detached, Detached, Flats, and Terraced properties). The out of sample data are quarterly starting from second quarter 2011 to last quarter 2013. The out of sample size has eleven observations. Table 4.27 has the forecasting performance in comparison to naïve mode. The naïve model is in equation 4.10.

$$\hat{y}_t = y_{t-1} \tag{4.10}$$

The naïve model assumes that best forecast value for the next quarter is equal to the previous observed quarter value.

The forecasting accuracy over the naïve model is judged by the root mean squared error which is in equation 4.11

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \tag{4.11}$$

The ARMA estimated model will be accepted if it beats out the naïve model out of sample. Table 4.27 compares the RMSE between the naïve model and the forecasted ARMA models.

Table 4.27: A comparison between the root mean squared error (RMSE) for the forecasted ARMA versus the Naïve forecast.

	ARMA	**NAIVE**	**(ARMA/NAÏVE) RMSE**
UK ALL	0.013028	0.007557	1.70
Semi-Detached	0.118359	0.167534	0.70
Detached	0.262608	0.419065	0.62
Flats	0.067862	0.119003	0.58
Terraced	0.092422	0.159651	0.56

The ARMA for the UK ALL out of sample has failed to improve on the forecast from the naïve mode. It is 70% worse than the naïve model. That is despite the model being acceptable within sample data. The picture is different for property types within the Manchester City. The ARMA models beat the naïve forecasting out of sample by 30%, 38%, 42%, and 44% for Semi-Detached, Detached, Flats, and Terraced properties.

4.10 Summary

The Chapter handled research problems one and two, lack of studies of real estate price determination by residential property types within a city and lack of studies of seasonality of real estate prices by property types within a city. The chapter applied the Box-Jenkins (1974) methodology to the UK

All house price index in addition to different property types (Semi-Detached, Detached, Flats and Terraced properties) in Manchester during the period from 1995 (quarter one) to 2011 (quarter one). Manchester was chosen since it is the second largest city in the UK after London according to the rules used by the office for national statistics, ONS (2011). Manchester was chosen since London house prices are more influenced by international factors than Manchester due to massive foreign investments. The literature review in Chapter two indicates that most research on real estate in the UK concentrated on correlations between regional properties and very few examined residential property investments. This thesis is the first to examine residential properties within a region or a city in addition to the UK All house price index. The autocorrelations functions (ACFs) and partial autocorrelations functions (PACFs) of price changes were examined for each property type. The results of the model building using Box-Jenkins (1974) methodology for the UK All house price index in addition to the four residential property types within Manchester (Semi-Detached, Detached, Flats and Terraced properties), are summarised in the following Table 4.28.

Table 4.28: Box Jenkins Analysis of The Properties

Property Type	**ARMA(p,q)**	**Parameter Estimates Along-with Their Significance Level. ***, **, and * for 1%, 5%, and 10%, respectively.**	**Seasonality. Significant Quarters only with Their Significance Level. ***, **, and * for 1%, 5%, and 10%, respectively.**	**Adjusted R-Squared and F-Test, ***, **, and * for 1%, 5%, and 10%, indicates significance respectively.**
UK All House	**ARMA (1,1)**	0.4943079, 0.5713402	None is significant	Adj-Rsqaured = Not Applicable
Semi-Detached	ARMA(5,0)	AR(1)= -0.446***, AR(2)= -0.220*, AR(3)= 0.109, AR(4)= 0.470***, AR(5)= 0.432***	None is significant	Adj-Rsqaured=0.34 F-test=6.95***
Detached	ARMA(4,0)	AR(1) = -0.656*** AR(2) = -0.607***, AR(3) = -0.226 AR(4) = -0.335**	Q(3)=0.0974***	Adj-Rsqaured=0.47 F-test=9.86***
Flats	ARMA(1,0)	-0.391***	None is significant	Adj-Rsqaured=0.14 F-test=11.38***
Terraced	ARMA(1,0)	-0.347***	None is significant	Adj-Rsqaured=0.11 F-test=8.60***

The UK All house price changes are modeled using ARMA(1,1) model. Adjusted R-squared is not well defined in the case of ARMA model due to the inclusion of past error terms in the regression of the model. Although the ARMA(1,1) passes all comprehensive diagnostic checks within sample, the model fails to predict out of sample. There was no seasonality quarters identified within the UK All house prices. Forecasting out of sample was more successful for property types within Manchester.

The results show that Manchester's most expensive housing type (Detached) experienced more negative price declines than the less expensive Semi-Detached houses and the least expensive properties such as Terraced houses and Flats. This is shown in Table 4.28 as Detached properties experienced four negative quarter correlations while the Terraced houses and Flats experienced just one negative quarter correlation. This means the price decline for Detached property took year to show positive price change while for Flats and Terraced properties it only took a quarter to show a positive price changes. The autocorrelations for the Semi-Detached properties which are closer to the most expensive type of property (Detached) than the least expensive (Flats) showed a mixed pattern with the correlations with quarters one and two being negative while the correlations with quarter four and five are positive. Semi-Detached property seems to recover faster than Detached properties but slower than the Terraced houses and Flats. The Detached and semi Detached property past history explains more of the variation in the current prices than for Flats and Terraced houses as indicated by the adjusted R-squared. The reason for this could be due that Flats and Terraced houses are much more affordable and therefore variation in prices could be more related to causal variables such as mortgage rates and credit conditions of the market. The Detached and Semi-Detached properties are much more expensive which may mean they can have less sensitivity to mortgage and credit markets conditions, and therefore more dependence on past history than Flats and Terraced properties.

One limitation of the time series analysis carried out in this Chapter is that the period of study from 1995 (quarter one) to 2011 (quarter one) include the credit crunch and the subsequent mortgage problems from quarter four 2008. Unfortunately, due to limited quarterly data it is not possible to divide the data into two sub-samples, pre and post-crash since time series techniques require at least fifty observations for estimation while we only have sixty four. The residual analysis of all models indicates adequacy with no identified patterns in ACFs and PACFs of the residuals. Dividing the sample pre and post credit crunch would have led to possible positive correlation pre and negative correlations post but with no impact on the stability of our models which have passed all comprehensive diagnostic tests. Causal models can be more successful at handling unusual period of trouble such as the credit crunch as the model tries to capture the cause and effect relationships between real estate prices and the factors that influence them such as mortgage rates, personal debt levels and unemployment. These causal models will be explored in more details in Chapter five.

Chapter 5: A New Socio-economic model for real estate price behaviour

5.0 Introduction

This chapter deals with the thesis research problem number three which is lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting. The chapter provides a new model to the literature by combining GIS tools with multiple regression techniques. The six steps methodology suggested in chapter three are implemented. The steps are statement of research problem, gathering relevant data, estimation methods, diagnostic tests for the residuals, model evaluation, and out of sample forecasting. Research problem three is reflected in the thesis objectives two and three which are; (1) to introduce a methodology that combines published data from the UK office for national statistics with the geographic maps (GIS) to identify causal relationships that can be used in forecasting real estate prices, and (2) to provide a new approach utilizing time series techniques, causal models and GIS tool to model and forecast real estate price changes.

Multiple regression methodology that emphasizes relationships between dependent and explanatory variables were presented and discussed in Chapter three. The objective from the causal modeling between real estate price as a dependent variable and its hypothesized explanatory variables is to use the causal relationship to forecast real estate prices out of the sample. The challenge is to develop an understanding of what drives real estate prices and can be used as explanatory variables. The literature review in Chapter two and the past causal models that were summarized in Chapter three use macro-economic variables such as gross domestic product (GDP) and consumer price index (CPI). This Chapter contribution is in exploring the causal relationship for real estate prices using Geographic Information Systems (GIS). The identified relationships from the GIS applications will be tested by the multiple regression methodology.

Real estate is not just about price fluctuation. There are factors that may influence real estate prices more than its fluctuations. For example, location is one such factor. The importance of location has been discussed in the literature review Chapter. Finding the real estate location that can have a positive net present value for the investor is an important question that the advanced technology of GIS has managed to address. A thematic map displays the picture of the potential area for investment as per a particular theme, say for example personal consumer debt. The thematic map makes a simple area of a map look meaningful by relating its location to causal socio-economic data from the Office for National Statistics (ONS). These causal factors are called explanatory variables. The focus of section 5.1 is on identifying a probable cause and effect relationship between socio-economic explanatory factors and location (and therefore real estate price). Examples of the relationship between socio economic parameters and real estate location and price are economic deprivation and personal debt. Individuals prefer to buy properties in locations where fewer residents live on council benefit and where people have less debt (Can, 1998).

Since housing is fixed to a specific location the importance of neighbourhood with respect to the operation of housing and mortgage markets cannot be ignored (Can, 1998). Location is a major issue in household residential satisfaction. Neighbourhoods are defined by Can, 1998 as separate objects which possess households and housing structures with the same features. Typically, households exhibit similar social, economic and demographic characteristics within neighbourhoods (Can, 1998). According to Can, 1998; there are four factors which lead to positive or negative mindset on residents: (1) accessibility to daily use social amenities like transport; (2) surrounding environment; (3) social, economy and demography; and (4) public service provision.

The use of GIS in the domain of real estate is measured by the availability of consistent information. With the help of GIS, comparable data with specific characteristics can be selected using Structured Query Language (SQL) (ESRI ArcGIS Manual, 2006). Through GIS, risk can be reduced through initial identification and visualization of the localities as per their corresponding socio-economic factors. GIS helps to map relationships between demographics, household income of a particular location and investment in real estate (De Man, 1998). A GIS combines location and socio-economic data to create thematic maps showing a wide range of data relating to population, housing and economic deprivation, council tax claimants etc.

GIS has been used in choosing a particular or a potential site from a group of related sites. The technology has been used in potential sites attribute analysis, created to help in the selection of sites, analysis for location of residential subcategories (Barnett and Okoruwa, 1993), and advanced systems for managerial performance have been incorporated. GIS has been developed to analyse whether real estate is situated in a locality, which is sensitive to natural calamities, and to calculate insurance rates based on analysis of the surrounding locality crime accounts and the distance to fire hoses, fire brigades, and police stations (Francica, 1993; Kochera, 1994).

GIS maps consist of a point, line and polygon (Marks, 1994; Podor, 2010). A layer can be a point, a polygon or a line. A line can be defined as a shape having length and direction but no area, connecting at least two points or geographic co-ordinates. Examples of such layers are roads, railways etc. A point is a digital representation of a place that has location but is too small to have area or length at a particular scale, such as a city on a world map or a building on a city map. A polygon has geometry which is significantly larger than a point on the map scale. The important layer in the thesis is the postcode polygon which has been geo-referred with xyzmaps.com and ordnance survey maps of the UK. Normally such a layer is created by creating an image of the main roads or streets of a locality. There are 46 such districts for Manchester. The districts go by the following distribution: M0-M9, M11-M35, M38, M40-M41, M43-46, M50, M60, M90 and M99.The maps in the appendix show the spatial extent of the district. These postal boundary maps will be used to link with the ONS database to generate the thematic maps of the respective areas.

Before the thematic maps are created, it is important to understand the classification of Manchester as per the Office for National Statistics U.K. This classification says Manchester can be divided into areas such as middle super output areas (MSOAs) and lower super output areas (LSOAs). Terminology such as LSOA and MSOA are important because the latter i.e. MSOA is a common attribute to both the area maps (mentioned in the above paragraph) and the socio-economic factors which are being defined by the ONS. The following Table (Table 5.1) obtained from the website of the ONS explains the LSOAs and MSOAs. The classification has been obtained from Office for National Statistics.

Table 5.1 ONS Definitions of lower super output areas (LSOAs) and middle super output areas (MSOAs)

Geography	**Minimum Population**	**Maximum Population**	**Minimum Number of Households**	**Maximum Number of Households**
LSOA	1000	3000	400	1200
MSOA	5000	15000	2000	6000

- Lower Super Output Area (LSOA)

  As per the Office for National Statistics, the LSOAs were originally built using Census data from groups of Output Areas (typically four to six). They had a minimum size of 1,000 residents and 400 households, but average 1,500 residents.

  The Manchester City has 297 such areas.

- Middle Super Output Area (MSOA)

  As per the Office for National Statistics, there are now 7,201 Middle Super Output Areas (MSOAs). Manchester has 60 such areas. In the thesis only MSOAs are used to create the thematic maps since the ONS database deals only with MSOAs. Figure 5.1 shows the MSOAs for Manchester. The base map has been obtained from the Ordinance Survey UK website (www.ordnancesurvey.co.uk). The red boundary identifies M009 as an example. ArcGIS 9.2 software is used.

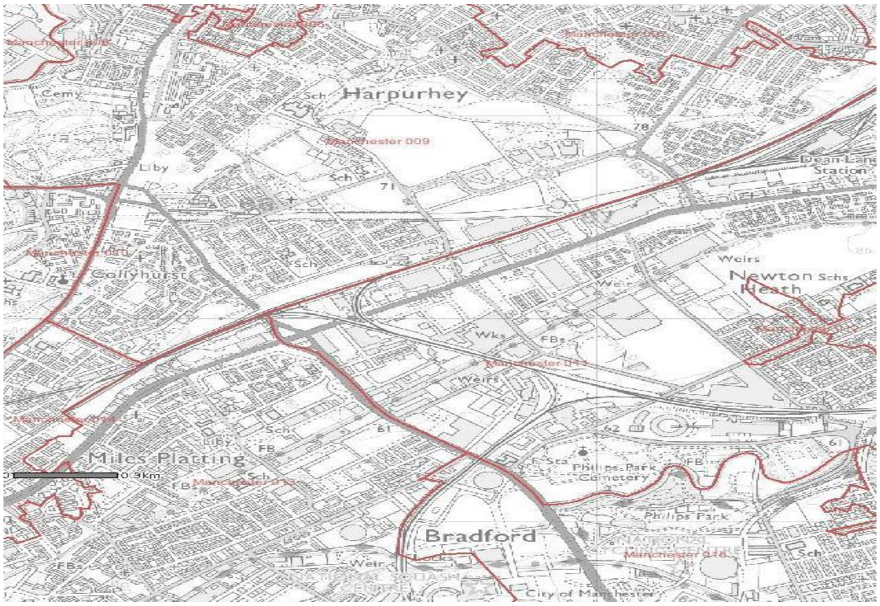

Figure 5.1: MSOAs of Manchester

Four MSOA areas of Manchester were chosen for the empirical analysis of the causal relationship between price and socio-economic variables. There are 21 Figures for each MSOA. To avoid repetition and save space, the maps for M002, M020, and M050 are not reported. The analysis can be easily extended to cover all MSOAs of Manchester. The causal relationship of these four MSOAs provides an answer to how the socio-economic factors behave with each other. The 4 MSOA areas are chosen so that two are adjacent in the north, one in the middle and one in the south of the Manchester map. The two adjacent areas, Higher Blackley (M001) and Blackley (M002), provide nearness of two localities. The one in Stanley Grove (M020) and the one in South Wythenshawe (M050) provide a basis for comparison with those in the north.

5.1 Estimation Methods

The objective of this section is to test using multiple regression analysis casual relationships between real estate prices on the one hand and explanatory variables such as socio-economic dimensions on the other. The established causal relationships will be tested using historical data in this chapter. The literature review behind the causal model was discussed in chapter two and the required data for analysis was discussed in data section in chapter four. The areas chosen for the analysis, as explained in section 5.0, are Higher Blackley (M001), Blackley (M002), Stanley Grove (M020) and South Wythenshawe (M050). The choice of these four areas was discussed in section 5.0. The seven domains used as explanatory variables were defined and discussed in chapter two section 2.6. The domains are:

- Crime and Safety
- Economic Deprivation
- Health Care
- Housing
- Personal Consumer Debt
- Social Grade
- Environment

These domains contain the data of each of the explanatory variables as per respective MSOA (middle super output area). These domains are then linked with the area maps (mentioned in 5.0, Introduction) through the ArcGIS software to generate thematic maps of each Middle Super Output Areas. The symbols and the legends for the thematic maps will be created by the default symbols and legend-shades present in the software. Each of the thematic maps will give us an entire picture of the full MSOA. The maps help to visualize the area remotely i.e. without actually going to the site one can see the picture of not only a particular locality but also localities surrounding the same.

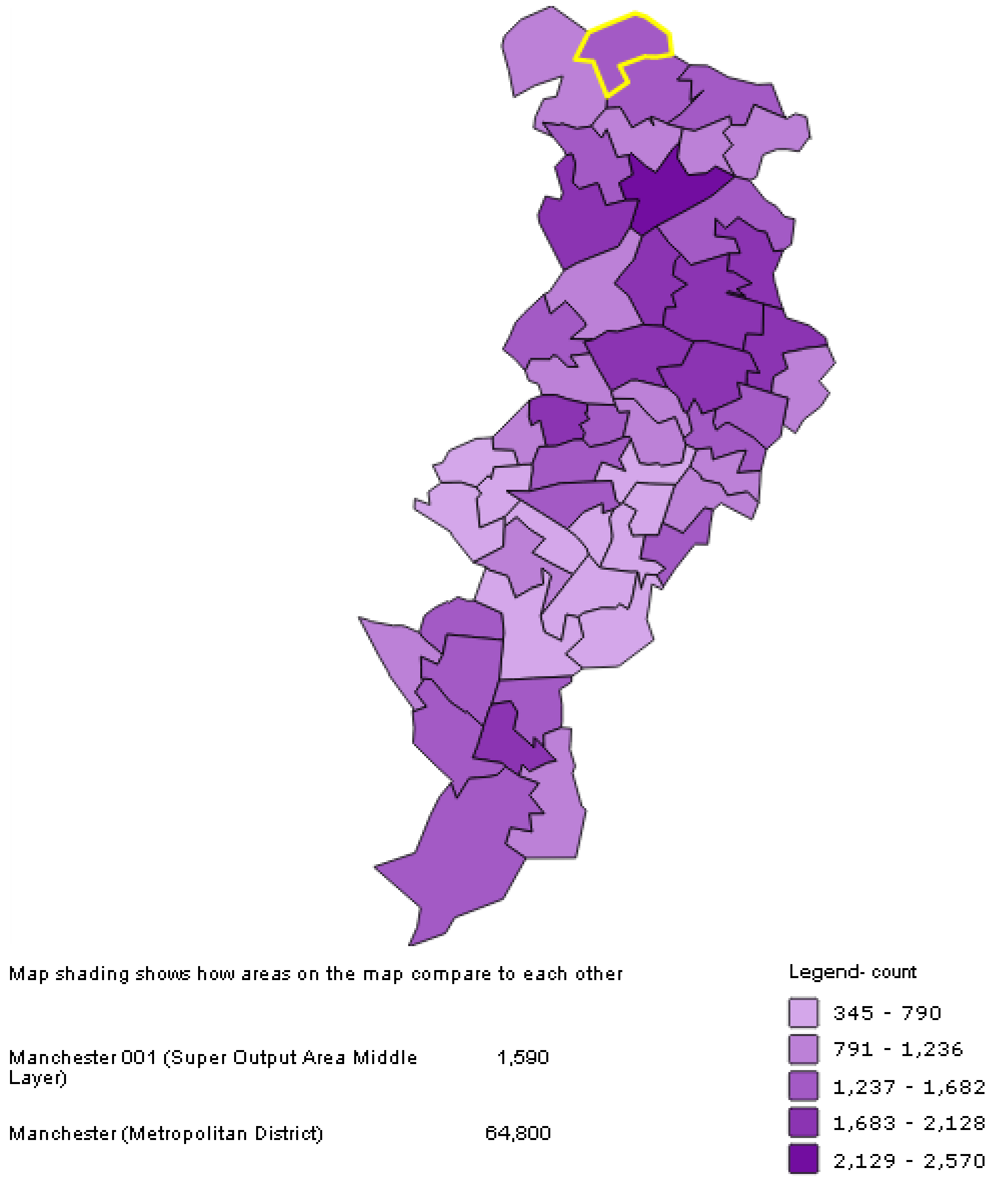

Figure 5.2: This map shows the total number of claimants in receipt of Housing Benefit or Council Tax Benefit where the legend count gives the number of claimants

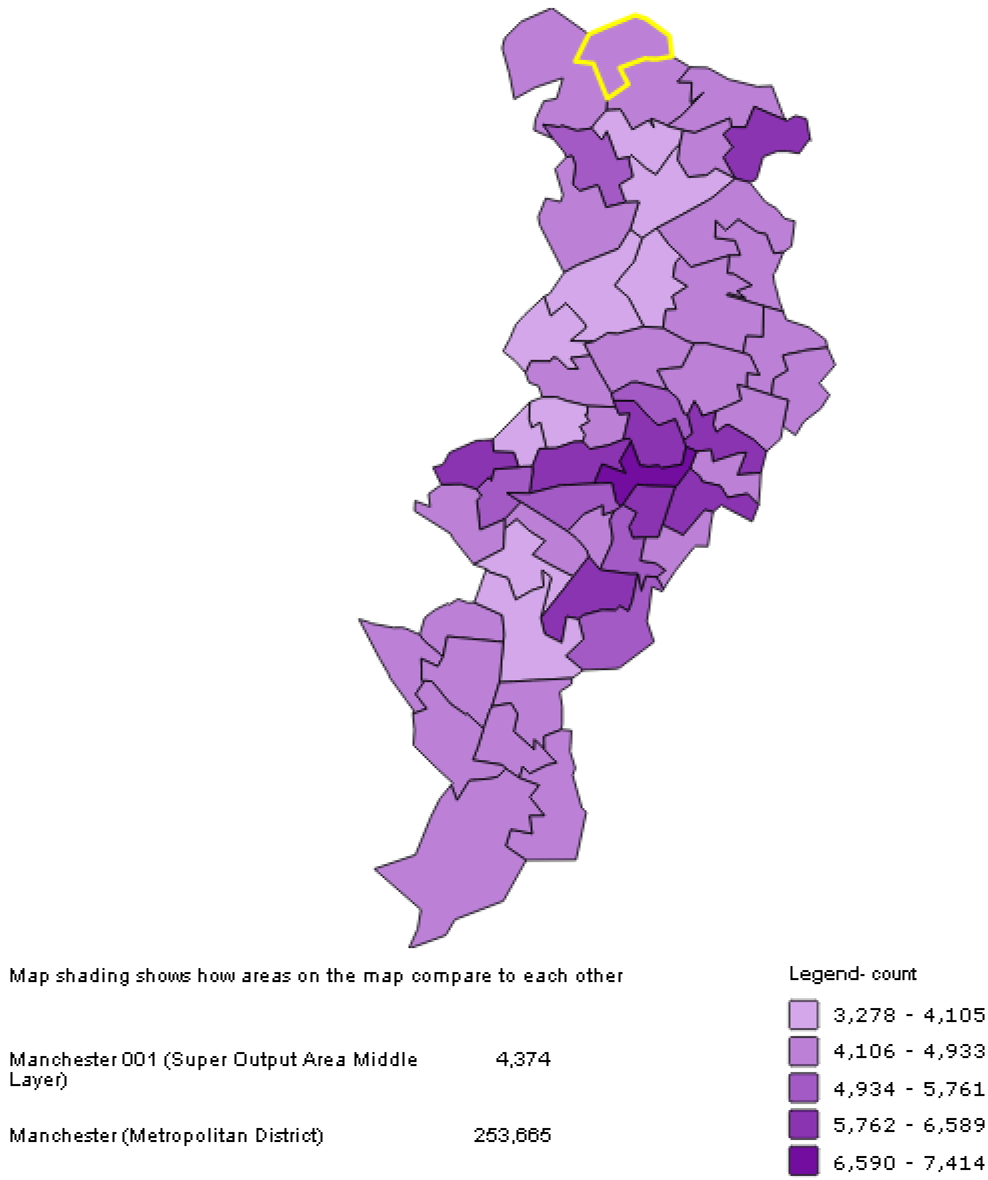

Figure 5.3: Individuals' counts according to good health as provided by the ONS

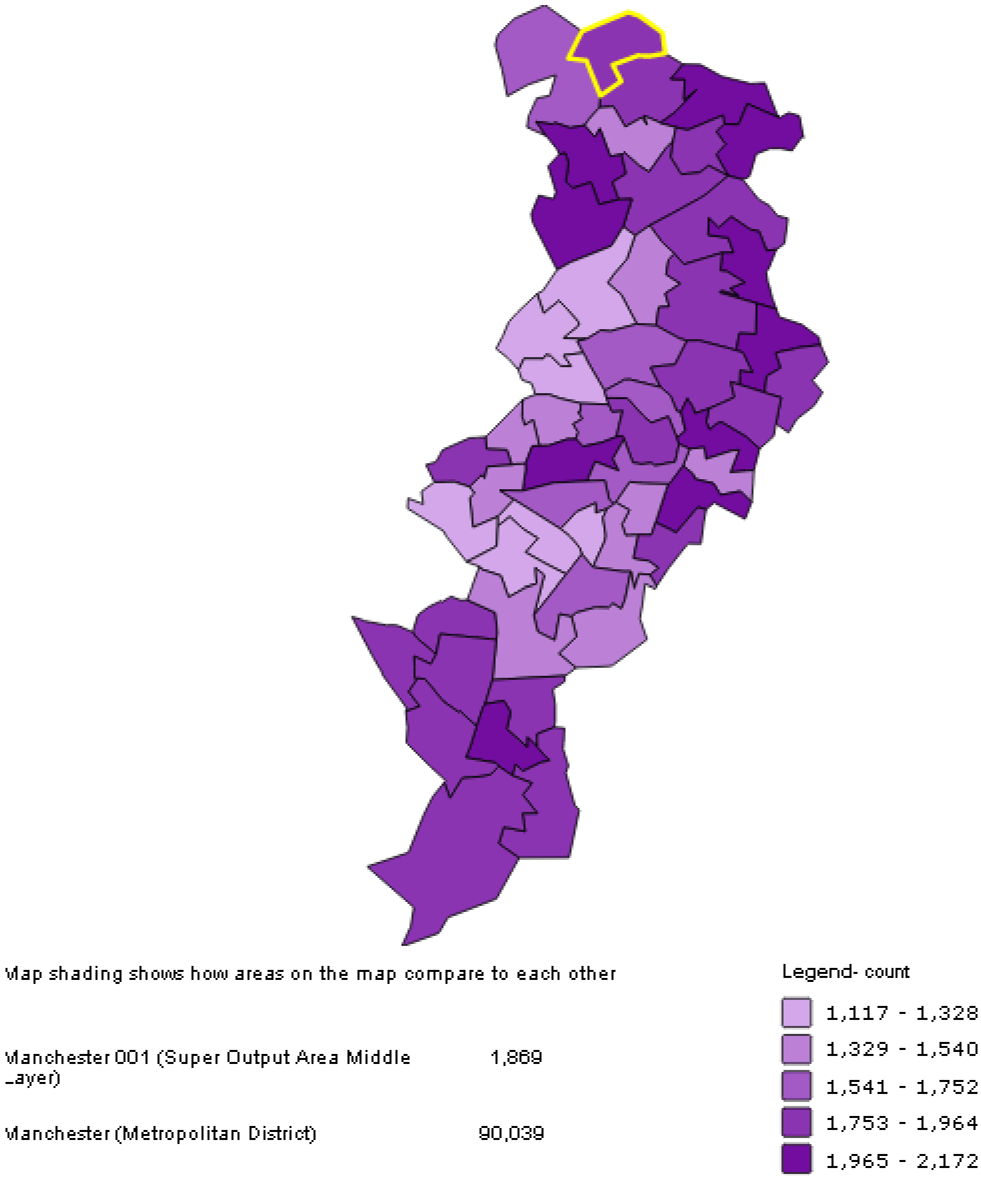

Figure 5.4: Individuals' counts according to fairly good health as provided by the ONS

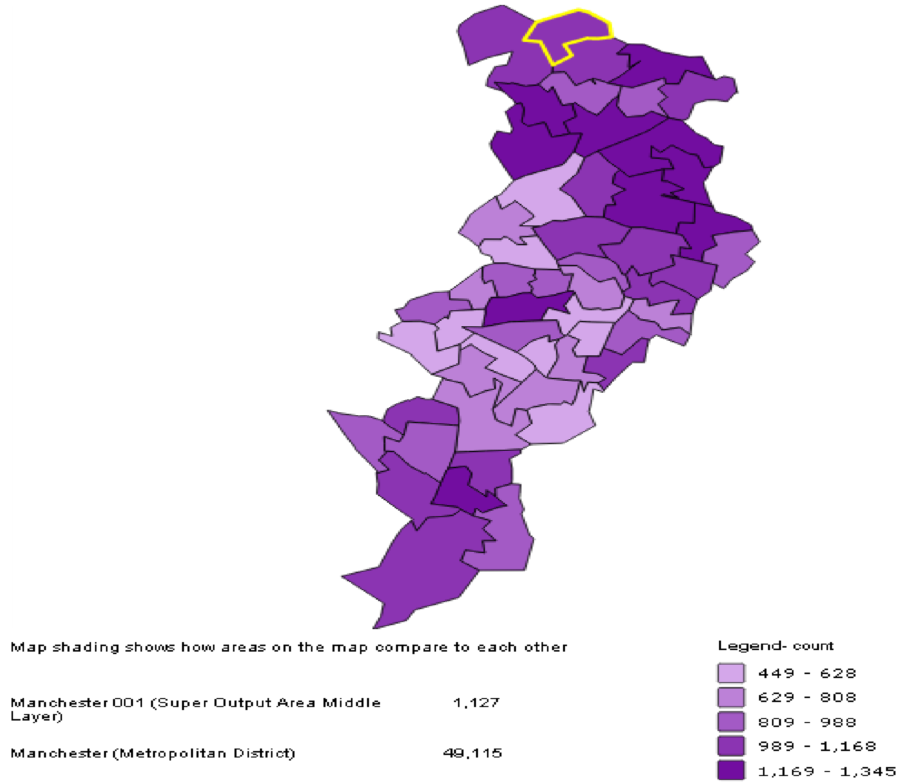

Figure 5.5: Individuals' counts according to not-good health as provided by the ONS

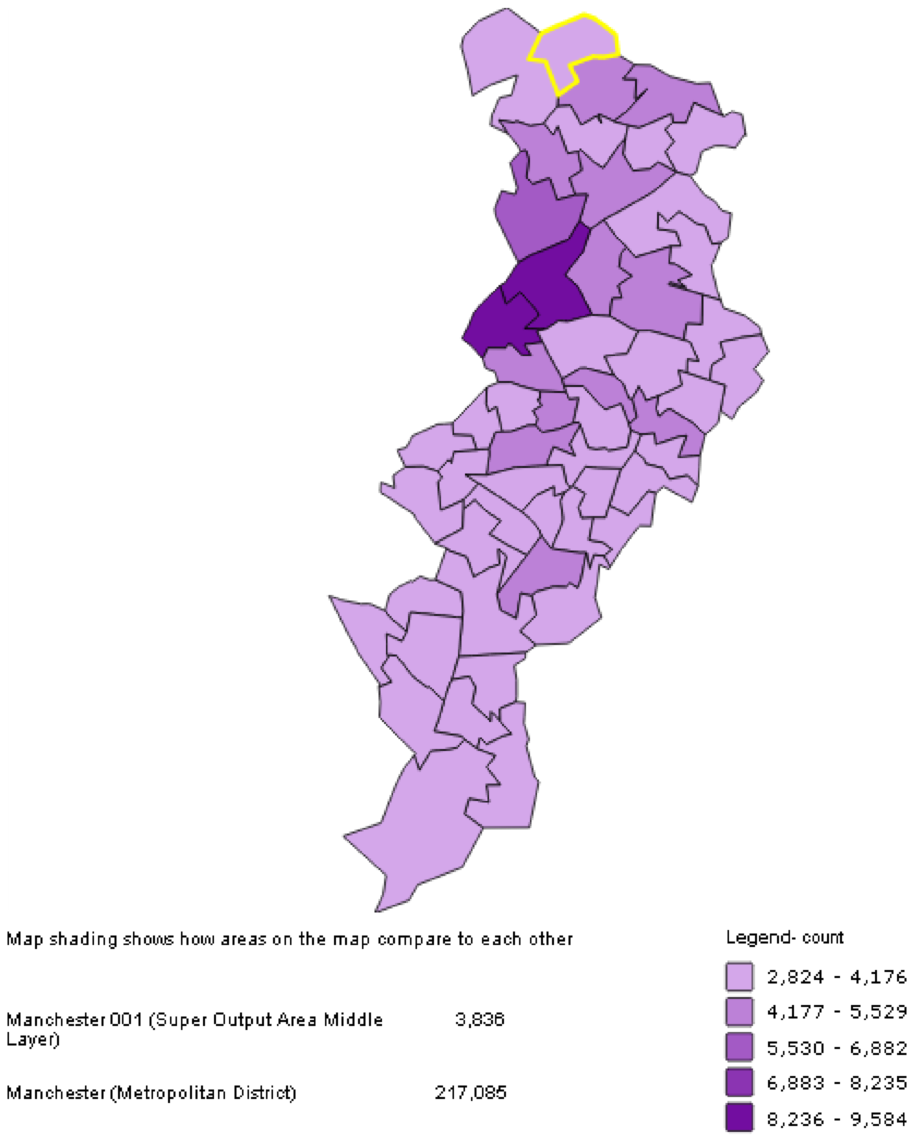

Figure 5.6: Map Showing Dwelling Stock (total number of real estate units)

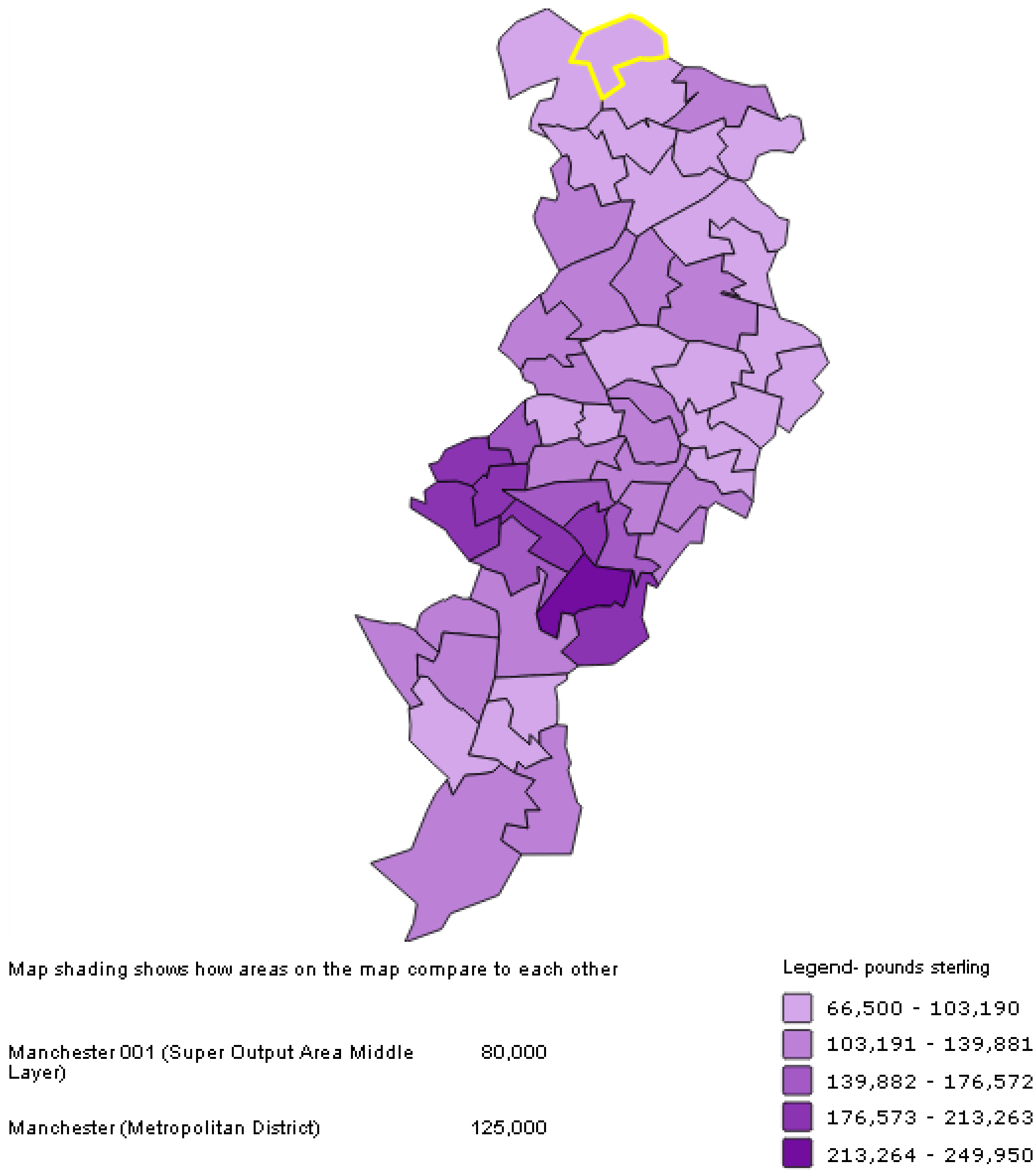

Figure 5.7: Map Showing Average Price for All Dwellings

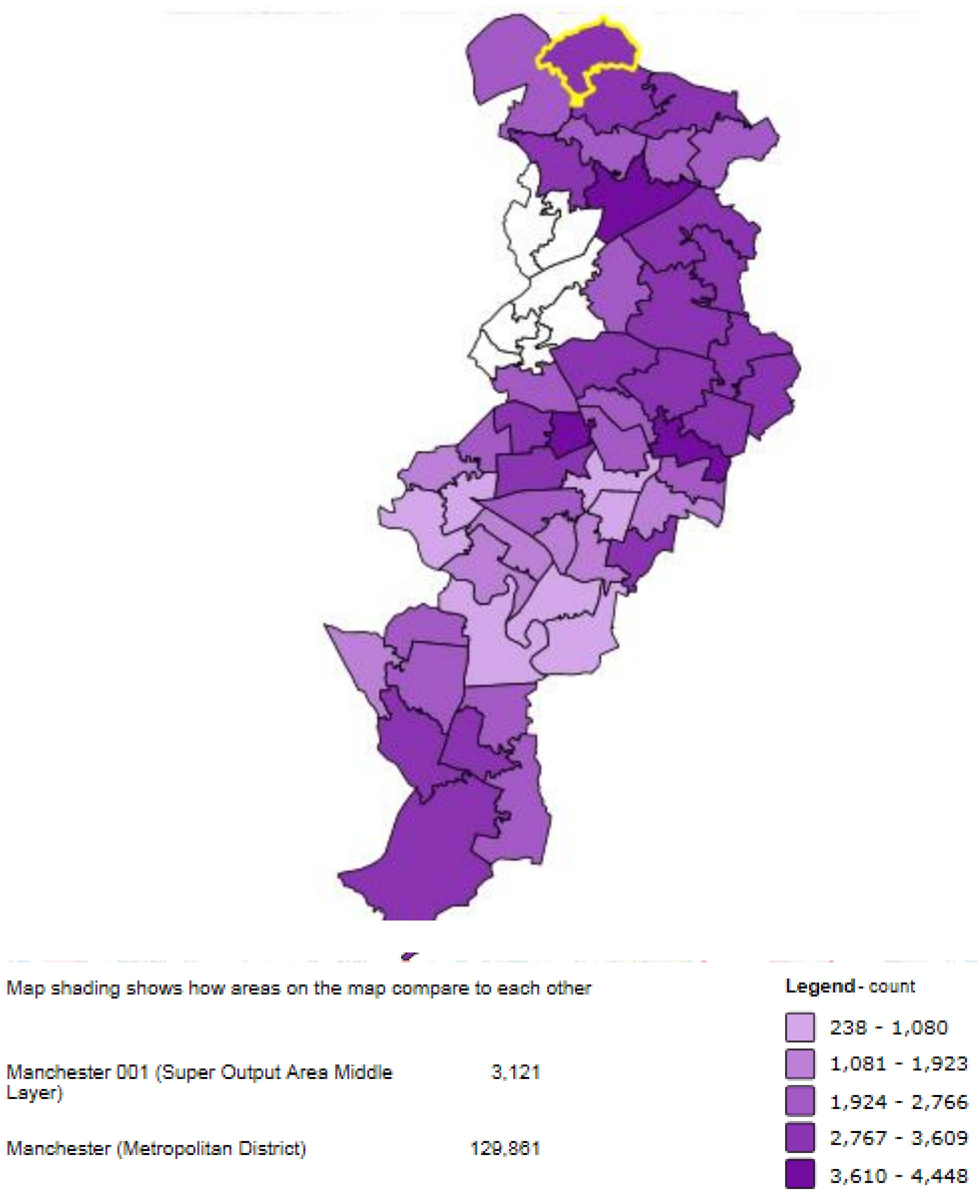

Figure 5.8: Map Showing Number of Dwelling Stock under Tax Band A; the legend count gives the number of dwellings

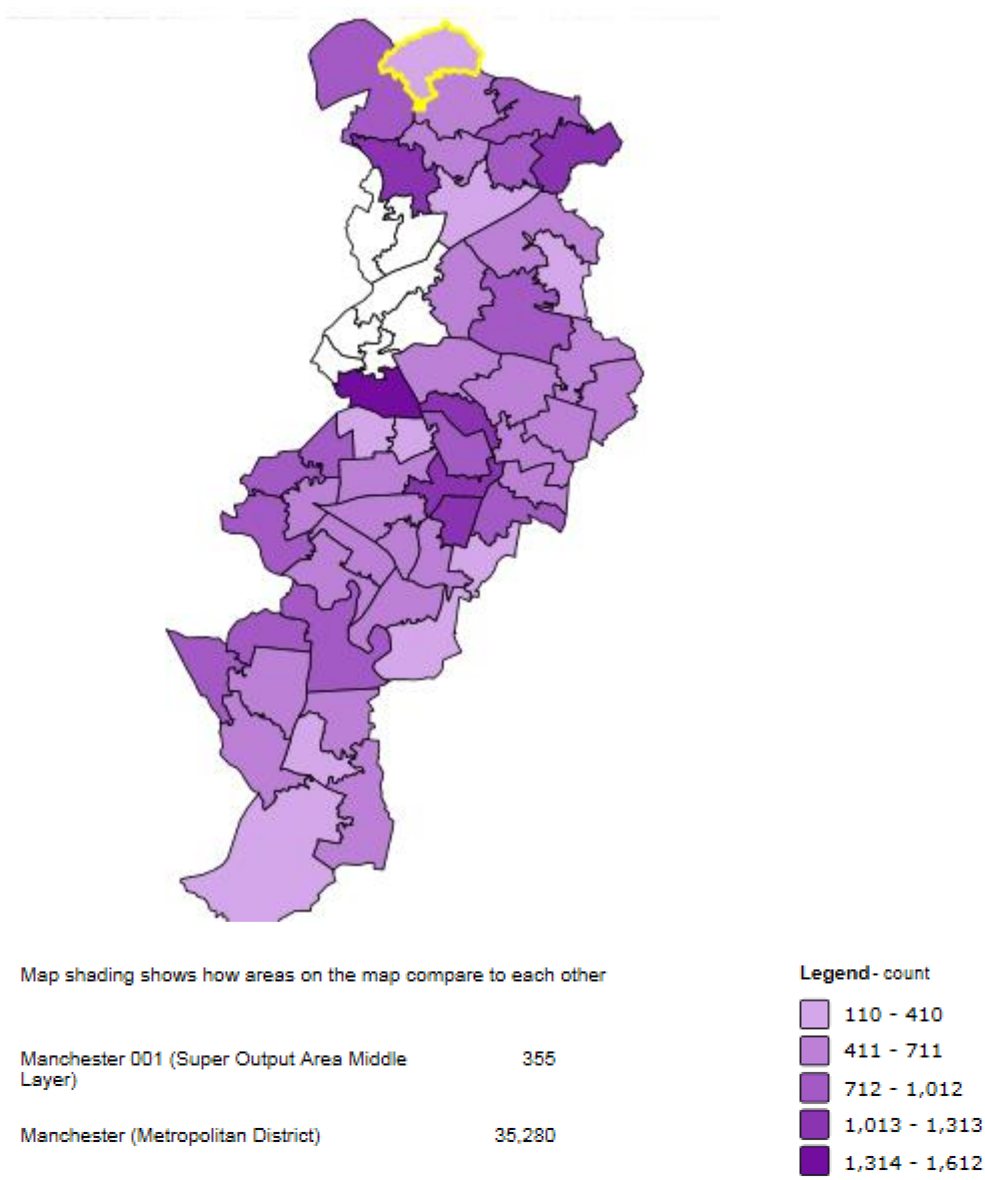

Figure 5.9: Map Showing the Number of Dwelling Stock under Tax Band B; the legend count gives the number of dwelling stock

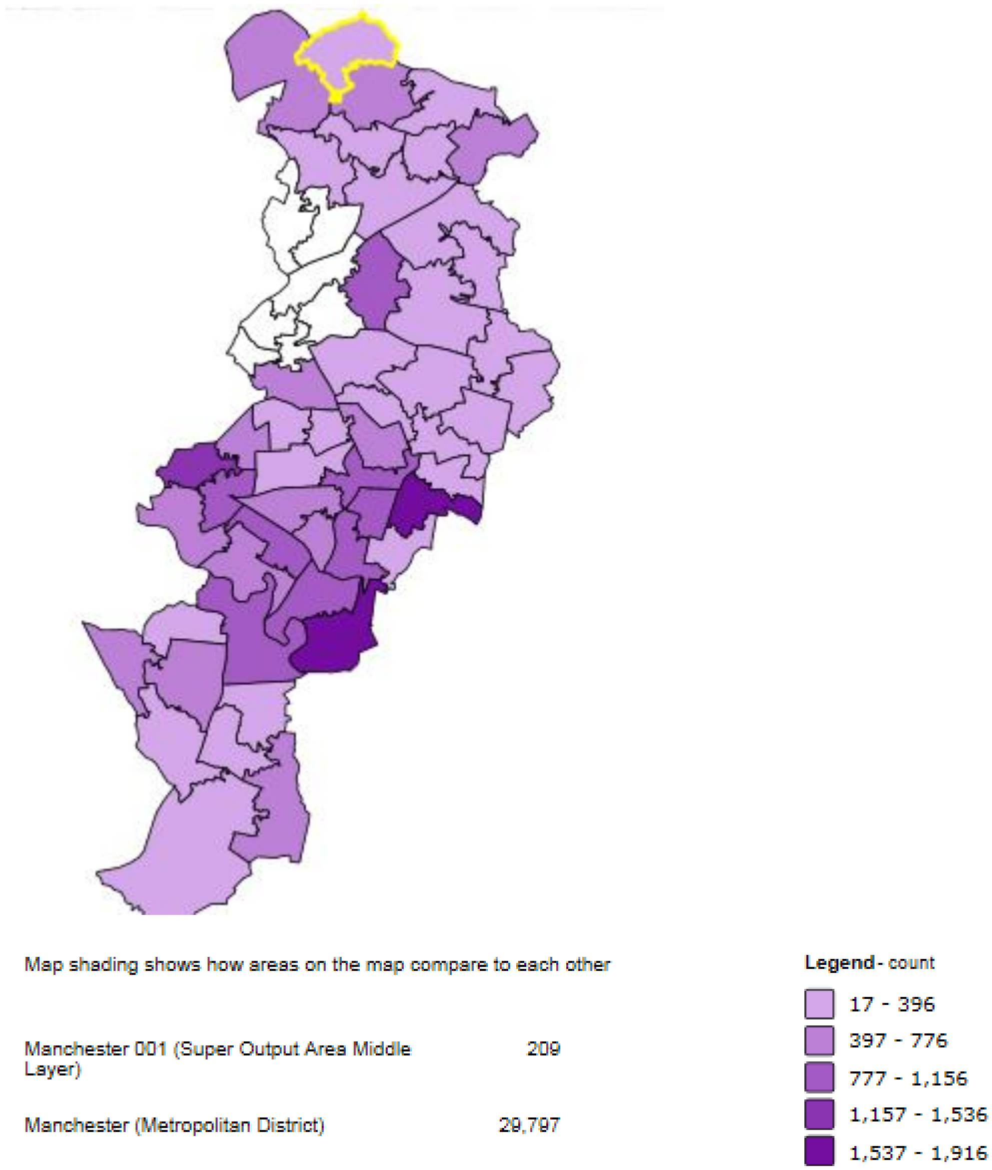

Figure 5.10: Map Showing Number of Dwelling Stock under Tax Band C and; legend count gives the number of dwelling stock

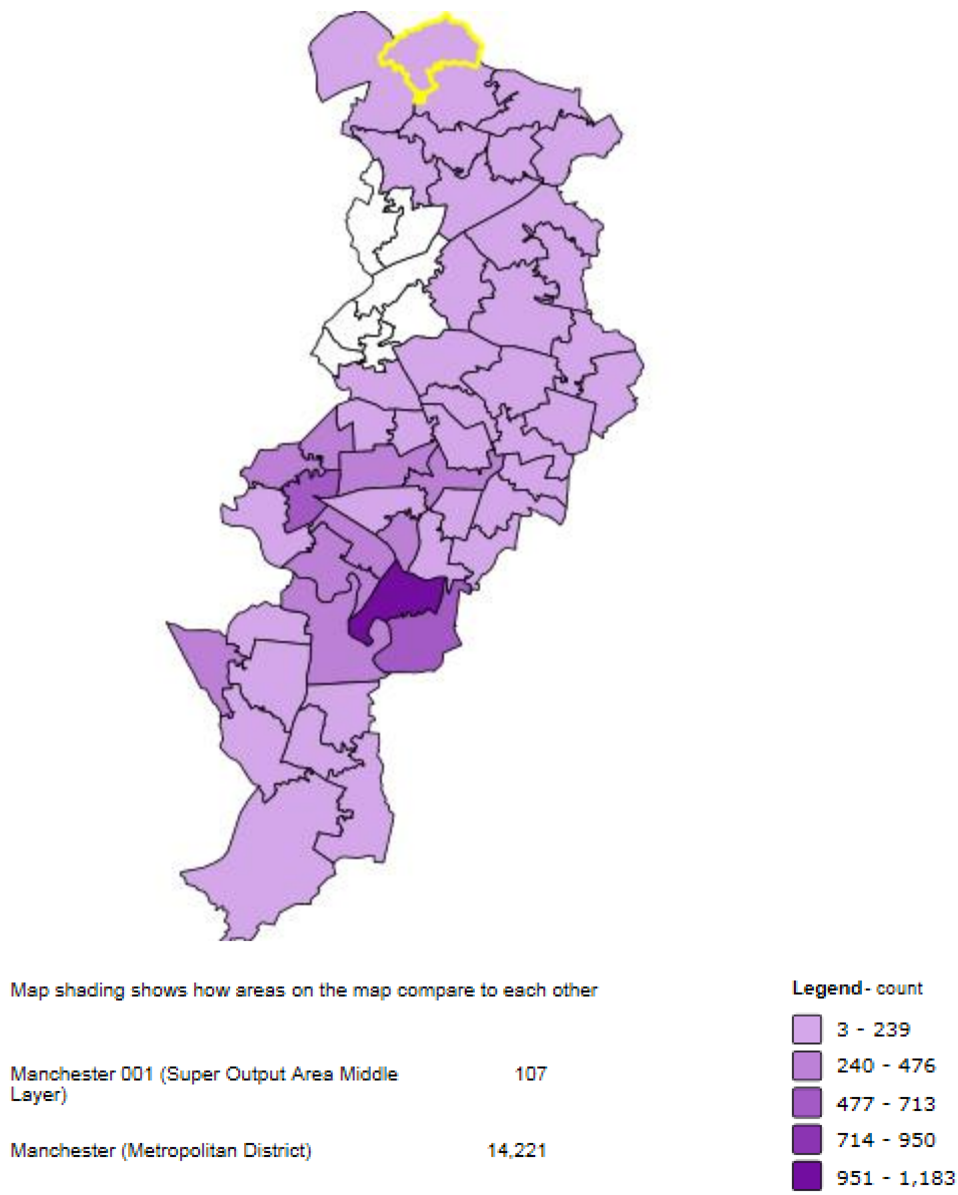

Figure 5.11: Map Showing Number of Dwelling Stock under Tax Band D; the legend count gives the number of dwelling stock

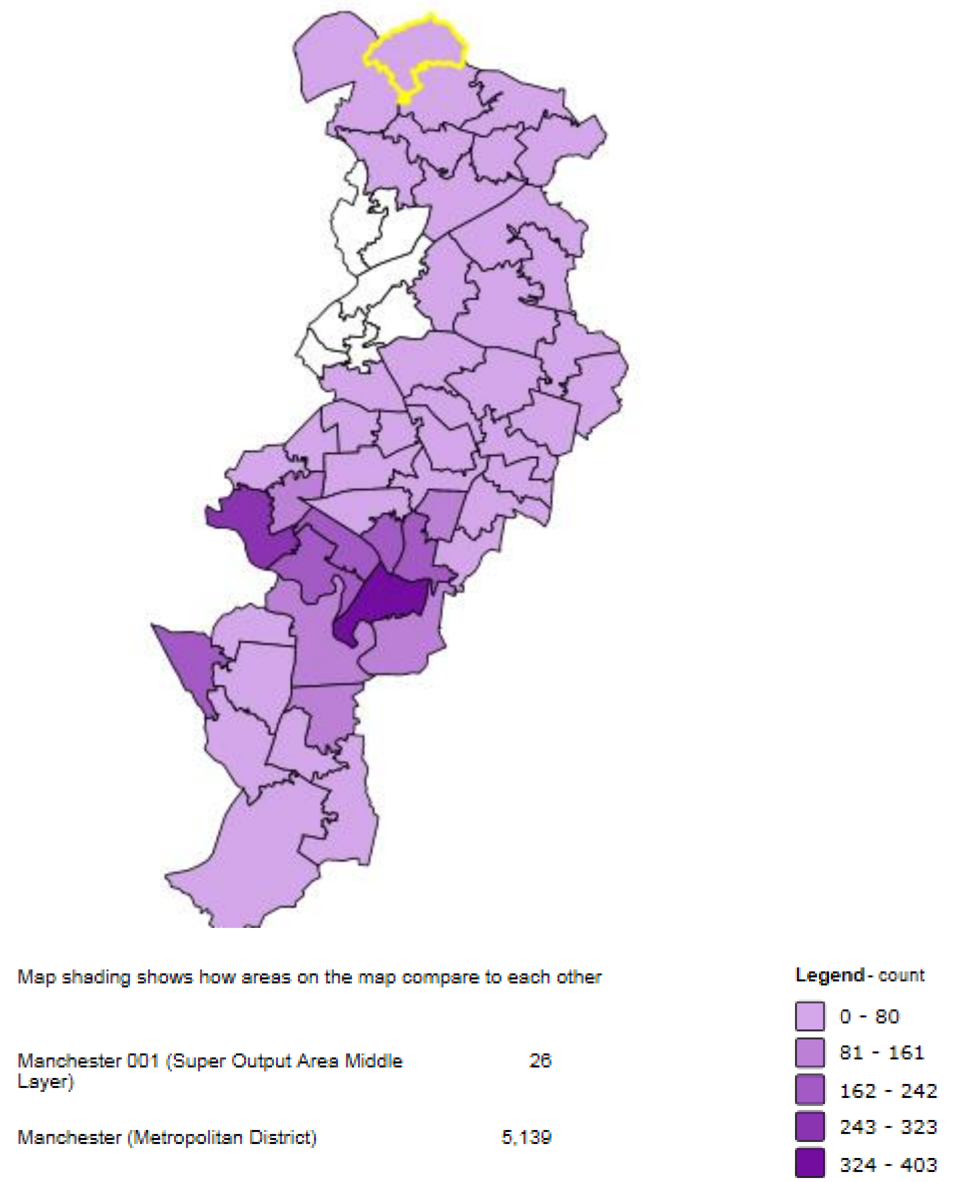

Figure 5.12: Map Showing Number of Dwelling Stock under Tax Band E; the legend count gives the number of dwelling stock

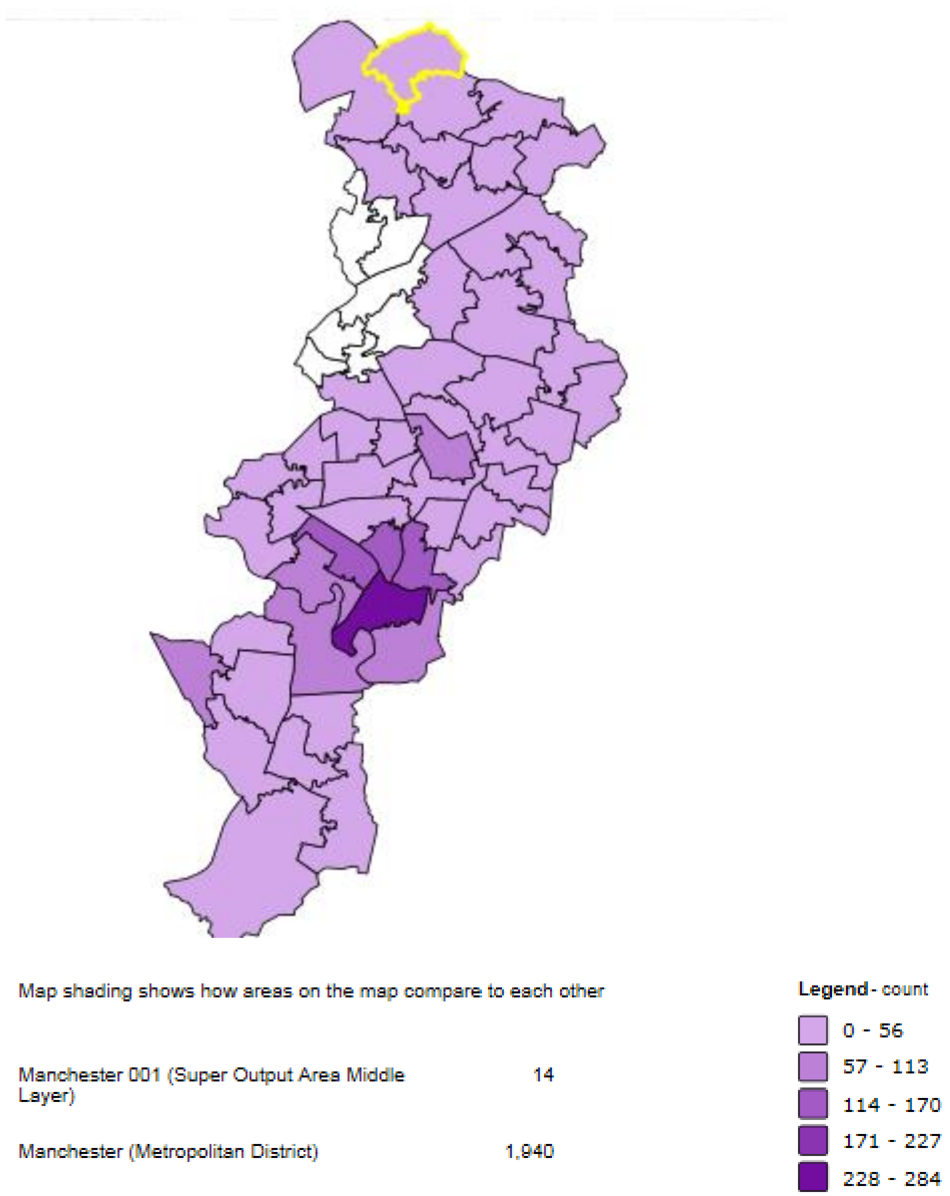

Figure 5.13: Map Showing Number of Dwelling Stock under Tax Band F; the legend count gives the number of dwelling stock

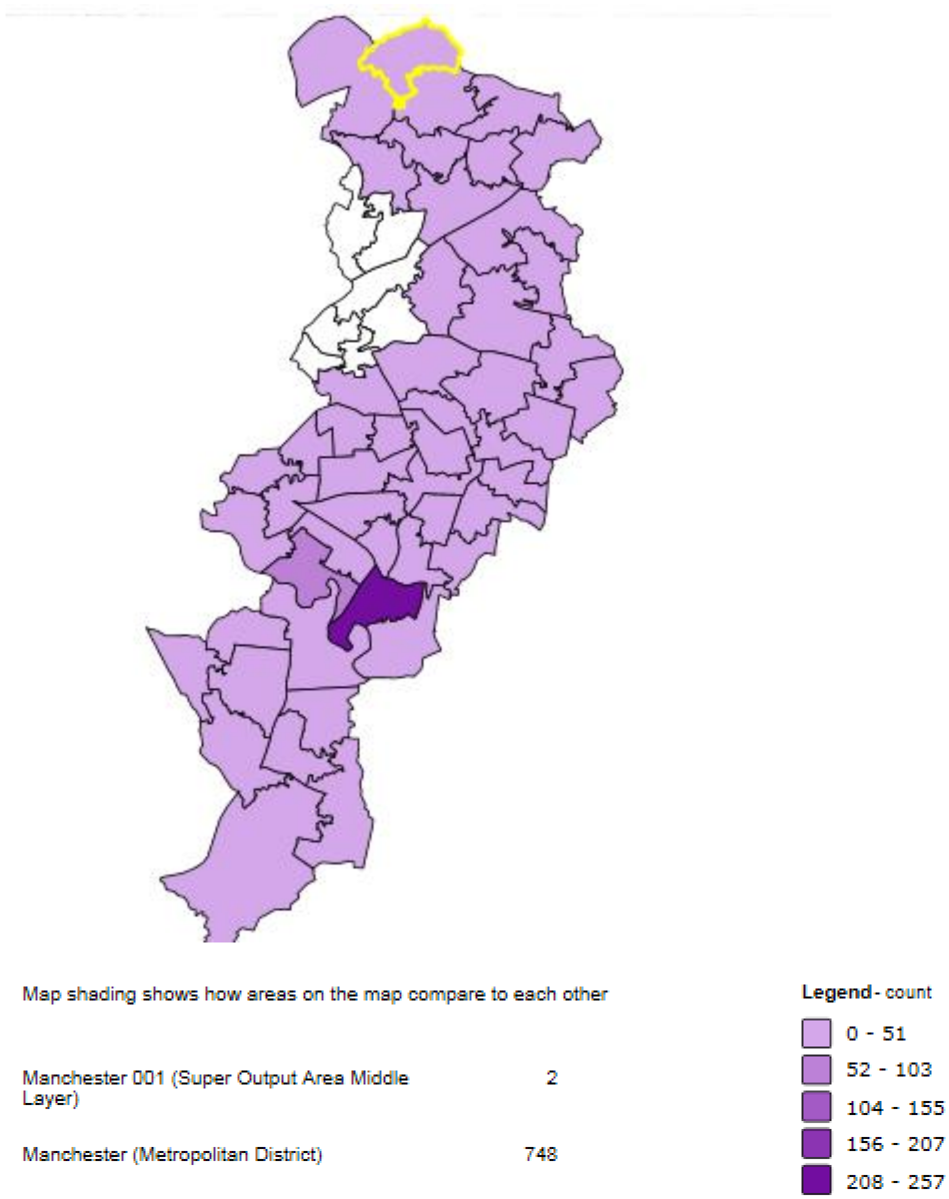

Figure 5.14: Map Showing Number of Dwelling Stock under Tax Band G; the legend count gives the number of dwelling stock

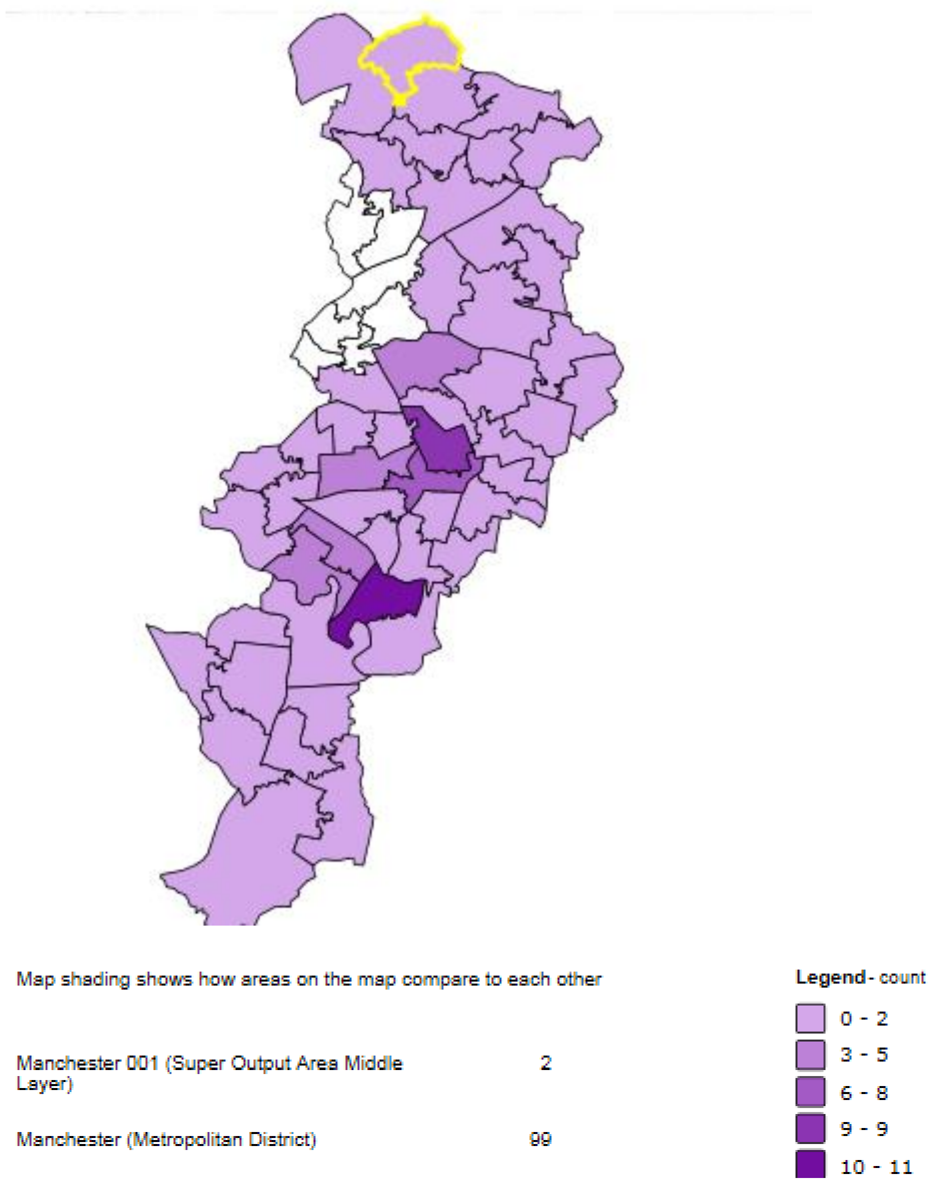

Figure 5.15: Map Showing Number of Dwelling Stock Under Tax Band H; the legend count gives the number of dwelling stock

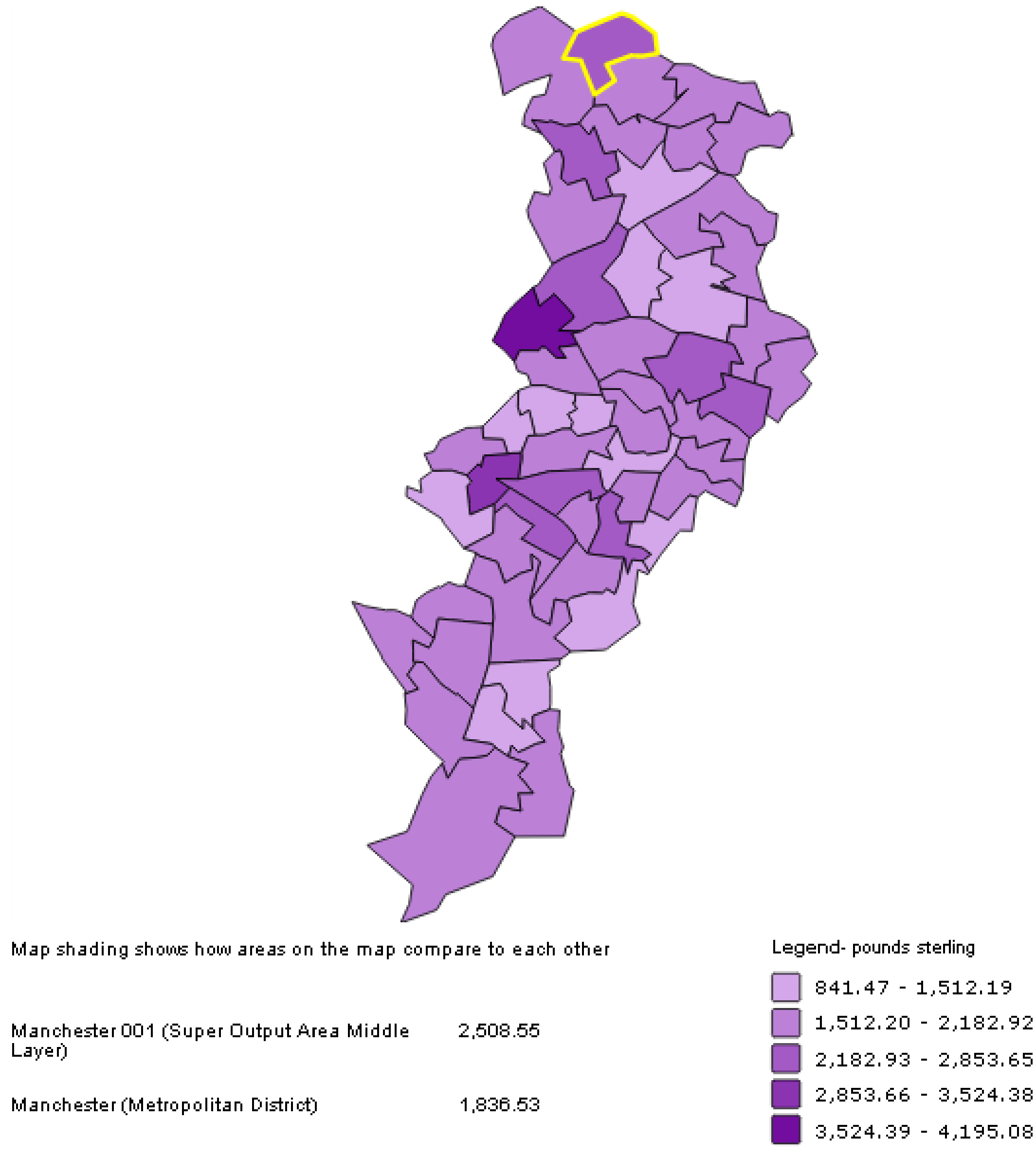

Figure 5.16: Map Showing Average Personal Consumer Debt; the legend count gives the value in Pounds Sterling

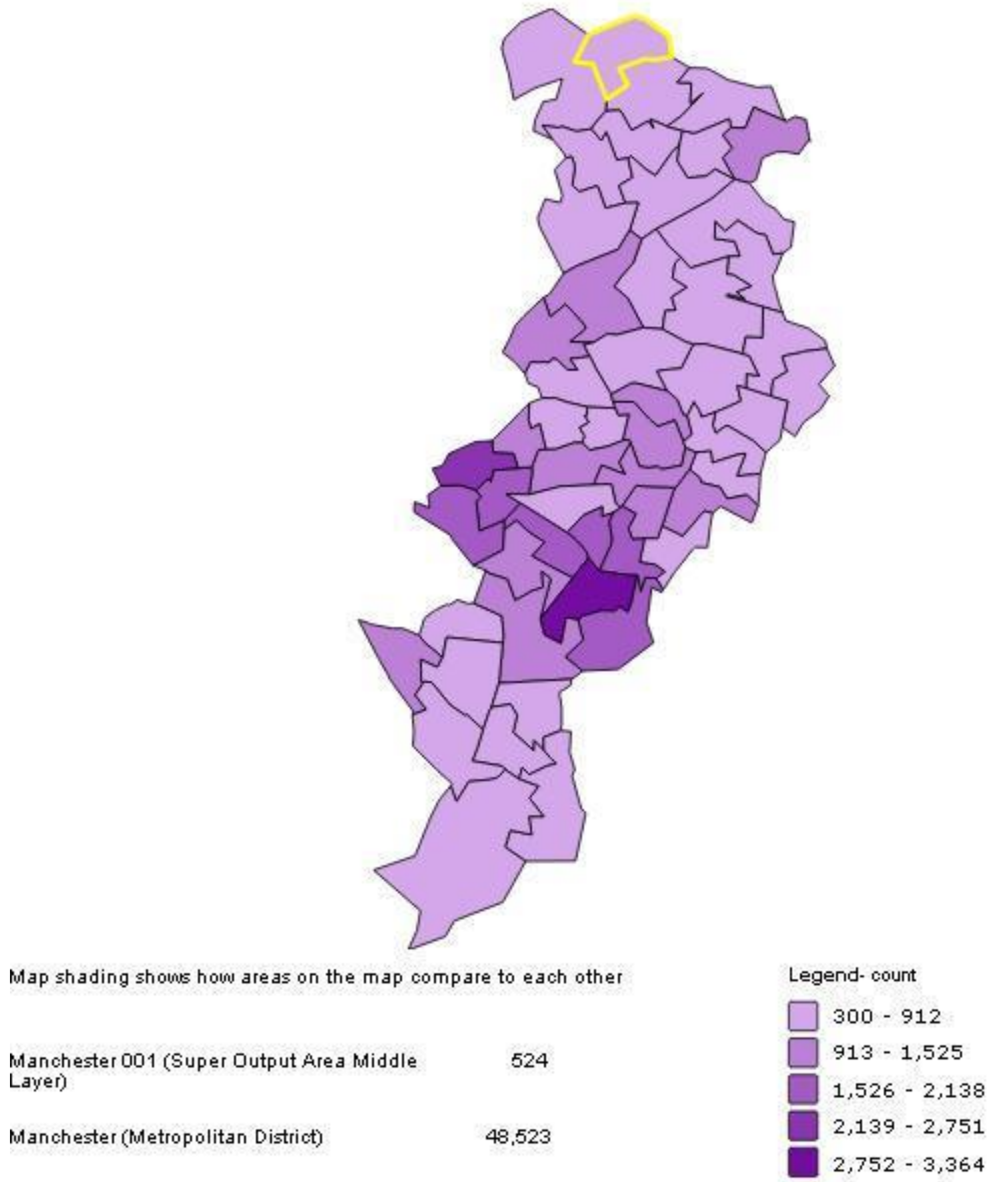

Figure 5.17: Map Showing Higher Professionals; the legend count gives the number of higher professionals

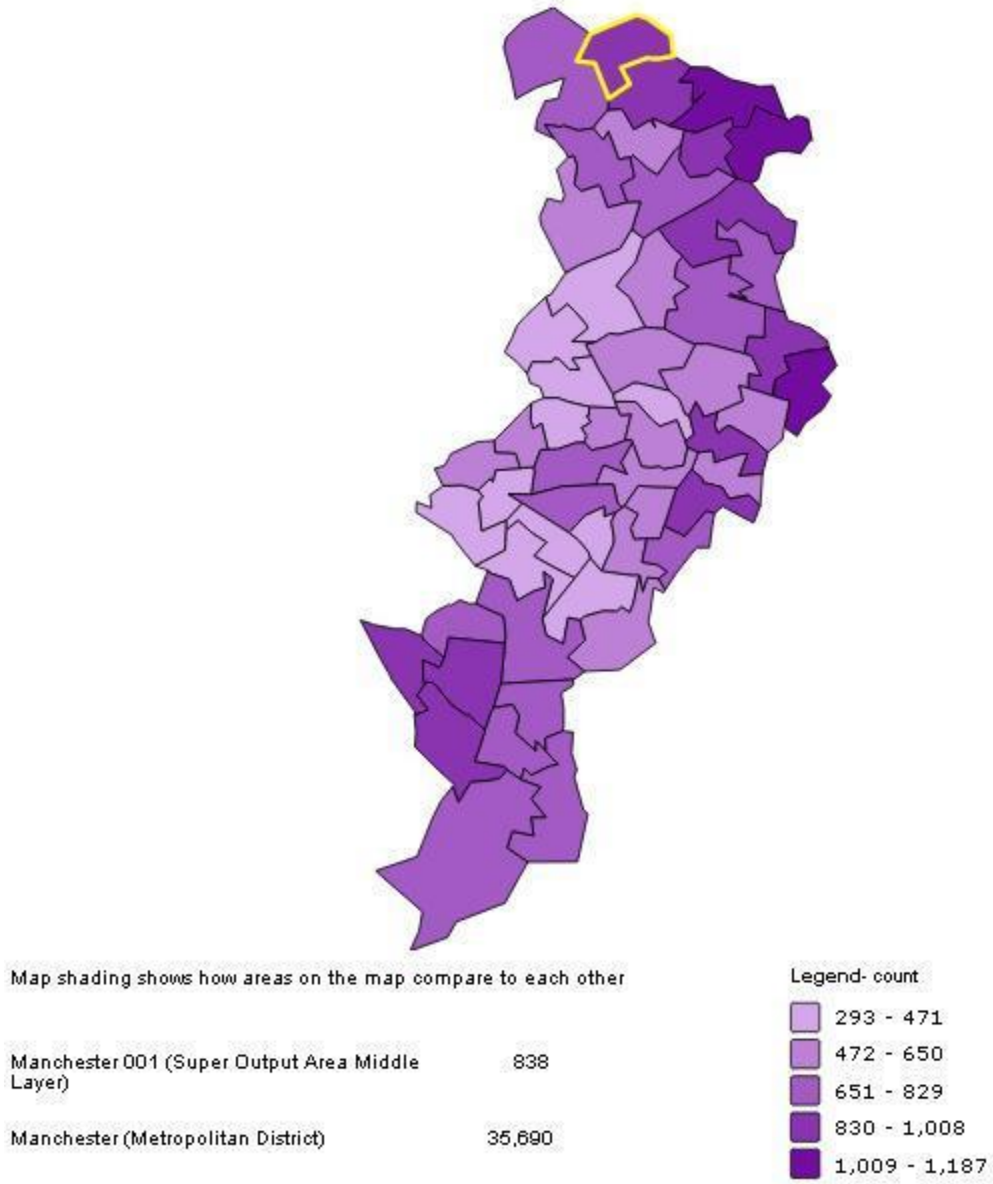

Figure 5.18: Map Showing Skilled Workers; the legend count gives the number of skilled workers

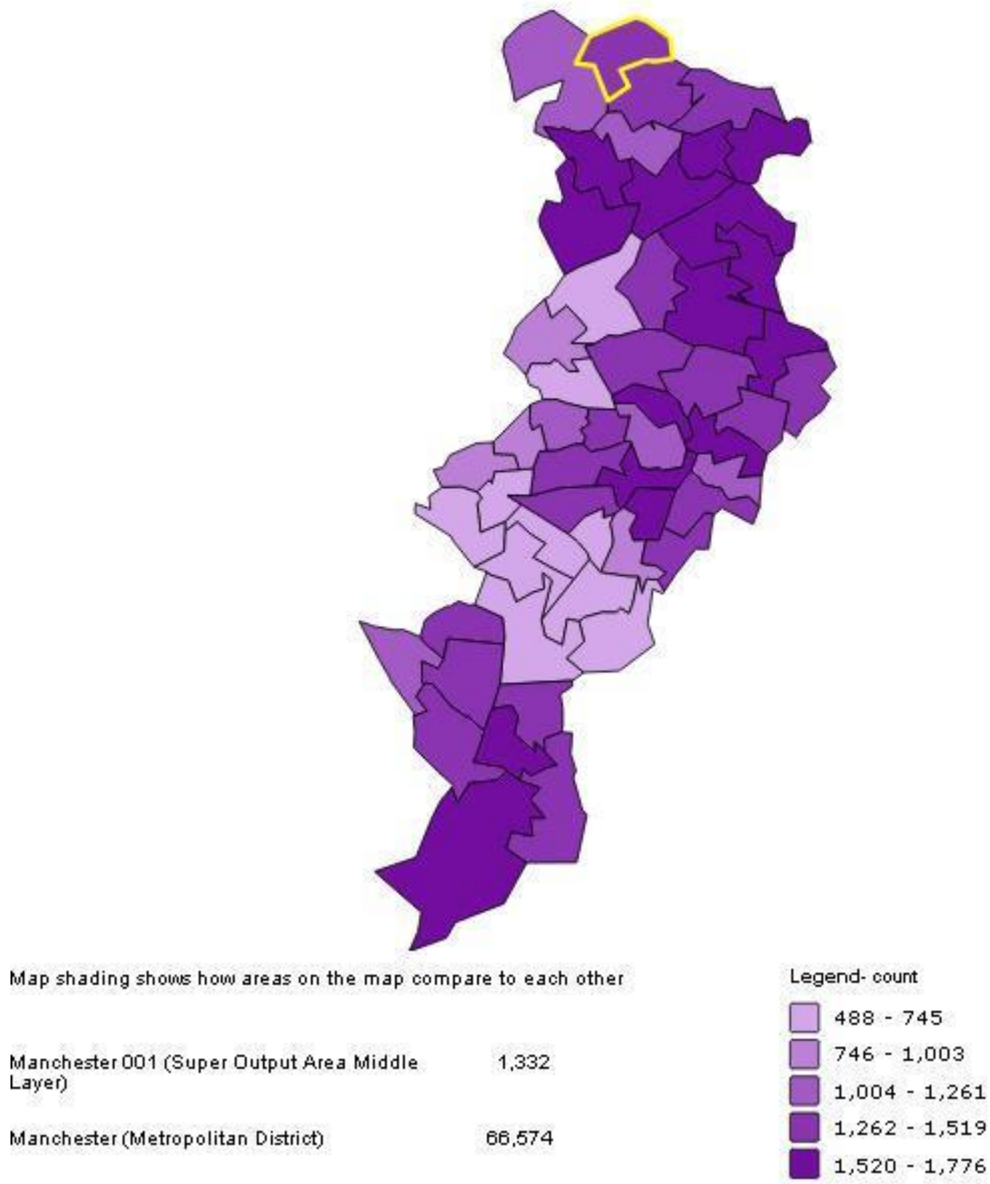

Figure 5.19: Map Showing Semi skilled and Unskilled Workers; the legend count gives the number of semi skilled and unskilled workers

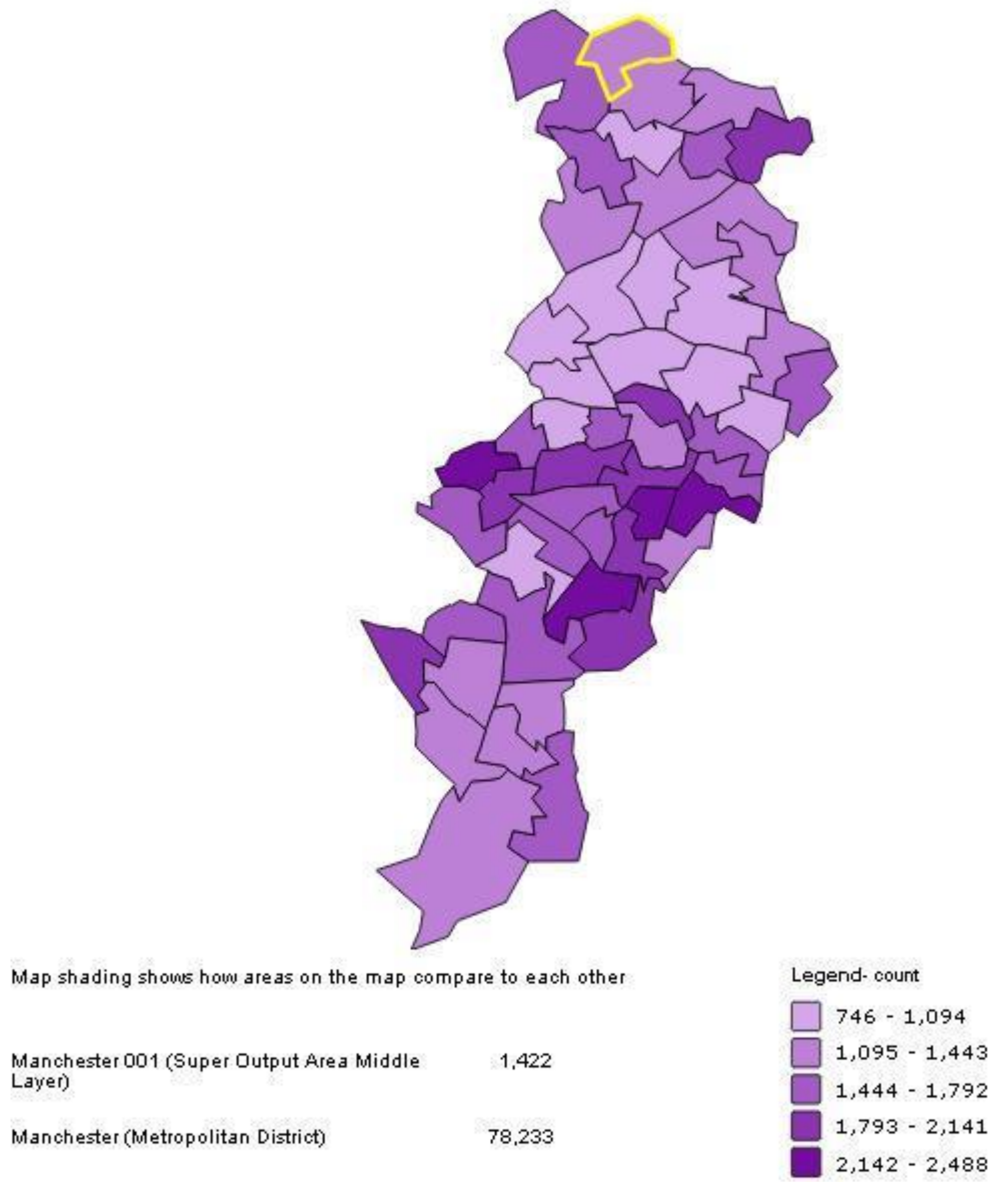

Figure 5.20: Map Showing Supervisory Professionals; the legend count gives the number of supervisory professionals

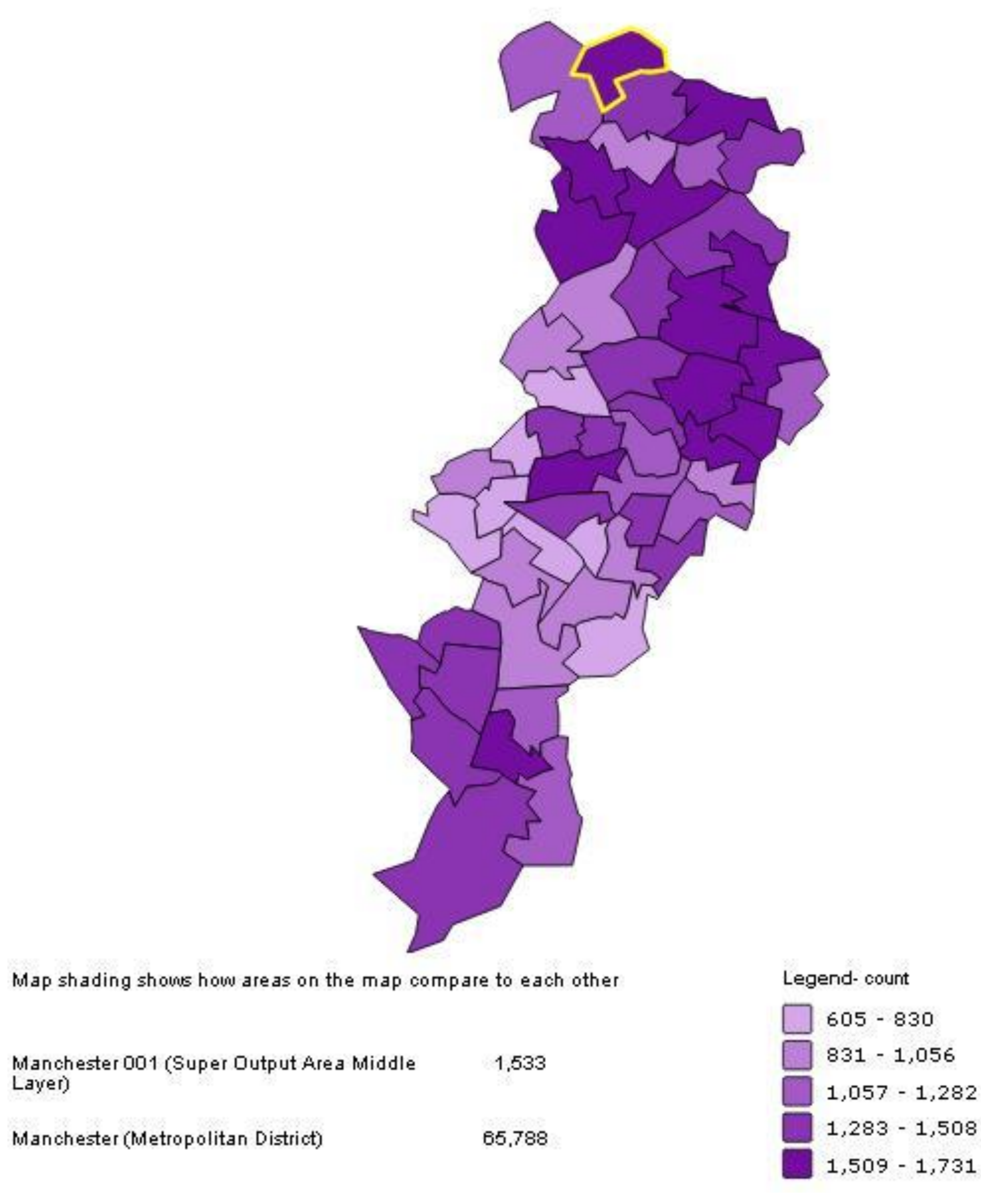

Figure 5.21: Map Showing People Living on State Benefit; the legend count shows number of such people

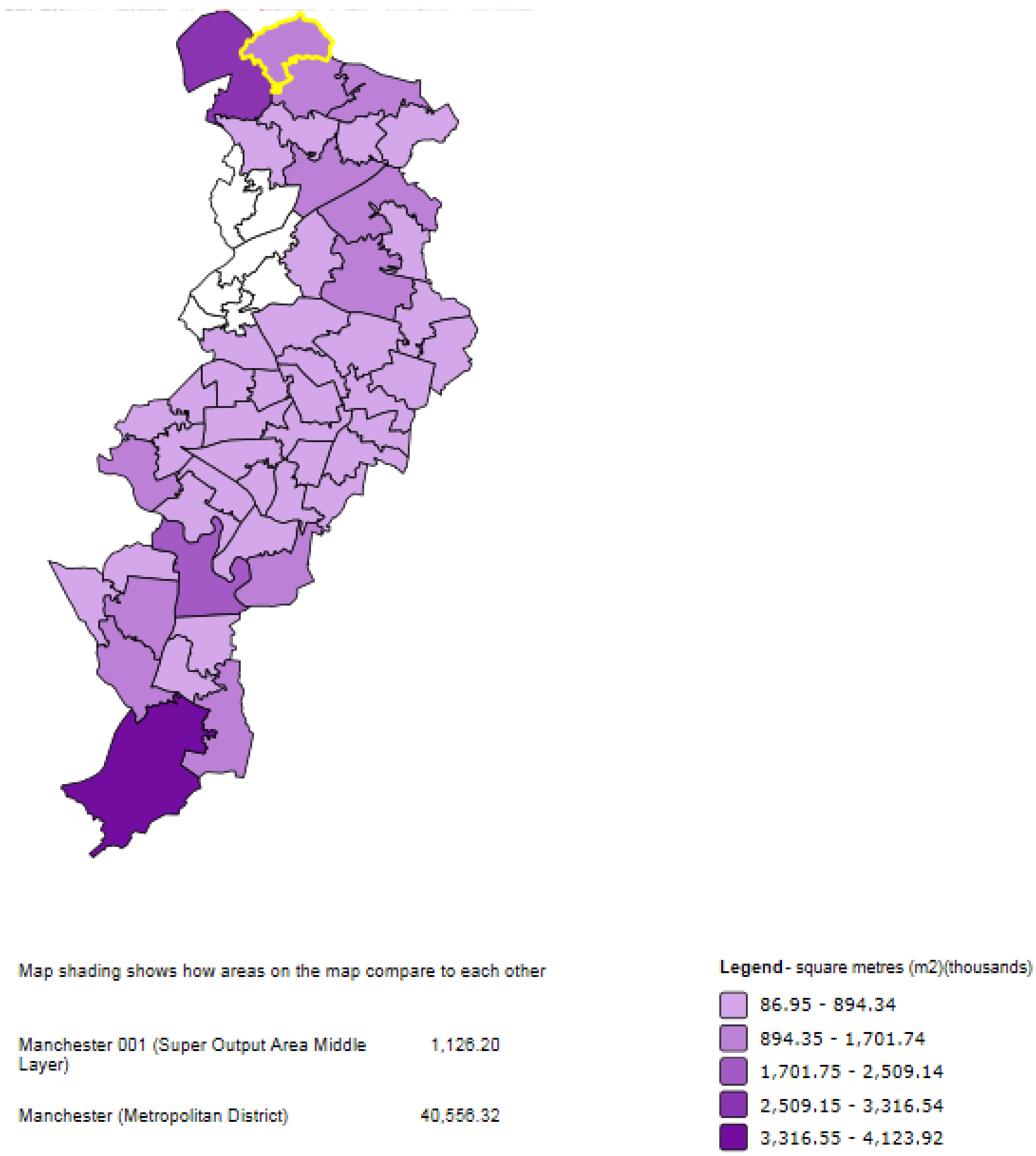

Figure 5.22: Map Showing Green Space; the legend count gives the estimate of such areas in thousand square meters

Table 5.2 contains an analysis of the six domains (since data on crime is not available) for the four areas. These domains provide individuals' wellbeing and behavior pattern trends as explained earlier.

Table 5.2: The neighborhood index domains comparison between M001, M002, M020, and M050

Domain	Sub Domain	M001	M002	M020	M050	Manchester	Conclusion
Economic Deprivation		1590 people claiming council tax benefit	1600 people claiming council tax benefit	1775 people council tax benefit	1820 people claiming council tax benefit	64,800 people council tax benefit	M001 and M002 are best in terms of less people claiming benefits.
Health Care	Good Health	4374 people are in good health	4700 people are in good health	4454 people are in good health	4380 people are in good health	253,665 people are in good health	M002 is best in terms of more people in good health.
	Fairly Good Health	1869 people are in fairly good health	1872 people are in fairly good health	1766 people are in fairly good health	2008 people are in fairly good health	90,039 people are in fairly good health	M050>M002>M001>M020 M002 is second to M050 in terms of people in fairly good health. M001 is third.
	Not So Good Health	1127 people are not so good health	1145 people are not so good health	1122 people are not so good health	1234 people are not so good health	49,115 people are not so good health	Inconclusive since numbers are very close to each other.
Housing	Dwelling Stock (total number of real estate units)	3836	4177	3705	3821	217,085	M002 has most stocks in line with its advance position in terms of economic deprivation and health care.
	House Price	80,000 GBP	85,000 GBP	85,000 GBP	84,000 GBP	1,25,000 GBP	M002 is demanding the highest average real estate price in line with its advance position in economics wellbeing and health care. Other factors must contribute to the house price in M020.
	Council Tax Band A Stock	3121 houses under Band A	2999 houses under Band A	3107 houses under Band A	3534 houses under Band A	1,29,861 houses under Band A	M002 has fewer houses in tax band A. There may be an opportunity for high end real estate developments in M002.
	Council Tax Band B	355 houses	635 houses	471 houses	217 houses	35,280 houses	M002 has more houses in band B. Same conclusion as

	Stock	under Band B	under Band B	under Band B	under Band B	under Band B	above, opportunities may exist in M002 for high end developments in tax band A.
	Council Tax Band C Stock	209 houses under Band C	423 houses under Band C	107 houses under Band C	62 houses under Band C	29,797 houses under Band C	M002 has more houses in band C. Same conclusion as above, opportunities may exist in M002 for high end developments in tax band A.
	Council Tax Band D Stock	107 houses under Band D	94 houses under Band D	15 houses under Band D	7 houses under Band D	14,221 houses under Band D	M002 is second to M001 as both have more houses in band D. Same conclusion as above, opportunities may exist in M002 for high end developments in tax band A.
	Council Tax Band E Stock	26 houses under Band E	19 houses under Band E	1 house under Band E	1 house under Band E	5139 houses under Band E	M002 is second to M001 as both they have more houses in band E. Same conclusion as above, opportunities may exist in M002 for high end developments in tax band A.
	Council Tax Band F Stock	14 houses under Band F	3 houses under Band F	1 house under Band F	0	1940 houses under Band F	M002 is second to M001 as both they have more houses in band F. Same conclusion as above, opportunities may exist in M002 for high end developments in tax band A.
	Council Tax Band G Stock	2 houses under Band G	4 houses under Band G	2 houses under Band G	0	748 houses under Band G	Numbers are too small to be conclusive.
	Council Tax Band H Stock	2 houses under Band H	0	1 house under Band H	0	99 houses under Band H	Numbers are too small to be conclusive.
Personal Consumer Debt		2508 GBP	1539.93 GBP	2630.71 GBP	1398.68 GBP	1836.53 GBP	M050>M002>M001>M020 Personal consumer debts are low in M050 followed by M002. However, the two are not far from each other especially in comparison to M001 and M020.
Social Grade	Higher Professional	524 people	655 people	462 people	342 people	48,523 people	M002 has the largest higher professional community. This supports earlier conclusion that the area may be suitable for high end developments that come under band A council

							tax.
	Skilled Workers	838 people	1422 people	1090 people	737 people	35,690 people	M002 has the largest higher skilled workers community. This supports earlier conclusion that the area may be suitable for high end developments that come under band A council tax.
	Semi Skilled and Unskilled Workers	1332 people	888 people	638 people	1652 people	66,574 people	M002 second best in having low semi-skilled workers.
	Supervisory Professional	1422 people	1494 people	1444 people	1098 people	78,233 people	With the exception of M050, it is difficult to distinguish between M001, M002 and M020.
	People Living on State Benefit	1533 people	1406 people	1528 people	1685 people	65,788 people	M002>M020>M001>M050 M002 has fewer people living on state benefit. This supports earlier conclusion that the area may be suitable for high end developments.
Environment (Green Space, Thousand Squared Meters)		1126.20	970.44	638.69	385.28	40,556.32	M001>M002>M020>M050 M002 is second best to M001 in terms of green space. However, M001 and M002 are not far away from each other especially in comparison with M020 and M050.

N.B.:Council Tax Band Details: Band A: Up to 40,000 GBP, Band B: 40,001 GBP-52,000 GBP, Band C: 52,001 GBP-68,000 GBP, Band D: 68,001 GBP – 88,000 GBP, Band E: 88,001 GBP- 120,00 GBP, Band F: 120,001 GBP-160,00 GBP, Band G: 160,001 GBP – 320,00 GBP and Band H: 320,001 GBP and above

The conclusions in Table 5.2 identify M002 and M020 as having the highest average real estate price of 85,000. M002 is identified as the best location from the four considered (M001, M002, M020 and M050). M002 has fewer claimants for council tax benefits, fewer houses in high council tax band A, more people in good health, a more professional and skilled work force, fewer people living on state benefits, higher average house price and more green space than M020 and M050. M002 looks like an attractive place for real estate investment since it provides a better environment for individuals’ well being. M020 is the second best area in terms of green space and the lowest in personal consumer debts, and is second to M002 in terms of having the fewest people on benefits. The conclusions establish a pattern to be tested that the real estate prices are positively correlated with higher income, fewer council tax claimants, good health and green space.

5.2 Model evaluation and diagnostic tests for the residuals

Section 5.2 has hypothesized positive causal relationships between real estate price changes on the one hand and higher income, fewer council tax claimants, good health and green space on the other. Data sources on employment, household incomes, inflation and council tax for the period from 1998 to 2013 were explained in chapter four, section 4.1. The following regression is used:

$$y_t=\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + u_t \tag{5.1}$$

Where,

y_t= change in price,

α= constant,

$\beta_1, \beta_2, \beta_3, \beta_4$= parameter estimates,

X_1= change in employment,

X_2= change in income,

X_3= inflation,

X_4= change in council tax,

u_t= residual term with normal distribution.

The cross correlations of the explanatory variables are presented in Table 5.3. The cross correlations are examined to avoid the problem of multicollinearity. It is generally accepted that as long as the cross correlation between two variables is less than 0.7, there is no collineraity problem and both variables can be included as explanatory variables.

Table 5.3: The Cross-Correlations between Change in Employment, Change in Income, Inflation and Change in Council Tax

	Change in employment	Change in Income	Inflation	Change in Council Tax
Change in employment	1			
Change in Income	0.1958	1		
Inflation	-0.1225	-0.3779	1	
Change in Council Tax	0.0417	0.5183	-0.3313	1

None of the correlations are larger than 0.70, which indicates lack of collinearity. The highest positive correlation is between change in income and change in council tax. The regression is run for the all the property types and the results are shown in *Table 5.4*.

Property Types	**Co-efficient and T-Test for Change in Employment**	**Co-efficient and T-Test for Change in Income**	**Co-efficient and T-Test for Change in Inflation**	**Co-efficient and T-Test for Change in Council Tax**	**Co-efficient and T-Test for Constant**	**Adjusted R-Squarred**	**F-test**
Semi-Detached	0.368 (0.25)	2.676 (0.23)	1.354 (0.86)	0.043 (0.03)	-0.046 (-0.63)	.0159	0.217
Detached	0.406 (0.14)	3.96 (1.64)	-2.393 (-0.76)	-1.979 (-0.72)	0.073 (0.5)	0.042	0.378
Flats	-1.469 (-0.87)	3.421** (2.48)	0.312 (0.17)	-1.118 (-0.71)	-0.007 (-0.08)	0.160**	0.216
Terraced	-1.657 (-0.76)	2.833 (1.59)	2.740 (1.17)	0.980 (0.48)	-0.082 (-0.76)	0.049	0.365

***, **, * indicates significance at the 1%, 5%, and 10%, respectively.

The Table indicates non significant relationships between price changes for each property type and the explanatory variables except for the real price changes for Flats and changes in income.

Table 5.5 shows the regression of Flats price changes as dependent variable on the changes in income as the explanatory variable.

Table 5.5: Regression Analysis between Flats and Change in Income

Source	**SS**	**Df**	**MS**	**No.Of Obs.**	16
Model	0.04223025	1	0.04223025	**F(1,14)**	6.34
Residual	0.093261627	14	0.006661545	**Prob>F**	0.0246
Total	0.135491877	15	0.009032792	**R-Squarred**	0.3117
				Adj R-Squarred	0.2625
				Root MSE	0.8162
Flats	**Coef.**	**Std.Err.**	**T**	**P> I t I**	
Change in Income	2.65174**	1.05319	2.52		0.025
Constant	-0.0141657	0.0384437	-0.37		0.718

***, **, * indicates significance at the 1%, 5%, and 10%, respectively.

Flat price changes are significantly positively correlated with change in income. A positive rise in income leads to a positive rise in the prices of Flats and vice versa.

5.3 Summary

This Chapter handled research problem three, lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting. The Chapter examined four areas in Manchester with the objective of identifying causal relationships between real estate price and socio economic parameters. One of the two areas that demanded the largest real estate average prices is found to be associated with fewer claimants of council tax benefits, fewer houses in high council tax band A, more people in good health, a more professional and skilled work force, fewer residents living on state benefits, and more green space that is consistent with individual well-being attributes. The other area demanding the highest average price is a lower number of people living on state benefits, which came second only to the best area in terms of green space. The Chapter discussed well-being and consumer behaviour attributes, and also examined the causal relationships identified in the case study for data for the entire Manchester City. Annual data were collected for employment, income, inflation, and council tax during the period from

1998 to 2013. Multiple regressions were estimated for price changes for each of the four property types as dependent variables, and change in employment, change in income, inflation, and change in council tax as independent variables. The only significant relationship is observed between price changes in the market for Flats and changes in income. The problem is that data is only available on an annual basis and for a relatively short time period. The sample size is small and changes in the underlying causal relationships can be captured less frequently and because of that out of sampling forecasting is not feasible. However, studying the four areas in Manchester provided us with a clear insight that causal models are important in determining real estate prices in case if data is available. It will be better than time series models since it can explain the causal effects, for example how a change in an independent variable can lead to changes in the dependent.

Chapter 6

Conclusions, Contributions, Limitations and Recommendations

6.0 Conclusions

The thesis identified a gap in the literature on real estate price forecasting that led to specific three research problems; (1) lack of studies of real estate price determination by residential property types within a city, (2) lack of studies of seasonality of real estate prices by property types within a city, and (3) lack of studies utilizing the potential of Geographic Information Systems in real estate forecasting. The three research problems led to the identification of the aim of the thesis as to provide an approach to real estate residential modeling and forecasting price and seasonality covering a property type's within a region or a city. The thesis covers residential real estate markets of different types with regard to time varying parameters, causal relationships and correlations between property types.

The content of the research problem and the aim of this study have been broken down into the three objectives of the research, which are; (1) investigate the time series properties of different property types (Flats, Terraced, Semi-Detached, and Detached) and compare their time series characteristics with each other as well as the UK all house price index, (2) introduce a methodology that combines published data from the UK office for national statistics with the geographic maps (GIS) to identify causal relationships that can be used in forecasting real estate prices, and (3) to provide a new approach utilizing time series techniques, causal models and GIS tool to model and forecast real estate price changes.

6.1 Contributions

The thesis main contribution on the theoretical side is that it is the first to combine three different tools into an integrated approach to model and forecast the behavior of real estate prices over time. The three tools are time series, cross sectional casual models, and Geographic information systems (GIS). The implications for the research community are that the approach is flexible to cope with different scenarios based on data availability. The practical contributions of the thesis are numerous. The thesis is the first to examine residential properties within a region or a city in addition to the UK All house price index. It is much more accurate to forecast by property types than for aggregate index. Property prices responses to market changes are different according to property type. The accuracy for forecasting is more for the less expensive property type (terraced and flats).

The results show that Manchester's most expensive housing type (Detached properties) experienced more negative price declines than the less expensive (Semi-Detached) properties and the least expensive such as Terraced properties and Flats. This means the price decline for Detached properties took a year to show positive price changes while for Flats and Terraced properties it took only a quarter to show positive price changes. The autocorrelations for the Semi-Detached properties which is closer to the most expensive (Detached) properties than the least expensive (Flats) showed a mixed pattern with the correlations with quarters one and two being negative while the correlations with quarters four and five being positive. The Semi-Detached properties seem to recover faster than the Detached but slower than the Terraced properties and Flats. The Detached and Semi-Detached past history explains more of the variation in the current prices than for Flats and Terraced properties as indicated by the adjusted R-squared. The reason for this could be that Flats and Terraced properties are much more affordable and therefore variation in their prices could be more related to causal variables such as mortgage rates and credit conditions of the market. The Detached and Semi-Detached properties are much more expensive which may mean they have less sensitivity to mortgage and credit markets conditions, and therefore more dependence on past history than Flats and Terraced properties.

The thesis is also first to concentrate on real estate property type seasonality and volatility in addition to modeling its time series attributes within a region or a city. Real estate prices for the Semi-Detached, Detached, Terraced properties and Flats indices for Manchester are more volatile than the UK All index. The main reason for this is that the UK All price index is smoother since it has far more coverage than Manchester. The behavior of the process of Terraced price change seems to be constant over time with the exception of the sharp rise in prices in 2000 quarter four when prices went up by more than fifty percent. With regard to the Flats price change, the price change can be categorized into three scales: low volatility, medium volatility and high volatility. The low volatility period ranges from 2002 Q2 to 2007 Q2; the medium volatility from 1995 Q2 to 2001 Q2; and the high volatility from 2008 Q1 to 2011 Q1. Detached property price changes can be categorized into three periods: low volatility, medium volatility and high volatility. The low volatility ranges from 1995 Q2 to 2001 Q2; medium volatility from 2002 Q2 to 2007 Q2 and high volatility from 2008 Q1 to 2011 Q1. The price change for Semi-Detached properties from the period 1995 Q2 to 2011 Q1 can be categorized into three scales: low volatility, medium volatility and high volatility. The low volatility ranges from 1995 Q2 to 2001 Q2; medium volatility from 2002 Q2 to 2007 Q2 and the high volatility from 2008 Q1 to 2011 Q1. With regard to seasonality, no seasonal patterns were found for any real estate price series once we have fitted the appropriate time series model.

The thesis is also the first to combine Geographic Information Systems (GIS) attributes with socio-economic factors to help in understanding causal relationships that can be used to forecast real estate prices. Four areas in Manchester were examined with the objective of identifying causal relationships between real estate price and socio economic parameters using GIS. We examined the causal relationships identified in the case study for data for the entire Manchester City. Annual data were collected for employment, income, inflation, and council tax during the period between 1998-2013. Multiple regressions were estimated for price changes for each of the four property types as a dependent variable, and change in employment, change in income, inflation, and change in council tax as independent variables. The only significant relationship is observed between Flat price changes and changes in income as Flats are more affordable and therefore more income sensitive than other property types.

The novelty of the thesis is, it is the first that pays attention to within city real estate price forecasts by property types. Previous studies concentrated on a macro decision making level.

Macro decisions are relevant to policy makers and real estate fund managers for predicting trends. However, no study has looked at the micro level, i.e. studying time series performance within a city and the correlation and performance by property type. Such a study can provide critical information to policy makers, developers, real estate fund managers and private investors when it comes to micro level management. This thesis attempts to fill the gap that exists in the field, as no previous studies have been found that cover modeling and forecasting real estate prices on a micro level within a city. Also, there has not been a study that examined real estate prices within a city and by properly types using the proposed methodology of Brooks and Tsolacos, 1999 for modeling seasonality. The thesis also tries to incorporate recent developments in geographic information systems (GIS) to identify multivariate causal models that can help in real estate forecasting within a city by property types. Previous research such as that of Breedon and Joyce, 1993 has analyzed the relationship between real estate prices, earnings, disposable income, demographic factors, and the rate of repossessions by lenders. There is no study that uses geographic maps to identify causal relationships that can be used in forecasting real estate prices. The thesis maps causal relationships between real estate prices and explanatory variables like higher income, council tax claimants, good health and green space. The thesis is the first to provide a framework that ties social and economic variables to the price of real estate property within a city.

The thesis managerial relevance is in providing a new approach for the issue of choosing real estate locations that can help in achieving higher economic value added. The thesis implements GIS to map causal factors that lead to higher real estate prices. Those causal factors were later used in a multiple regression format that can help in forecasting future real estate price changes. The thesis examined four areas in Manchester with the objective of identifying causal relationships between real estate price and socio economic parameters. One of the two areas that demanded the largest real estate average prices is found to be associated with fewer claimants of council tax benefits, fewer houses in high council tax band A, more people in good health, a more professional and skilled work force, fewer residents living on state benefits, and more green space that is consistent with individual well-being attributes. The other area demanding the highest average price is a lower number of people living on state benefits, which came second only to the best area in terms of green space. Annual data were collected for employment, income, inflation, and council tax during the period from 1998 to 2013. Multiple regressions were estimated for price changes for each of the four property types as dependent variables, and change in employment, change in income, inflation, and change in council tax as independent variables. There is a significant relationship between real estate price changes as dependent variable and changes in income in the market for Flats. Studying the four areas in Manchester provided us with a clear insight that causal models are important in determining real estate prices in case if data is available. It will be better than time series models since it can explain the causal effects, for example how a change in an independent variable can lead to changes in the dependent. Time series may help in forecasting but does not provide much insight about the causal relationships that explain those changes.

6.2 Limitations

The limitations of the thesis are as follows:

6.2.1 Crime related data has not been considered because of its sensitivity. This data has to be considered while making a decision on site selection.

6.2.2 The models developed in this thesis assume stationary time series. While we allow for trends and cycles and we try to consider their effects in shaping our decisions, we assume that there have been no structural changes in the time period of the study. Structural changes are defined to be outside shocks to the model which will lead to a shift up or down in the mean and/or variance of the time series data. Relaxing this assumption is difficult since real estate data frequency is not as high as in the stock market. Stock market prices can be observed on minute by minute, day by day etc. However, data for real estate is available quarterly. There are more chances for structural shifts in case of minute by minute or day by day data in comparison with quarterly data which tend to be smoother. The time series analysis covers a period from 1995 (quarter one) to 2011(quarter one) including the credit crunch and the subsequent mortgage problems from quarter four 2008. Unfortunately, due to limited quarterly data it is not possible to divide the data into two sub-samples, pre and post-crash since time series techniques require at least fifty observations for estimation while we only have sixty-four. The residual analysis of all models indicate adequacy with no identified patterns in ACFs and PACFs of the residuals. Dividing the sample pre and post credit crunch would have led to possible positive correlation pre and negative correlations post but with no impact on the stability of our models which have passed all comprehensive diagnostic tests.

6.3 Recommendations

The recommendations are directly related to the limitations stated in section 6.2.

6.3.1 Obtaining data on crime rate and including it in the analysis for price/location relationship. This will help in building a more accurate multiple regression models for forecasting real estate prices.

6.3.2 Expanding the data to include more observations beyond 2011. This would allow for a structural analysis in the real estate price determination pre and post credit crunch.

6.3.3 The suggested time series models have been proved successful in tracking the random changes in quarterly real estate prices in Manchester. One possible area for future research is to allow for multivariate time series models that would take correlations between different property (Semi-Detached, Detached, Flats, Terraced) types into account. For example, instead of modeling each property type individually, a multivariate model can model the price behavior for the four property types simultaneously allowing for the correlations between property types to be used in the modeling process.

6.3.4 The socio-economic preferences issues discussed in thesis can be further modified by designing and targeting clients within Manchester (UK). Future researchers can work on survey design and analysis in order to match clients' preferences with spatial and socio-economic factors of locations.

6.3.5 The GIS analysis has depended on the choice of four randomly locations within Manchester. Future research can target locations with high/low prices and model their price behavior as a function of different socio-economic factors. The objective is to establish explanatory models that can be used to forecast future price changes given future expectations of relevant socio-economic factors.

References

- Abraham, A., Sommerhalder, K. and Abel, T. (2010) "Landscape and well-being: a scoping study on the health-promoting impact of outdoor environments", *International Journal of Public Health* 55: 59-69.
- Addae-Dapaah, Kwame, Gek, W.S. and Ebrahim, M.S. (2002) "Real Estate Portfolio Diversification by Sources of Return", *Journal of Real Estate Portfolio Management*, Vol. 8, No. 1, 2002; pp. 1-15.
- Akintoye, A. S. and Macleod, M. J. (1997) "Risk Analysis and Management in Construction", *International Journal of Project Management*, Vol. 15, No. 1: 31-38.
- Alba, Joseph W. and Howard Marmorstein (1987), "The Effects of Frequency Knowledge on Consumer Decision Making", *Journal of Consumer Research*, 14: 14-25.
- Anuar Alias and SoiTho C.Y (2011), "Performance Analysis of REITS: Comparison Between M-REIT and UK-REIT", *Journal of Surveying, Construction and Property* Vol.2 Special Issue 2011.
- Alexander, S. (2000), "Economic Implications of Using a Mean –VaR Model for Portfolio Selection; A comparison with Mean –Variance Analysis", Journal of the American Statistical Association.
- Allen, Franklin and Gale, Douglas. (2007), "Systemic Risk and Regulation"; *National Bureau of Economic Research.*
- Alpert and Mark I. (1971), "Identification of Determinant Attributes", *Journal of Marketing Research*, 8:184-191.
- Amy, K., Ming, Y. and Yuan, L. (2000), "The natural vacancy rate of the Singapore office Market", *Journal of Property Research* 17 (4), 329-38.
- Applebaum, W. (1965), "Can Store Location Research be a Science?", *Economic Geography* 41(3): 234-237.

- Applebaum, W. (1966), "Methods for Determining Store Trade Areas, Market Penetration and Potential Sales", *Journal of Marketing Research*, 3: 127-141.
- Anderson R. I. and Springer, T. M. (2005). "Investor Perception of Retail Property Risk: Evidence from REIT Portfolios", *Journal of Shopping Center Research*, 12: 104-120.
- Armstrong, Martin A. (2009), "A Forecast For Real Estate", *Journal of Real Estate Finance Economics*.
- Bailey, M. J., Muth, R.F., Nourse, H.O. (1963), "A regression method for real estate price index construction", *Journal of the American Statistical Association*, 58 933-942.
- Barker, K. (2004) '*Review of Housing Supply: Final Report – Recommendations*,' London: HMSO.
- Barker, K. (2006) '*Barker Review of Land Use Planning; Final Report – Recommendations*,' London: HMSO.
- Barthélémy, Fabrice and Prigent, Jean-Luc (2009); "Optimal Time to Sell in Real Estate Portfolio Management", *Journal of Real Estate Finance and Economics*, 38:59–87.
- Barras, R. (1983), "A simple theoretical model of the office-development cycle", *Environment and Planning*, Vol. 15 No. 6, pp. 1381-94.
- Baum, A (2002), "European private property market", *International Direct Real Estate Investment*.
- Baum and Key, T (1999), "Systems for Attributing Real Estate Returns", *European Real Estate Society Conference*.
- Birch, John W., and Sunderman, Mark A. (2003), "Estimating Price Paths for Residential Real Estate", *Journal of Real Estate Research*, 25(3), 277-300.
- Berndt, E., B. Hall, R. Hall, and J. Hausman (1974), "Estimation and inference in nonlinear structural models", *Annals of Economic and Social Measurement.* 3:653-665.
- Beth Anne Grib (2011), "The Importance of Demographics in Real Estate".
- Blume, Marshall E. and Friend, Irwin E. (1974). "Risk, Investment Strategy and the Long-Run Rates of Return" *The Review of Economics and Statistics* Vol. 56, No. 3, pp. 259-269.
- Booth, D. G. and Fama E. F. (1992). "Diversification Returns and Asset Contributions", *Financial Analysts Journal*, 48, 26-32.

- Box, George and Jenkins, Gwilym (1970), “Time series analysis: Forecasting and control”.
- Brockwell P. J. and Davis R. A. (1986), “Time Series: Theory and Methods”, *Springer Series in Statistics*.
- Breedon, F. J. and Joyce, M. A. S. (1993), "House prices, arrears and possessions: a three equation model for the UK", Bank of England Working Paper Series No 14.
- Brereton, F., Clinch, J. P. and Ferreira, S. (2008), “Happiness, geography and the environment”, *Ecological Economics, 65(2),* 386–396.
- Brown, G.R. (1997), “Reducing the dispersion of returns in U.K. real estate portfolios”, *Journal of Real Estate Portfolio Management*.
- Brown, Roger J. (2004), “Risk & Private Real Estate Investments”, *Journal of Real Estate Portfolio Management*, Vol. 10, No. 2, pg. 113 (16).
- Brooks, C., and Tsolacos, S. (1999), “The impact of economic and financial factors on UK property performance”, *Journal of Property Research* 16(2), 139-52.
- Brooks, C., and Tsolacos, S. (2010), “Real Estate Modelling and Forecasting”.
- Bruce, Brian R. (1991), “Real Estate Portfolio Management”, *Probus Publishing Co. Chicago IL*.
- Bryan, T. B. and P. F. Colwell (1982), “Housing Price Indexes, In C. F. Sirmans (Ed.)”, *Research in Real Estate, A Research Annual*, Greenwich, CT: JAI Press.
- Buckley, Michael P. (1994), “Portfolio Disposition Strategies: The Institutional Decision of the Decade”, *CRE, Real Estate Issues*.
- Burrough, P. A. (1986), “Principles of Geographic Information Systems for Land Resources Assessment”, *Clarendon Press, Oxford*.
- Burrough, P. A. (1991), “Intelligent geographical information systems”, *Increased Sustainable Agricultural Productivity in Africa Through the Use of Intelligent Geographical Information Systems*, 14-18.
- Byrne, P and S. Lee (1995), “Is There a Place for Property in the Multiasset Portfolio?”, *Journal of Property Finance*, 6:3, 60-83
- Can, A. (1998), “GIS and Spatial Analysis of Housing and Mortgage Markets”, *Journal of Housing Research*, Volume 9, Issue 1.

- Cameron, G., Muellbauer, J. and Murphy, A. (2006), “Was There a British House Price Bubble? Evidence from a Regional Panel”, *Centre for Economic Policy Research, London.*
- Capozza, D. R. and Seguin, P. J. (1995), "Expectations, Efficiency, and Euphoria in the Housing Market," *NBER.*
- Capozza, D., Hendershott, P., Mack, C., and Mayer, C. (2002), “Determinants of real house price dynamics”, *NBER.*
- Case, Karl E. and Shiller, Robert J. (1990), “Returns and risk on real estate and other investments: More evidence”, *Journal of Real Estate Portfolio Management*, Vol.8, No. 3: 265-279.
- Case, Karl E.; Quigley, John M.; and Shiller, Robert J. (2005), "Comparing Wealth Effects: The Stock Market versus the Housing Market," *Advances in Macroeconomics:* Vol. 5: Iss. 1, Article 1.
- Castro, Jose Camilo Otalora. (2008), “Strategic Framework for Real Estate Investment in Emerging Markets: The Case of Commercial Markets in Bogota, Columbia”
- Cieślik Andrzej (2009), “Bilateral trade volumes, the gravity equation and factor proportions”, *Journal of International Trade & Economic Development, Taylor and Francis Journals*, 18(1), pp. 37-59.
- Chari.V. V., Lawrence Christiano, and Patrick J. Kehoe(2008), “Facts and Myths about the Financial Crisis of 2008.”
- Chau (2008), K. W., S. K. Wong, and G. Newell. (2003), “Performance of Property Companies in Hong Kong: A Style Analysis Approach”, *Journal of Real Estate Portfolio Management*, Vol. 9, No. 1, pp. 29- 16.
- Cheng Y, Han X (2013), “Does large volatility help? – stochastic population forecasting technology in explaining real estate prices”, *Journal of Population Economics.*
- Chin, Lawrence, & Fan, Gang-Zhi. (2005), “Autoregressive analysis of Singapore's private residential prices”, *Property Management*, 23(4), 257-270.
- Chow, H. K. and Choy, K. M. (1993), “A leading economic index for monitoring the Singapore economy”, *Singapore Economic Review*, Vol. 38 No. 1, pp. 81-94.

- Chua, A. (1999), “The Role of International Real Estate in Global Mixed-Asset Investment Portfolios”, *Real Estate Portfolio Management*, 5:2, 129–37.
- Ciochetti, B. A., Craft, T. M. and Shilling, J.D. (2002), “Institutional Investors’ Preferences for REIT Stocks”, *Real Estate Economics*, 30: 4, 567–94.
- Clapp, J. (1993),“Dynamics of Office Markets: Empirical Findings and Research Issues”, *Washington, DC, Urban Institute Press*.
- Clapp, J. M. and C. Giaccotto (1992), “Estimating Price Indices for Residential Property: A Comparison of Repeat-Sales and Assessed Value Methods”, *Journal of the American Statistical Association,* 87, 300-06.
- Clapp, J. M. (1997), “How GIS Can Put Urban Economic Analysis on the Map”, *Journal of Housing Economics,* 6, 368–86.
- Clayton Jim (2009), “Cap Rates and Real Estate Cycles”, *Journal of Real Estate Finance and Economics.*
- Clayton, J. and MacKinnon, G. (2003), “The relative importance of stock, bond and real estate factors in explaining REIT returns”, *Journal of Real Estate Finance and Economics*, 27(1), 39–60.
- Clements, M. P. and Hendry, D. F. (1996), “Intercept corrections and structural change”, *Journal of Applied Econometrics*, 11, 475-94.
- Coleman S., Leone V. (2011), “Time series characteristics of UK commercial property returns: Testing for multiple changes in persistence”, *Journal of Real Estate Finance and Economics*.
- Cooper, Downs and Paterson. (1999), “REIT Selection & Portfolio Construction using Operating Efficiency”, *Journal of Real estate Portfolio Mgt,* p. 3.
- Cotter, John and Stevenson, Simon. (2005), “Multivariate Modelling of Daily REIT Volatility”, *Journal of Real Estate Finance and Economics*, 32:3, 305-25.
- Carter, Jame R. (1989), “On Defining the Geographic Information System”.
- Castle, G. (1993), “GIS: A Time-saving Tool for Targeting Candidate Properties”, *Business Geographics*, Vol.1 (2):20.
- Crawford, G. and Fratantoni, M. (2003), “Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices”, *Real Estate*

Economics, 31, 223-43.

- Cowley, M (2007), "Project Viability Studies and Market Forecasts- A Survey of Property Developers", *Urban Developer,* Issue No.1, p. 24
- Dangermond, J. (1988), "A review of digital data commonly available and some of the practical problems of entering them into a GIS", *American Society of Photogrammetry and Remote Sensing*, Vol. 5, pp. 1-10.
- Das, S., R. Gupta, and A. Kabundi (2009), "Could We Have Predicted the Recent Downturn in the South African Housing Market?" *Journal of Housing Economics* 4:325-335.
- Davis Richard and Brockwell J.Peter (1986), "Time Series: Theory and Methods". *2nd Edition by Springer-Verlag New York, Inc*.
- Davenport, M. L. and S.E. Nicholson (1993), "On the relation between rainfall and the Normalized Difference Vegetation Index for diverse vegetation types in East Africa", *International Journal of Remote Sensing* 14:2369-2389.
- Deardorff A. V. (1998), "Determinants of bilateral trade: Does gravity work in a neoclassical world?", *The regionalization of the world economy. Chicago: The University of Chicago Press*.
- Densham, Paul J., and Michael F. Goodchild. (1989), "Spatial Decision Support Systems: A Research Agenda", *American Society for Photogrammetry and Remote Sensing*.
- De Man, E. (1998), "Establishing a Geographical Information System in Relation to Its Use: A Process of Strategic Choice", *International Journal of Geographical Information Systems* 2: 245-61.
- Denzer R. (2005), "Generic integration of environmental decision support systems-state-of-the-art", *Environmental Modelling & Software* 20 (10) (2005) 1217-1223.
- Diavatopoulos, D., J. Doran, and D. Peterson. (2008), "The Information Content in Implied Idiosyncratic Volatility and the Cross-section of Stock Returns: Evidence from the Option Markets", *Journal of Futures Markets*, 28:11, 1013-16.
- DiBartolomeo, Dan., Gold, Richard., BelevEmilian and Baldwin, Ken. (2005), "A New Approach to Real Estate Risk", *Northfield Information Services*.

- Dickinson, H. J., and Hugh W.Calkins. (1988), "The Economic Evaluation of Implementing a GIS", *International Journal of Geographical Information Systems* 2(4): 307-27.
- Diebold, F. X. (1989), "Forecast combination and encompassing reconciling two divergent literatures", *International Journal of Forecasting*, Vol.5, pp. 589-92.
- Dolan, P., Peasgood, T., and White, M. (2008), "Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective wellbeing", *Journal of Economic Psychology*, 29, 94-122.
- Doan, T. (1994), *Regression Analysis of Time Series: User Manual*, 4th edn., Evanston, IL, Estima.
- Donlon K. H. (2007)," Using GIS to Improve the Services of a Real Estate Company".
- Dramowicz, E. (2005), "Retail Trade Analysis Using the Huff Model", *Directions Magazine*.
- Driffill, John and Sola, Martin (1998), "Intrinsic Bubbles and Regime Switching", *Journal of Monetary Economics*, Vol. 42, 357-373.
- DuaPami (2004)," Analysis of Consumers Perceptions of Buying Conditions for Houses".
- Engle, Robert F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica* 50:987-1007.
- Engle and Bollerslev, (1986), " Modelling The Persistence of Conditional Variances", *Econometric Reviews*, 5(1), 1-50.
- Efron, B., and Tibshirani, R. J. (1994), "An Introduction to the Bootstrap", *Boca Raton, FL: CRC-Press*.
- ESRI, ArcGIS Manual (2005). Published by ESRI, California.
- Fabozzi, F. and Tunaru, R. (2009), ‟Pricing Models for Real-Estate Derivatives‟, *International Center for Finance*.
- Fama, E. and French, K. (1989), "Business conditions and expected returns on stocks and bonds", *Journal of Financial Economics*, Vol. 25 No. 1, pp. 23-49.
- Fama, E. F., and French, K. (1993), "Common Risk Factors in the Returns on Stocks and Bonds"; *Journal of Financial Economics*, 33, 3-56.

- Farinella J. A., Graham J.E.,Markowski J., Schuhmann P.W. (2013), “The Relationships Between the Real Estate and Stock Markets in Poland”.
- Fei, Peng. (2009), “Risk Dynamics of Housing Market: Cross-sectional Variations, Time Variations and Economic Fluctuations”; *Real Estate Research Institute (RERI).*
- Feldman, B. E. (2003), “Investment Policy for Securitized and Direct Real Estate”, *Journal of Portfolio Management*, Special Real Estate Issue, 112-21.
- Ferson, W. and Campbell, R. H. (1991), “The variation of economic risk premiums”, *Journal of Political Economy*, Vol. 99 No. 2, pp. 385-415.
- Fik T. C. Sidman, B. Sargent, and R. Swett (2005), “A Methodology for Delineating A Primary Service Area For Recreational Boaters Using A Public Access Ramp: A Case Study Of Cockroach Bay”, *The Florida Geographer* 36: 23-40.
- Findlay, M. D, R. D McBride, S. D. Messner and J.S Yourmark (1979), “Optimal Real Estate Portfolios”, *Journal of the American Real Estate and Urban Economics Association* pp. 298-317.
- Fisher, J. D. and Liang, Y. (2000), “Is Sector Diversification More Important Than Regional Diversification?”, *Real Estate Finance*, vol. 17, no. 3, pp. 35-59.
- Fisher, Jeffrey D. and David Geltner(2000), "De-Lagging The NCREIF Index: Transaction Prices And Reverse-Engineering," *Real Estate Finance*, v17, 7-22.
- Fisher, Jeffrey, Ling, David C. and Naranj, Andy (2006), “The Dynamics of Capital Flows and Property Returns: A Disaggregated Analysis of Metropolitan Areas and Property Types”, *Center for Real Estate Studies, Kelley School of Business, Indiana University, USA.*
- Flavin, Marjorie and Takashi Yamashita (2002), “Owner-occupied housing and the composition of the household portfolio”, *American Economic Review* 92, 345–362.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003) “Do financial variables help forecasting inflation and real activity in the euro area?” *Journal of Monetary Economics* 6 (2003): 1243-1255.
- Francica, Joseph R. (1993),” Profiting from a Geographic Information System”, *GIS World*, 75-84.

- French, N. and Gabrielli, L. (2004), “The uncertainty of valuation”, *Journal of Property Investment & Finance*, Vol. 22 No. 6, pp. 484-500.
- Fujiwara, D. and Campbell, R. (2011), “Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches”.
- Fuller, W. A. (1976),” Introduction to Statistical Time Series”, *New York, John Wiley*.
- G. A. F. Seber (1977), “Linear Regression Analysis”, *Wiley, New York.*
- Gallin, J. (2006), “The long-run relationship between house prices and income: Evidence from local housing markets” *Real Estate Economics*, 34(3):417–438.
- Georgi Georgiev (2002), “The Benefits of Real Estate Investment”, *Journal of Property Management*.
- Gehner E., Halman J. I. M. and de Jonge, H. (2007),“Risk Management in the Dutch Real Estate Development Sector: A Survey”.
- Geltner, D. (1993), "Estimating Market Values from Appraised Values Without Assuming an Efficient Market", *Journal of Real Estate Research* 8(3) 325-346.
- Geltner, David M., and Norman G.Miller, (2006), “Commercial Real Estate Analysis and Investments”, *Cincinnati: South-Western Educational Publishing*.
- Geltner, David M. (1998), "How Accurate Is The NCREIF Index As A Benchmark, And Who Cares?," *Real Estate Finance,* v14 (4, Winter), 25-37.
- German, J. C. (1999), “An Introduction to Quantifying Location’s Value in Real Estate Assessment Models Using Global Respone Surface Techniques”.
- Gibler M. Karen and Nelson L. Susan, (1998), “Consumer Behaviour Applications To Real Estate”.
- Gilberto, M., Foort, H., Hoesli, M and MacGregor B. (1999, “Optimal Diversification Within Mixed-Asset Portfolios Using a Conditional Heteroskedasticity Approach: Evidence from the U.S. and U.K”, *Journal of Real Estate Portfolio Management*, 5, 31-45.
- Giliberto, M. (1990), “Equity Real Estate Investment Trusts and Real Estate Returns”, *Journal of Real Estate Research*, 5:2, 259-63.
- Giliberto, S. M. (1993), “Measuring Real Estate Returns: The Hedged REIT Index”, *Journal of Portfolio Management*.

- Glaeser, E. L. and Gyourko, J. (2006), “Housing dynamics”, *National Bureau of Economic Research*.
- Gloudemans, R. J. (1990), “Adjusting for Time in Computer-Assisted Mass Appraisal”, *Property Tax Journal*, 9, 83-99.
- Girouard N and Blondal S. (2001), “House Prices and Economic Activity”, *OECD Economics Department* 279-2001.
- Goodchild, Michael F. (1992), “Geographical Information Science”, *International Journal of Geographical Information Systems* 6(1): 31-45.
- Gooijer J. G., Hyndman R.J. (2006), “25 years of time series forecasting”, *International Journal of Forecasting*, 443-473.
- Goetzmann, William and Wachter, Susan (1996), “The Global Real Estate Crash: Evidence from an International Database”.
- Goldhill O. (2014), 'Best city for foreign property investment opportunities' *The Telegaph (www.telegraph.co.uk) 6th Jan*.
- Gordon J. (2009), “From Markowitz to Modern Risk Management”, *The European Journal of Finance*, Vol. 15, Nos. 5-6, 451-461.
- Gordon, J., Canter, T and Webb, J. R. (1998), “The Effect of International Real Estate Securities of Portfolio Diversification”, *Journal of Real Estate Portfolio Management*, 4: (2), 83–91.
- Graff, Richard A. and Michael S. Young (1996), “Real Estate Return Correlations: Real World Limitations On Relationships Inferred From NCREIF Data,” *Journal of Real Estate Finance and Economics*, 121-142.
- Greene WH (2008), “Econometric analysis”, *6. Prentice Hall, Englewood Cliffs*.
- Grenadier, Steven R., Neng Wang. (2006), “Investment under Uncertainty and Time-Inconsistent Preferences”, *Journal of Financial Economics*, 84: 382-399.
- Grissonm T., Kuhle J. and Walther C. (1987), “Diversification Works in Real Estate, Too”, *The Journal of Portfolio Management*, vol. 13. No. 2, pp. 66-71.
- Gupta, S. K. and Cozzolino, J.M. (1974), “Fundamentals of Operations Research for Management”, *Journal of Property Management*.

- Gu, A. Y. (2002), “The predictability of house prices”, *Journal of Real Estate Research*, 24, 214-33.
- Hancock Beverly (2002), “An Introduction to Qualitative Research”, *Journal of Project Management.*
- Harrell, Frank E., Jr. (2001), "Sections 9.2, 10.5, Regression modeling strategies”, *New York: Springer-Verlag,* ISBN 0387952322.
- Hasan S. Mohammad (2006), “Business Cycles, Mortgage Rates and Housing Starts in the United Kingdom-An Empirical Analysis”, *Journal of Economic Literature*, Issue No. 67.
- Hepsen, Ali and Vatansever, Metin (2010), “Forecasting Future Trends in Dubai Housing Market by Using Box-Jenkins Autoregressive Integrated Moving Average International”, *Journal of Housing Markets and Analysis*, 4(3), 210-223.
- Hernandez, Tony (2007), “Enhancing Retail Location Decision Support”, *Journal of Retailing and Consumer Services*.
- Herbst, A. F. (1992), “Analyzing and Forecasting Futures Prices”, *John Wiley & Sons, Inc., New York, NY.*
- Hilber, C. A. L. and WouterVermeulen (2009), “The effects of supply constraints on housing costs: Empirical evidence for England and assessment of policy implications” *NHPAU Research Innovation Fund Project Interim Report, London School of Economics.*
- Hilber, C. A. L. and Robert-Nicoud, F. (2009), “On the Origins of Land Use Regulations: The Influential Landowner Hypothesis”.
- Hirotugu Akaike (1974), “A new look at the statistical model identification”, *IEEE Transactions on Automatic Control* 19 (6): 716–723.
- Hishamuddin, Mohd Ali. (2006), ”Modern Portfolio Theory: Is There Any Opportunity for Real Estate Portfolio?”, *Journal of Real Estate Portfolio Management*.
- Holland, Steven A. (1995), “Uncertainty and the Rate of Commercial Real Estate Development”, *RERI.*
- Hollies, R. (2007), “International variation in office yields: a panel approach”, *Journal of Property Investment and Finance,* 25(4), 370-87.

- Holt Jeff (2009), "A Summary of the Primary Causes of the Housing Bubble and the Resulting Credit Crisis", *The Journal of Business Inquiry*, 8, 1, 120-129.
- Hull, John C. and Basu, Sankarshan. (2010), "Options, Futures, and Other Derivatives", *Dorling Kindersley* (India) Pvt. Ltd. pp. 288.
- Hung, Szu-Yin (Kathy) and Glascock, John L. (2009, "Volatilities and Momentum Returns in Real Estate Investment Trusts", *Journal of Real Estate Finance and Economics*, Volume 41, Number 2: 126-149.
- Hutchison Norman E., Adair A.S. and Lehay, Iain. (2005), "Property Research Priorities in Australia", 8.2: 127-39.
- Hoyt, Homer. (1970), "One Hundred Years of Land Values in Chicago", *NY: Arno Press & The New York Times*.
- Ishijima, H. and A. Maeda (2012), "Real Estate Price Modeling and Empirical Analysis," *The International Journal of Economic Policy Studies*.
- Jensen, D. L. (1991), "Modeling Time Effects in Computer-Assisted Mass Appraisal", *Property Tax Journal*, 10, 95–123.
- Johnson and E. Weber (2008), "Decision under uncertainty: Psychological Economic & Nuroeconomic Explanation of Risk preference", *Behavioral Economics & the Brain*.
- Johnson, K. H., J. A. Wiley, and Z. Wu. (2007), "On the Relationship between Commercial Property Price and its Selling Time", *Journal of Real Estate Portfolio Management*, 379-388.
- Johnson, (2001), "Using ArcGIS Spatial Analyst. Redlands", *ESRI Press*.
- Johnston, H. J. (2000), "The Dictionary of Human Geography", *Oxford: Blackwell Publishers.*
- Kearl, J. H. (1979), "Inflation, Mortgages, and Housing," *Journal of Political Economy*, vol. 87, p. 1-29.
- Kelliher, C. F. and Mahoney, L. S. (2000), "Using Monte Carlo simulation to improve long term investment decisions", *The Appraisal Journal*, Vol. 68 No. 1, pp. 44-56.

- Kenneth A., Scharfstein, David S and Stein, Jeremy C. (1992), "Risk Management: Coordinating Corporate Investment and Financing Policies"; *Journal of Finance,* Vol.48 (December): pp. 1629-1658.
- Kochera, Andrew (1994), "Home Characteristics and Property Insurance", *Housing Economics* 42(12): 9-12.
- Kochin, L. A. and Parks, R. W. (1982), "Vertical equity in real estate assessment: A fair Appraisal", *Economic Inquiry,* 20(4): 511-532.
- Kuhle, J. L. and Moorehead, J. D. (1989), "Applying the Bootstrap Technique to Real Estate Appraisal: An Empirical Analysis," *The Journal of Real Estate Research* Vol. 5 No. 1, pp. 33-40**.**
- Lane, Philip (2001), "The New Open Economy Macroeconomics: A Survey", *Journal of International Economics,* 54, 2: 235-266.
- Leamer Edward E., Levinsohn James (1995), "International Trade Theory, the Evidence", *Handbook of International Economics 3. North-Holland.*
- Lee, S. L. (2005), "The Return Due to Diversification of Real Estate to the U.S. Mixed-Asset Portfolio", *Journal of Real Estate Portfolio Management*, Vol. 11, No. 1, pp. 19- 28.
- Li, H., Kong, C. W., Pang, Y. C. and Yu, L. (2003), "Internet-based geographical information system for E-commerce application in construction material procurement", *Journal of Construction Engineering and Management* 129, 689–97.
- Liang, Y., Chatrath, A., and Webb, J. R. (1996), "Hedged REIT Indices", *Journal of Real Estate Literature.*
- Ling, D. C. and Naranjo, A. (1999), "The Integration of Commercial Real Estate Markets and Stock Markets", *Real Estate Economics*, 27:3, 483-515.
- Ljung, G. M. and Box, G. E. P. (1978), "On a measure of lack of fit in time series models". *Biometrika* 65 (2): 297-303.
- Lo, A. W. and Wang J. (1995), "Implementing option pricing models when return assets are predictable" *Journal of Finance*, 50: 87-129.
- Lynn David and Wang Tim (2011), "Applied Analytics and quantitative methods for private real estate investing", *Real Estate Mathematics*.

- Lyons, T. and Skitmore, M. (2004), "Project Risk Management in the Queensland Engineering Construction Industry: A survey", *International Journal of Project Management*, Vol. 22, No. 1: 51-61.
- MacGregor, B. D. and N. Nanthakumaran (1992), "The Allocation to Property in the Multi-asset Portfolio. The Evidence and Theory Revisited", *Journal of Property Research*, 9:1,5-32.
- Maguire, David J. (1991), "An overview and definition of GIS", *ESRI Press.*
- Malliaris, Anastasios (2008), "Stochastic Optimal Control", *The New Palgrave Dictionary of Economics*.
- Malpezzi, S. (1999), "A simple error correction model of house prices", *Journal of Housing Economics*, 8(1): 27-62.
- Markowitz, H. M. (1952), "Portfolio Selection", *Journal of Finance*, 7:1, 77–91.
- Marks, A., C. Stanley, and G. Thrall (1994), "Criteria and Definitions for the Evaluation of Geographic Information System Software for Real Estate Analysis", *Journal of Real Estate Literature*, Vol.2 (2): 227-244.
- Martin A. Smith (2010), "Diversifying a Portfolio with Real Estate; Investment Management", *Journal of Real Estate Portfolio Management*.
- McGreal, Stanley, Adair, Alastair and Webb, James R. (2009), "Optimal Diversification in US/UK Private Real Estate Only Portfolios: The Good, the Bad, and the Uncertain", *Journal of Real Estate Portfolio Management*: 1083-5547.
- Mc Gough, T., Tsolacos, S and Olkkonen, O. (2000), "The predictability of office property returns in Helsinki", *Journal of Property Valuation and Investment* 18(6), 565-85.
- McNamara, M. P. M; Paul, F. (1997), "Issues in the development and application of property market forecasting: the investor's perspective", *Journal of Property Finance*, 8.4:363-376.
- McNeely, J. A., K. R. Miller, W. Reid, R. A. Mittermeirer and T. Werner (1990), "Conserving the World's Biological Diversity", WRI, *IUCN, Conservation International and World Bank, Washington DC*.

- Gland Miles, M. and N. Tolleson. (1997), "A Revised look at How Real Estate Compares with Other Major Components of Domestic Investment Universe", *Real Estate Finance*, 14, 11–20.
- Mehmedovic, M. H., Arslanagic, M., Kadic, S. and Cicic, M. (2010), "Decision Making in Real Estate Purchase The Case of Bosnia-Herzegovina".
- Meese, R. and Wallace, N. (1991), "Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices". *Real Estate Economics*, 19(3): 308-332.
- Miller, N. and M. Sklarz (1986), "A Note on Leading Indicators of Housing Market Price Trends", *Journal of Real Estate Research*, 1:1, 99–109.
- Mollart, R. G. (1988), "Monte Carlo simulation using LOTUS 123", *Journal of Property Valuation*, Vol. 6 No. 4, pp. 419-433.
- Mueller, Andrew G and Mueller, Glenn R. (2003), "Public and Private Real Estate in a Mixed-Asset Portfolio", *Journal of Real Estate Portfolio Management*.
- Mueller, G. R., K. Pauley, and W.A. Morrell. (1994), "Should REITs Be Included in a Mixed-Asset Portfolio?", *Journal of Real Estate Finance*, 11, 23-8.
- Muellbauer, J. and Murphy, A. (1997), "Booms and busts in the UK housing market", *The Economic Journal*, 107:1701-1727(27).
- Mughees, Shaukat (2010), "The Benefits and Importance of Commercial Real Estate", *MPRA Paper Number 28268*.
- Mun, J. (2002), "Real Options Analysis: Tools and Techniques for Valuing Strategic Investments and Decisions", *Wiley Finance, New York*.
- Newell, G. and Kamineni, R. (2007), "The significance and performance of real estate markets in India", *Journal of Real Estate Portfolio Management* 13(2): 161-172.
- Newman, William H. (1935), "The Building Industry and Business Cycles", *Chicago: University of Chicago Press*.
- Nitsch, Harold (2006), "Pricing Location: A Case Study of the Munich Office Market", *Journal of Property Research*, 94-96.

- Ng, L. H. (1993), “The importance of Geographic information system for Malaysia’s future development with particular reference to real estate development”.
- OECD (2004), “The contribution of housing markets to cyclical resilience, *OECD” Economic Studies* 38(1), 125-156.
- Oldfield, George S. and Santomero, Anthony M. (1997), “The Place of Risk Management in Financial Institutions”, *The Wharton School, University of Pennsylvania*.
- Ortalo-Magne, Francois and Rady, Sven. (2003), “Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints”, *University of Wisconsin, Madison, Mimeo*.
- Pace, R. K., Barry, R., Gilley, O. and Sirman, C. F. (2000), “A method for spatial temporal forecasting with an application to real estate prices”, *International Journal of Forecasting*, 16, 229-46.
- Pagourtzi E., Assimakopoulos V., Hatzichristos and French N. (2003), “Real Estate Appraisal: a review of valuation methods”, *Journal of Property Investment and Finance*, 21,4.
- Pearson K. Jesse (2007), “A Comparative Business Site-Location Feasibility Analysis using Geographic Information Systems and the Gravity Model”.
- Pearson K (1929), “Methods of Moment and Methods of Maximum Likelihood”, *Biometrika* 21: 370-375.
- Peyton, Martha S. and Bardzik, Steven. (2008), “What is Risk Management and how does it apply to Real Estate”, *Real Estate White Paper, TIAA-CREF Asset Management, Fall Edition.*
- Podor, Andrea (2010), *Scientific Journal of Riga Technical University Economics and Business. Economy: Theory and Practice*.
- Pribam, Karl (1940), "Residual, Differential, and Absolute Urban Ground Rents and Their Cyclical Fluctuations", *Econometrica* 8: 62-78.
- Property Industry Alliance Data Report 2012.
- Pyhrr, Stephen A., Born, Wales I. and Webb, James R. (1990), “Development of a Dynamic Investment Strategy Under Alternative Inflation Cycle Scenarios”, *The Journal of Real Estate Research*, Vol. 5, No. 2; pp. 177- 193.

- Quan, D. C. and S. Titman, (1999), “Do Real Estate Prices and Stock Prices Move Together? An International Analysis”, *Real Estate Economics*, 27:2, 183–207.
- REA (The Real Estate Almanac Inc.), 1999, Automatic Valuation Model (AVM) design methods and methodologies.
- Rabianski, Joseph, Market analyses and Appraisals, Real Estate Issues, Winter 1995, 45-49.
- Reinhart, Carmen, and Kenneth Rogoff, 2009, “This Time is Different: Eight Centuries of Financial Folly”, Princeton *University Press: Princeton, New Jersey*.
- Rodrigue, J-P., Comtois, C., Slack, B. (2009), *“The Geography of Transport Systems”*, *London, New York.*
- Rose A. (2000), “One money, one market: the effect of common currencies on trade”, *Economic Policy*, 30, pp. 9-45.
- Rossi, B. (2012), “Advances in Forecasting under Instabilities”, *Handbook of Economic Forecasting, Elsevier-North Holland*.
- Reed, Richard (2002), “The Cyclical Nature of Residential Housing Markets”, Paper presented in the Eighth Annual Conference of Pacific Rim Real Estate Society, Christchurch, 21-23 January.
- Richie, Stewart (2008), “Estimating the volatility of property assets”, *Volatility of Property Assets* Version 1.2.
- Saiz, A. (2008), ‘On local housing supply elasticity,’ The Wharton School.
- Saks, R. E. (2008), ‘Job creation and housing construction: Constraints on metropolitan area employment growth,’ *Journal of Urban Economics* 64, 178-195.
- Salter, S. P. (2006), “NASDAQ REITs and Optimal Investment Policy: Evidence from 1995 to 2005”, *Journal of Real Estate Portfolio Management*, Vol. 12, No. 3: 201-207.
- Schumpeter, Joseph A (1983), “The Theory of Economic Development”, *New Brunswick: Transaction Books*.
- Schwarz, Gideon E. (1978), "Estimating the dimension of a model", *Annals of Statistics* 6 (2): 461–464.

- Seiler, M. J., Webb, J. R. and Myer, F. C. N. (1999), "Diversification Issues in Real Estate Investment", *Journal of Real Estate Literature*, 7:2, 163-79.
- Serrano, C. and Hoesli, M. (2007), "Forecasting EREIT Returns", *Journal of Real Estate Portfolio Management*, Vol. 13, No. 4, pp. 293-309.
- Sharpe, William F. (1992), "Management Style and Performance Management", Journal *of Portfolio Management*, Winter: 7-19.
- Sharpe, W .F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk". *Journal of Finance*, 19: (3), 425– 42.
- Simon Emma (2013), "Revealed UK's Top 10 Buy to Let Hotspots". *The Telegraph*.
- Sirmans, Stacy G., Friday, Swint H. and Zietz, Emily N. (2003), "The Environment and Performance of Real Estate Investment Trusts", *Journal of Real Estate Portfolio Management*, Vol. 9, No. 2: 127-165.
- Sklarz, Michael A., Miller, Norman G., and Gersch, Will (1987), "Forecasting Using Long Order Autoregressive Processes: An Example Using Housing Starts", *Real Estate Economics,* 154(4), 374-388.
- Skole, D. and C. Tucker (1993), "Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988", *Science*, 260:1905-1910.
- Srinivasan, V., A. D. Shocker (1973), "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgements", *Psychometrika*, 38, 473-493.
- Stephen C. Guptill (1995), "Elements of Cartography, 6thed", *New York, John Wiley and Sons, Inc*.
- Stevenson, Simon. (2007), "A comparison of the forecasting ability of ARIMA models", *Journal of Property Investment & Finance*, 25(3), 223-240.
- Stockman, Alan and Tesar, Linda (1995), "Tastes and Technology in a Two Country Model of the Business Cycle: Explaining International Co-movements", *Americans Economic Review*, 85, 1: 168-85
- Stock, J. H., and M. W. Watson (2003), "Forecasting Output and Inflation: The Role of Asset Prices", *Journal of Economic Literature* 3(2003): 788-829.
- Sweet, A. J. (1995), "Land Investment, The Handbook of Real Estate Portfolio Management, Burr Ridge, IL", *Irwin Professional Publishing*.

- Thrall, G. and A. Marks (1993), “Functional Requirements of a Geographic Information System for Performing Real Estate Research Analysis”, *Journal of Real Estate Literature*, Vol.1 (1): 49-61.
- Thrall, G. I. (2002), “Business Geography and New Real Estate Market Analysis. New York”, *Oxford University Press*.
- Thrall, G. I. and del Valle, Juan (1996), ”Retail Location Analysis: Antecedents”, *Geo Info Systems*, 6 (6): 48-52.
- Thrall, G. I. and del Valle, Juan (1997), “The Calculation of Retail Market Areas: The Reilly Model”, *Geo Info Systems,* 7 (4): 46-49.
- Thrall, G. I. (2002), “Business Geography and New Real Estate Market Analysis”, *Oxford Press: Oxford.*
- Timmermann, A. and Perez-Quiros, G. (2001), “Business cycle asymmetries in stock returns: evidence from higher order moments and conditional densities”, *Journal of Econometrics*, 103, 259-306.
- Trippi, R. R. (1977), “Estimating the Relationship between Price and Time to Sale for Investment Property”, *Management Science*, 23:8, 838–43.
- Tse, Raymond. (1997), “An application of the ARIMA model to real-estate prices in Hong Kong”, *Journal of Property Finance,* 8(2), 152-163.
- Uher, T. E. and Toakley, A.R. (1999), “Risk Management in the Conceptual Phase of a Project”, *International Journal of Project Management*, Vol. 17, No. 3: 61-69.
- Varian, H. (2005), “Bootstrap Tutorial”, *Mathematica Journal*, 9, 768-775.
- Vasanthi Peter, Raja Peter (2006), “Risk Management Model: An Empirical Assessment of the Risk of Default”, *International Research Journal of Finance and Economics*.
- Vishwakarma, V. K. (2013), “Forecasting Real Estate Business: Empirical Evidence From The Canadian Market”, *Global Journal of Business Research,* Vol.7, No.3.
- Viezer, Timothy W. (2010), “The Application of Modern Portfolio Theory to Real Estate: A Brief Survey”, *Handbook of Portfolio Construction*, Part 3, 733-760.
- VonHippel, Paul T. (2005), “"Mean, Median, and Skew: Correcting a Textbook Rule”", *Journal of Statistics Education* 13 (2)

- Waggle, D and P. Agrawal. (2006), "The Stock–REIT Relationship and Optimal Asset Allocations", *Journal of Real Estate Portfolio Management*, Vol. 12, No. 3: 209- 13.
- Webb, B and O'Keefe, J. (2002), "The Case for Global Real Estate", *UBS Global Asset Management*.
- Weber, B. R. (2001), "Market Value Without a Market", *A Sequel, Integrating GIS and CAMA*.
- Weeks, S. (2003), "The devaluation of capital budgeting in real estate development firms Point of view", *Journal of Real Estate Portfolio Management*, Vol. 9 No. 3, pp. 265-268.
- West, K. D. (2006), "Handbook of Economic Forecasting", *Forecast Evaluation*, pp. 101-133 (North-Holland).
- William C. Wheaton, Raymond G. Torto, CRE, Petros S. Sivitanides, Jon A. Southard, Robert E. Hopkins, James M. Costello (2001), "Real Estate Risk: A Forward Looking Approach".
- Wilkens, Kathryn A., Heck, Jean L., and Cochran, Steven J. (2006), "The effects of mean reversion on alternative investment strategies", *Managerial Finance*, 32, 1.
- Williams, J. E. (1995), "Real Estate Portfolio Diversification by Sources of Return", *In A.L. Schwartz Jr. and S.D. Kapplin (Eds.), Alternative Ideas in Real Estate Investment, Boston, MA: Kluwer Academic Publishers.*
- Wilson P. J., Okunev J. (2001), "Enhancing information use to improve predictive performance in property markets", *Journal of Property Investment and Finance*.
- Winkler, R.L. (1989), "Combining forecasts: a philosophical basis and some current issues", *International Journal of Forecasting*, Vol.5 No.4, pp. 605-10
- Worthington Andrew C. and Higgs Helen (2003), "Co-movements in UK regional property markets: a multivariate co-integration analysis", *Journal of Property Investment & Finance*, 2003; 21,4.
- Wofford, L. E. (1978), "A simulation approach to the appraisal of income producing real estate", *Journal of the American Real Estate and Urban Economics Association*, Vol. 6 No. 4, pp. 370-394.

- Woychuk Ian (2011), “Exploring Real Estate Investments: Finding Investment Value”, *Investopedia*.
- Wright, Jeff R., Kurt A. Buehler (1993), “Probabilistic Inferencing and Spatial Decision Support Systems”, *In Expert Systems in Environmental Planning*.
- Wurtzebach, Charles H., Mueller, Glenn R. and Machi, Donna (1991). “The Impact of Inflation and Vacancy of Real Estate Return”, *Journal of Real Estate Research*.
- Yobaccio, Elizabeth, Ruben Jack H., Ketcham David C. (1989), “The Inflation-Hedging Properties of Risk Assets: The Case of REITS”.
- Young, M.S. and Graff, R.A. (1995), “Real estate Is Not Normal: A Fresh Look at Real Estate Distributions”, *Journal of Real Estate Finance and Economics*, Vol. 10: 225-259.
- Ziering, B., D. C. Liang, and W. McIntosh (1999), “REIT Correlations with Capital Market Indices: Separating Signal from Noise”, *Real Estate Finance*, 15, 61–7.
- Ziobrowski, A. J., Cheng, P. and Ziobrowski, B. J. (1997), “Using a Bootstrap to Measure Optimum Mixed-Asset Portfolio Composition”, *Real Estate Economics*, 25: 695-705.
- Zhou, Q., and Charnpratheep, K. (1996), “Fuzzy expert system and GIS for solid waste disposal siting in regional planning”, *Geographical Information Sciences*, 2, 37-50.
- www.esri.com. Accessed December 2012.
- wwww.NCREIF.org. Accessed September 2011.
- www.reit.com. Accessed November 2012.
- www.lightstone.co.za. Accessed October 2012.
- Manchester.gov.uk. Accessed November 2012.
- www.landregistry.gov.uk Accessed August 2011.

Appendix

- M002 Neighbourhood Thematic Maps

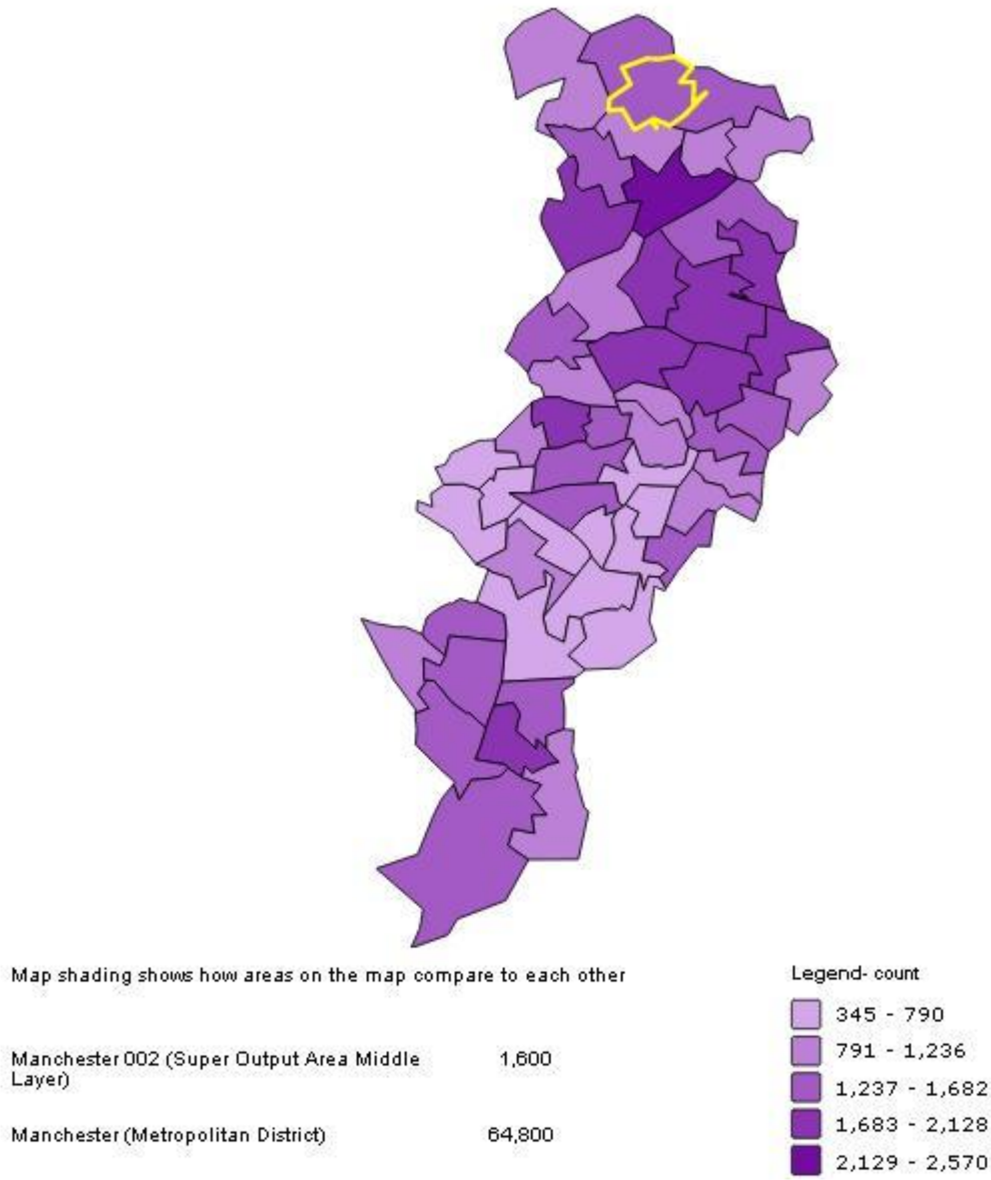

Figure: 7.43: Map Showing People Living on Benefit (Economically Deprived People)

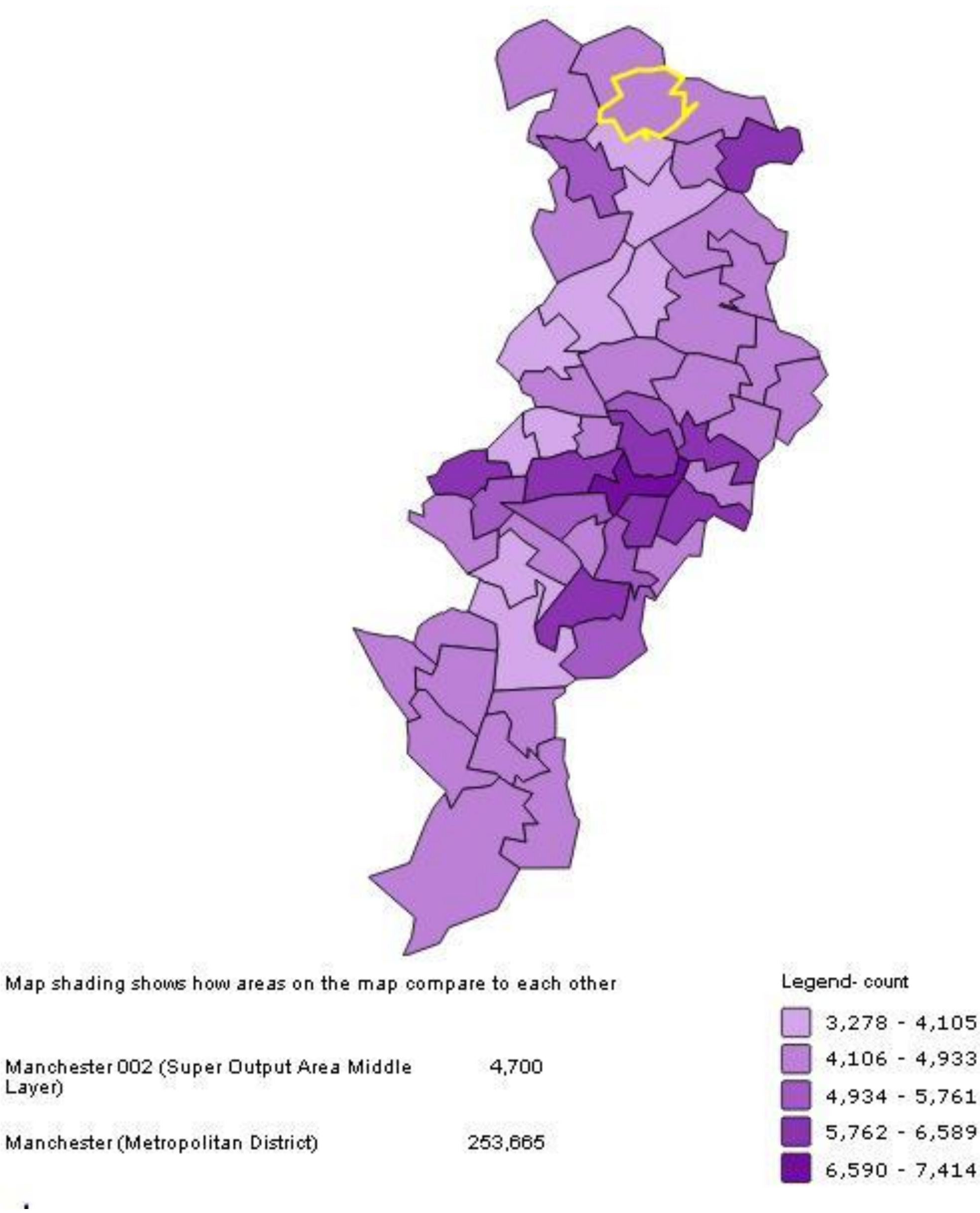

Figure 7.44: Map Showing People with Good Health

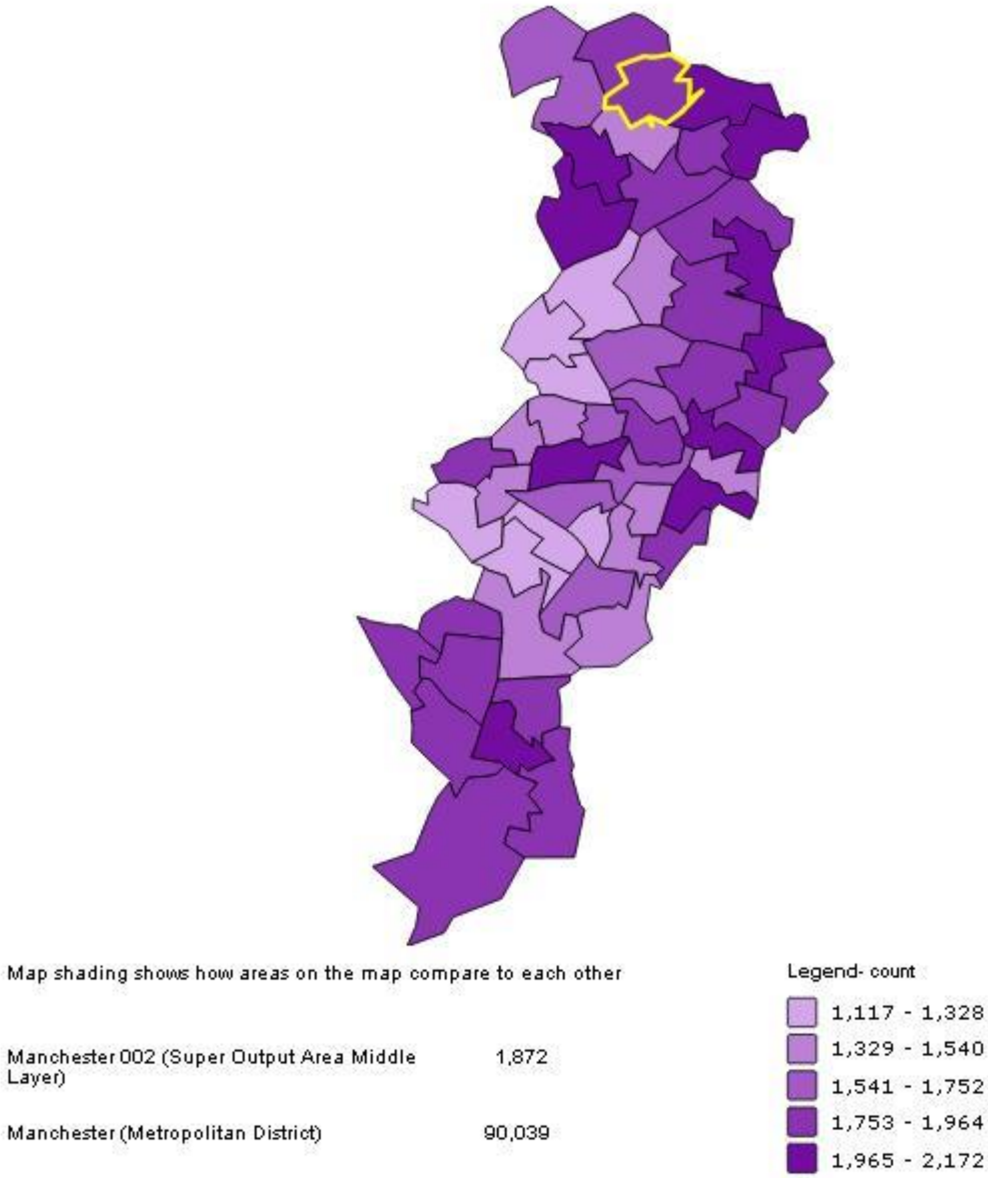

Figure 7.45: Map Showing People with Fairly Good Health

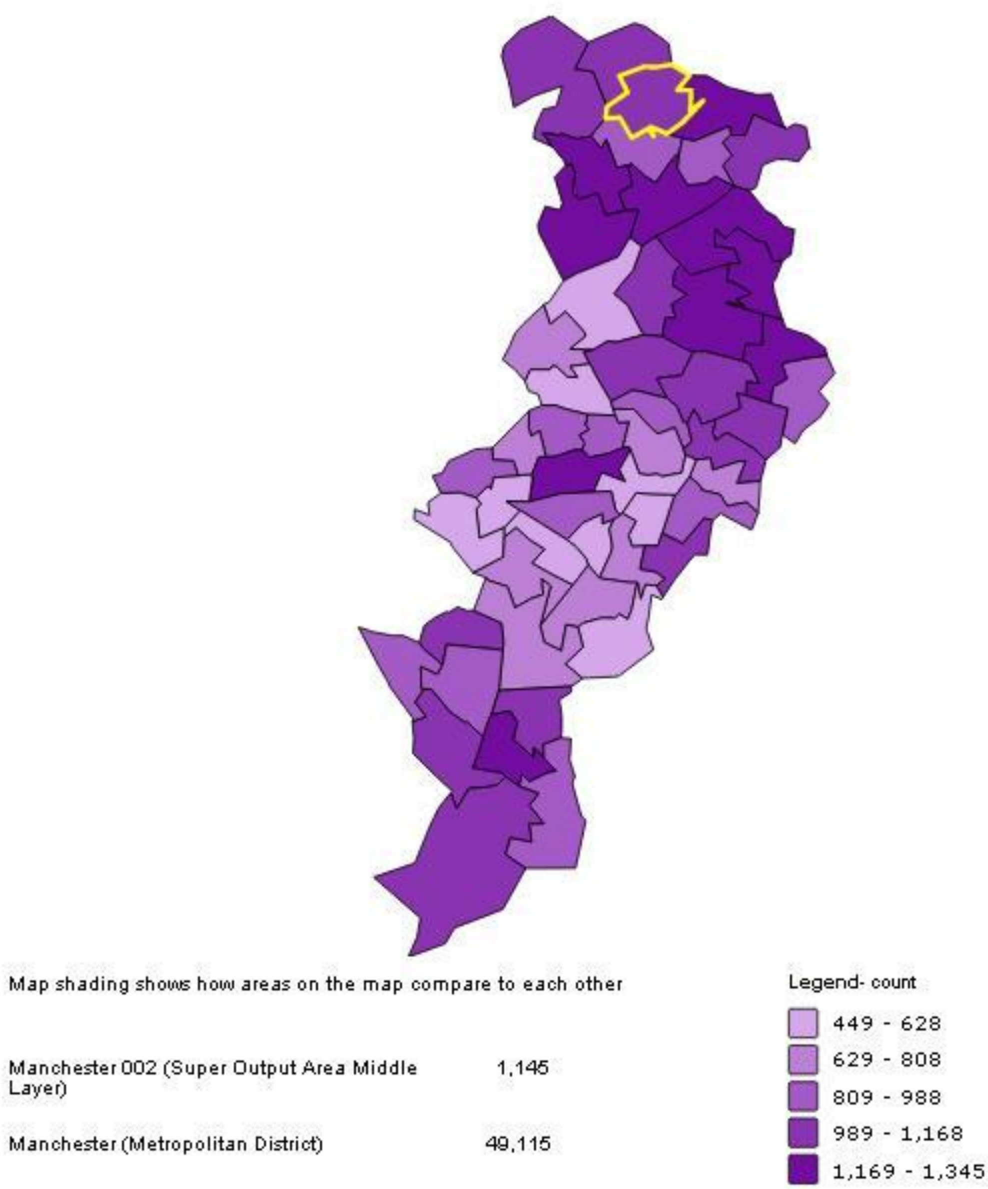

Figure 7.46: Map Showing People with Not Good Health

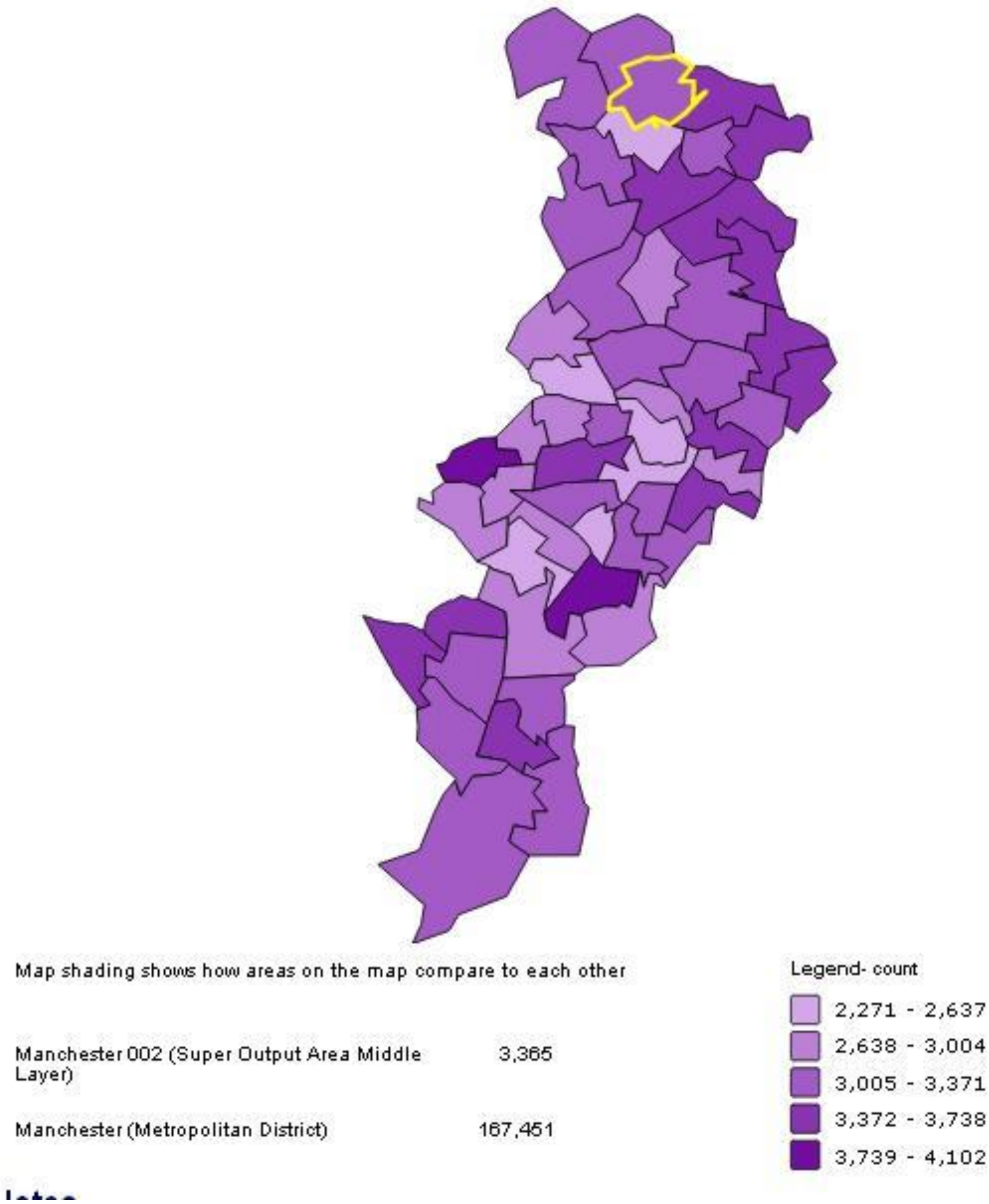

Figure 7.47: Map Showing Occupied Households

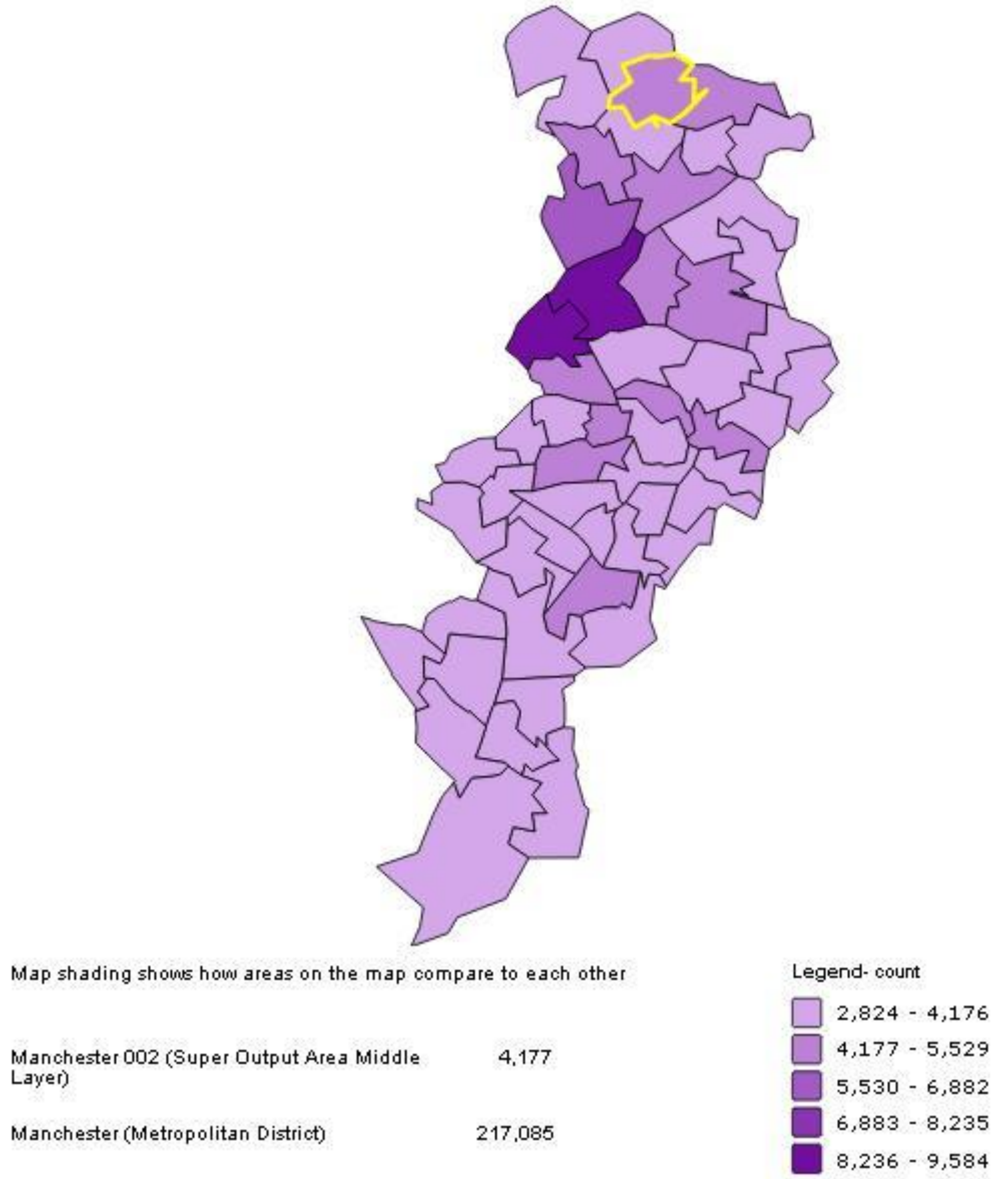

Figure 7.48: Map Showing Dwelling Stock

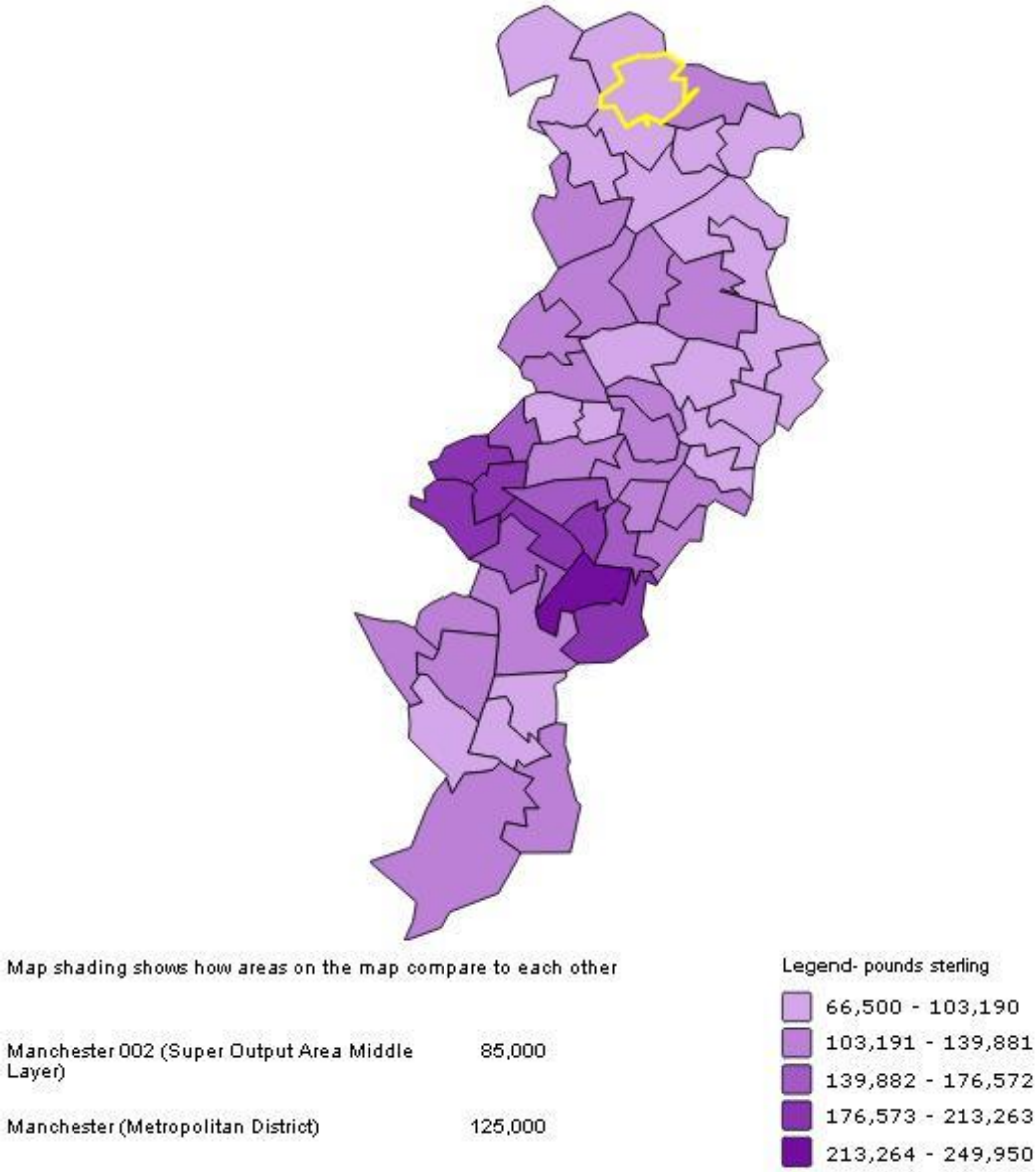

Figure 7.49: Map Showing Dwelling Price

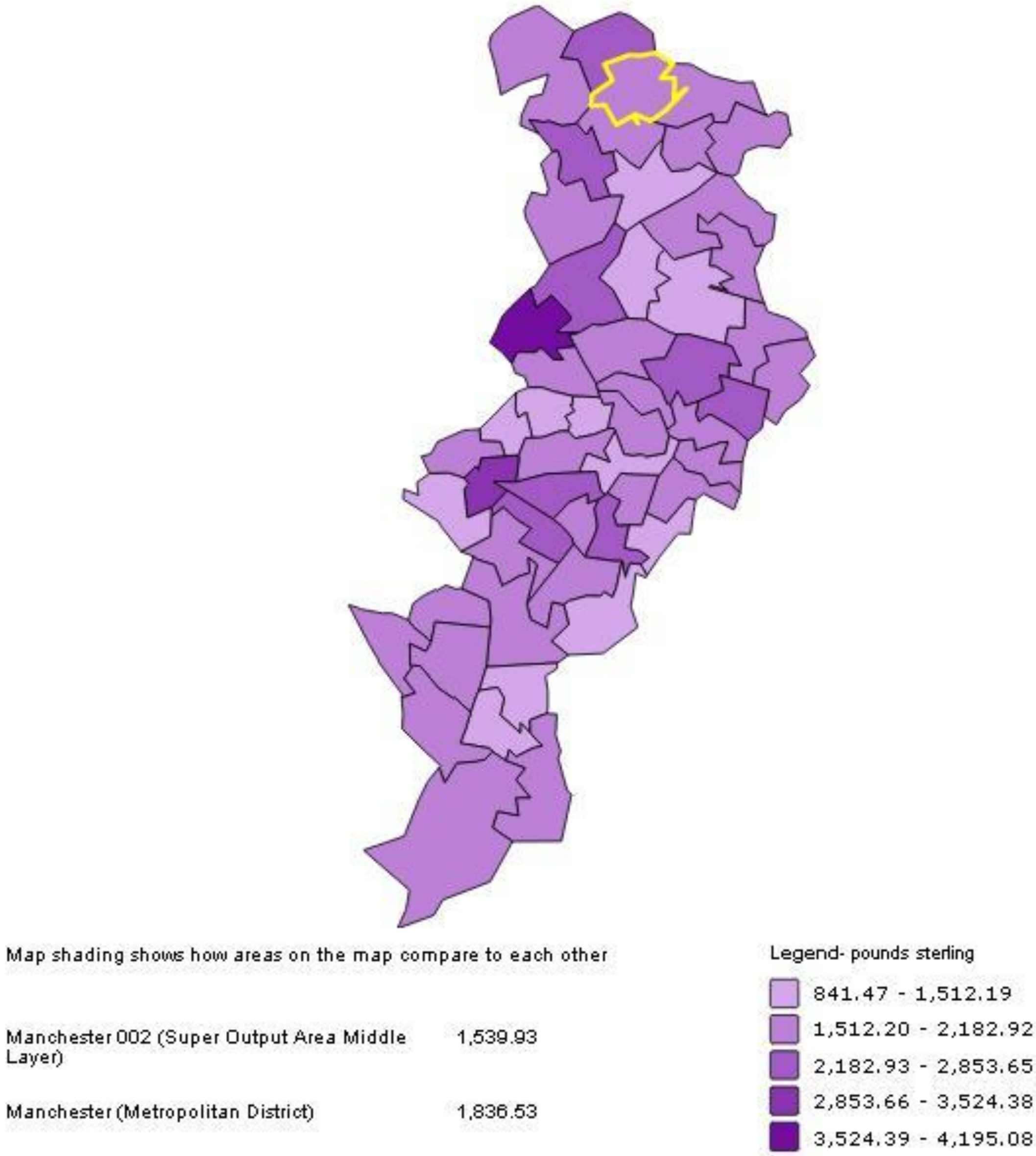

Figure 7.50: Map Showing Personal Debt

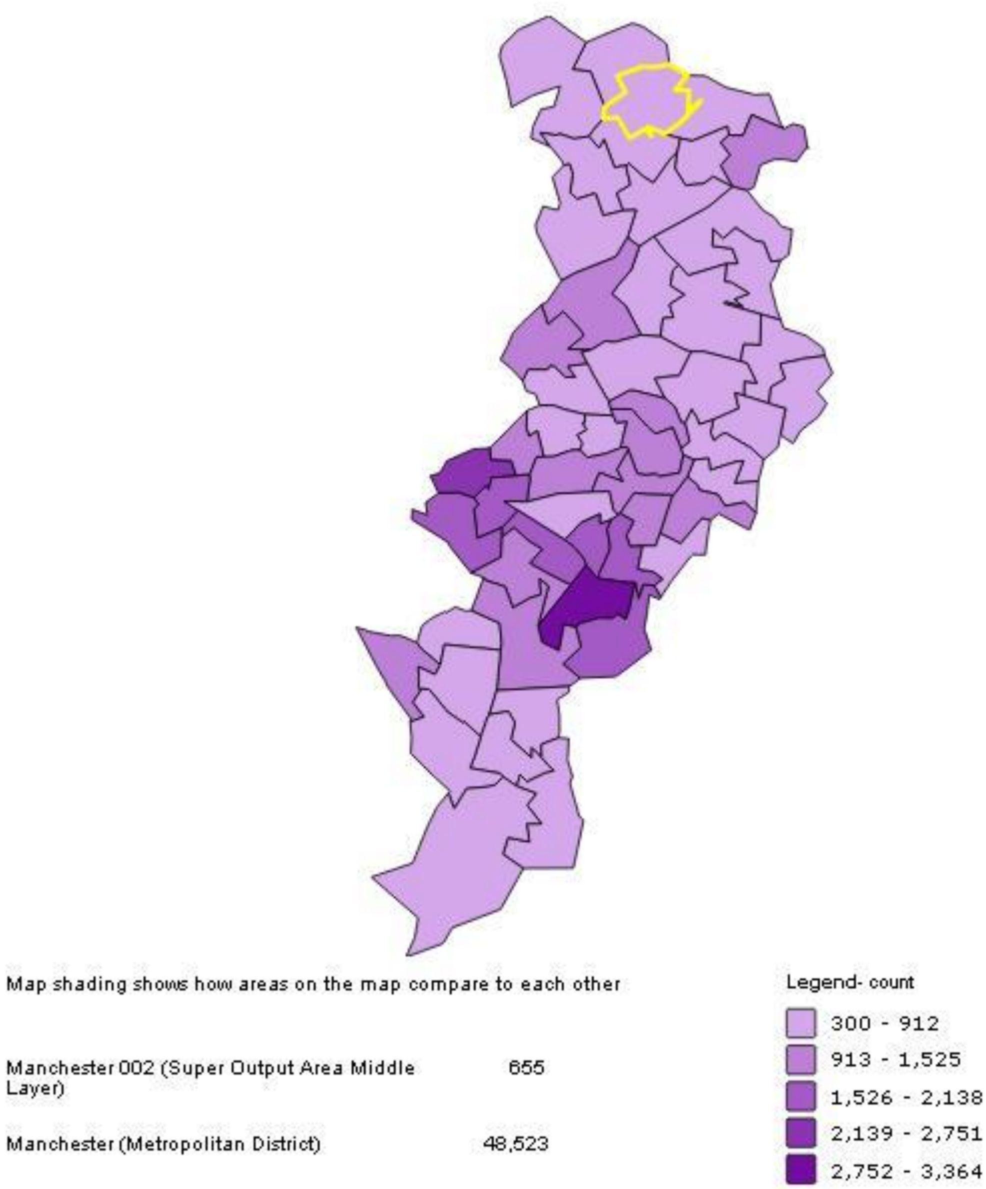

Figure 7.51: Map Showing Higher Professional

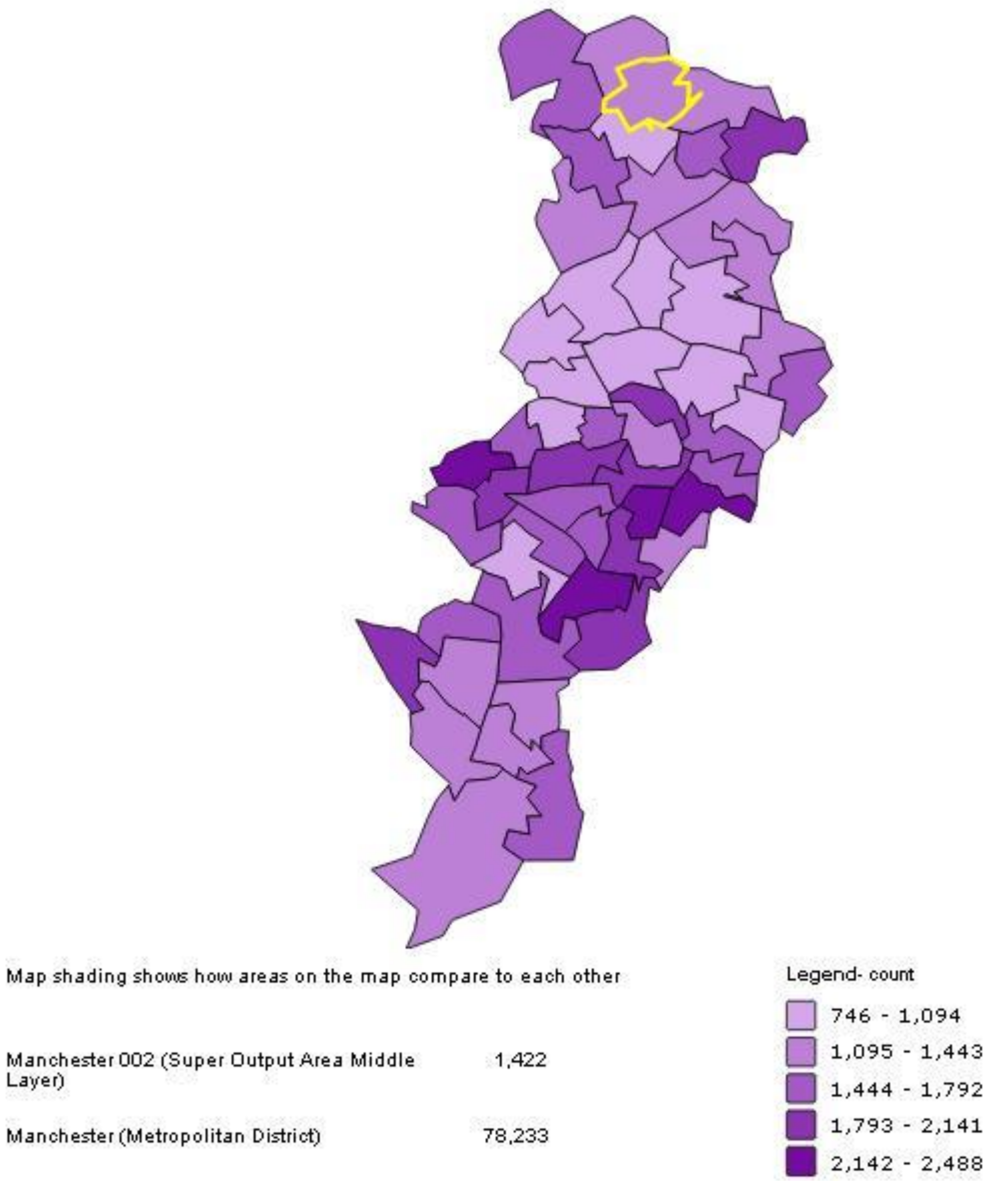

Figure 7.52: Map Showing Supervisory Professional

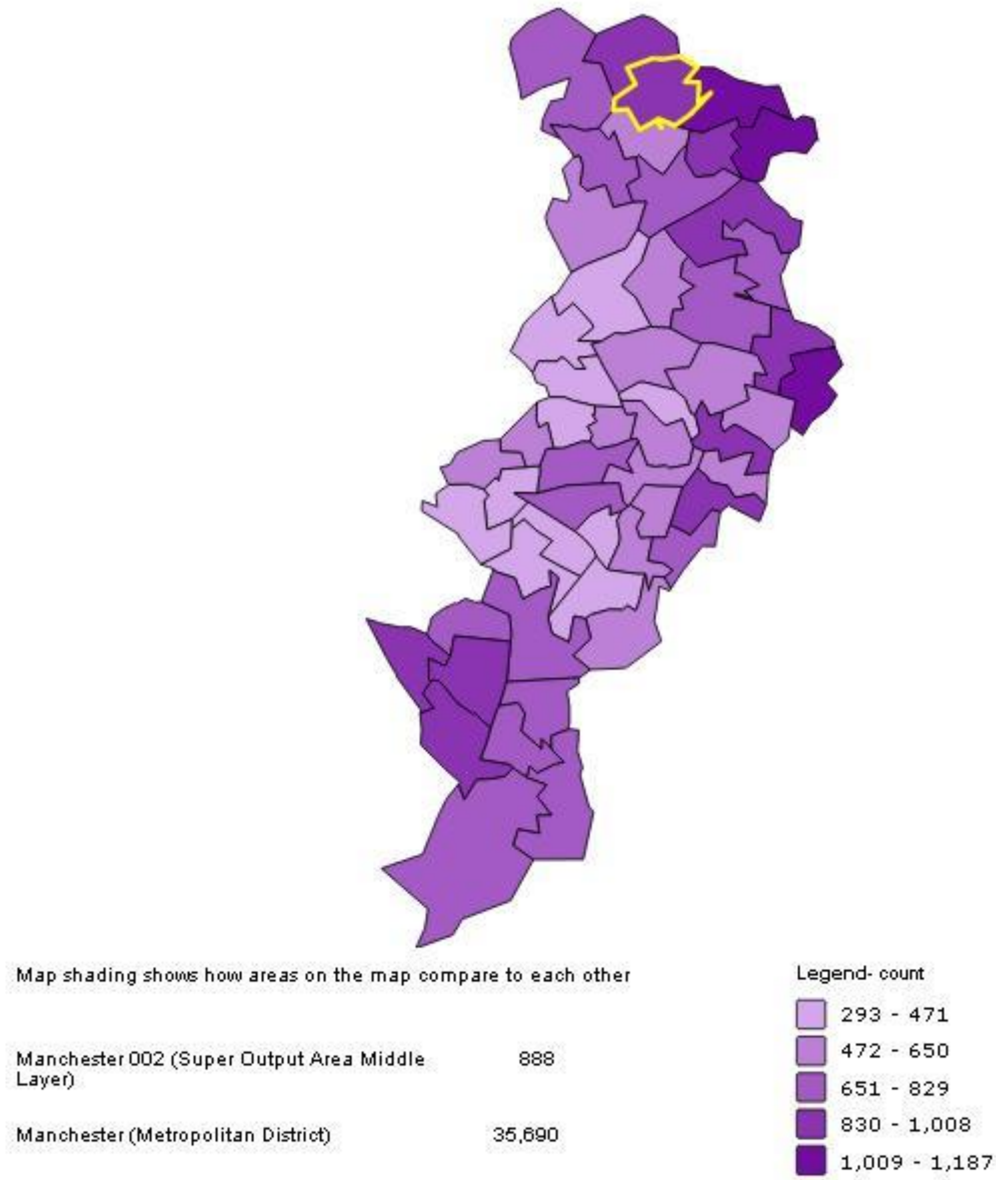

Figure 7.53: Map Showing Skilled Workers

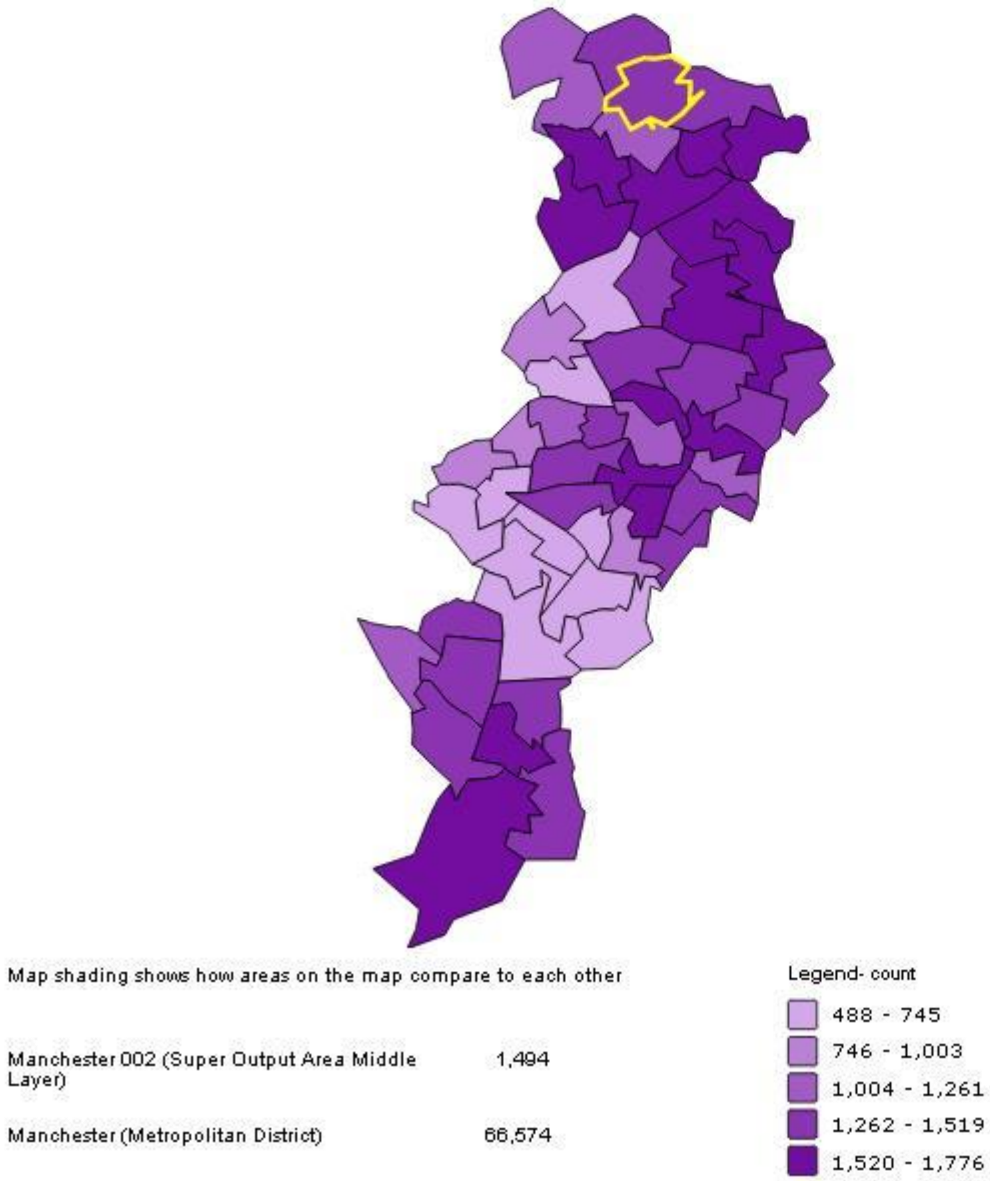

Figure 7.54: Map Showing Semi Skilled Workers

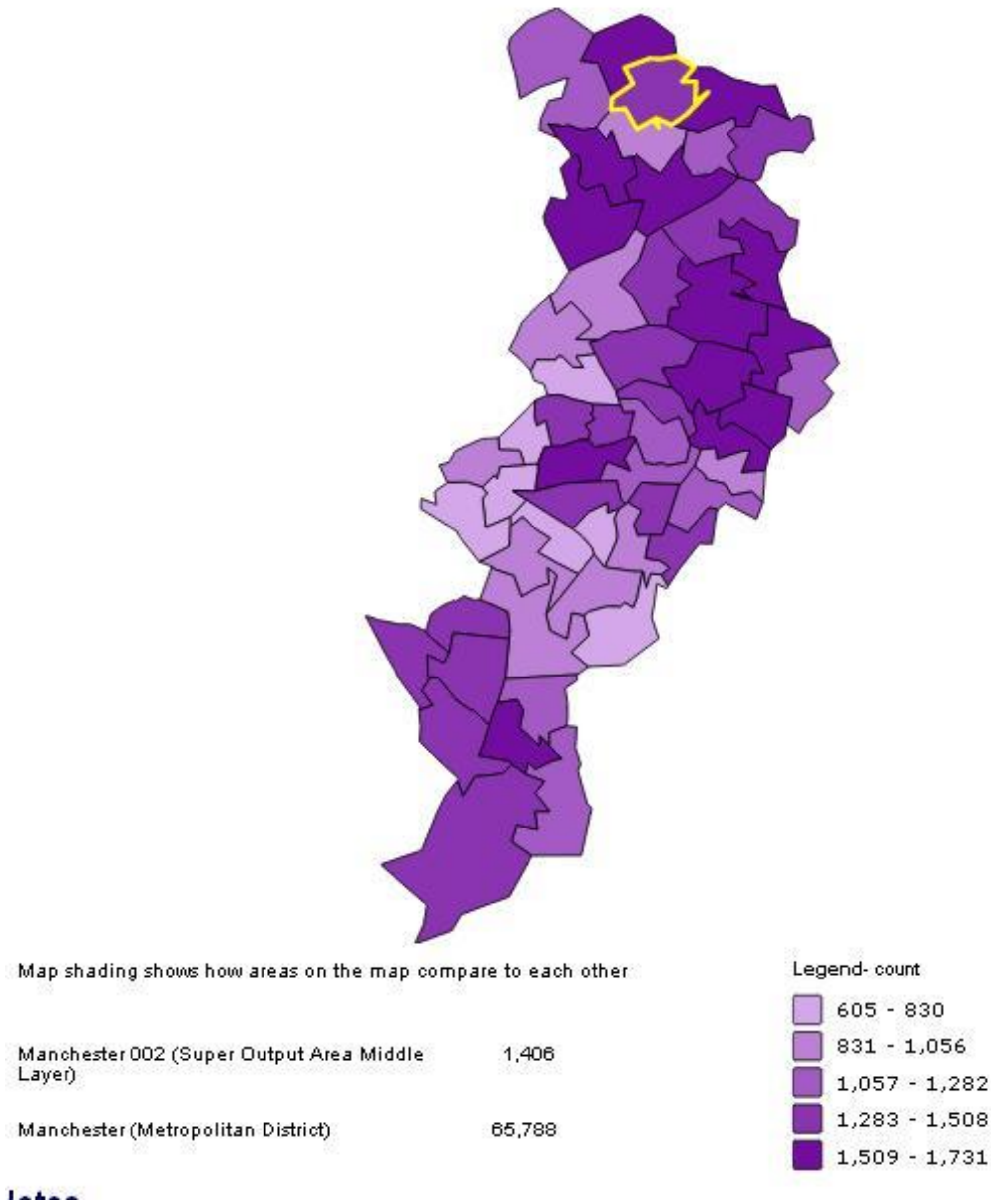

Figure 7.55: Map Showing People Living on State Benefit

- M020 Neighbourhood Thematic Maps

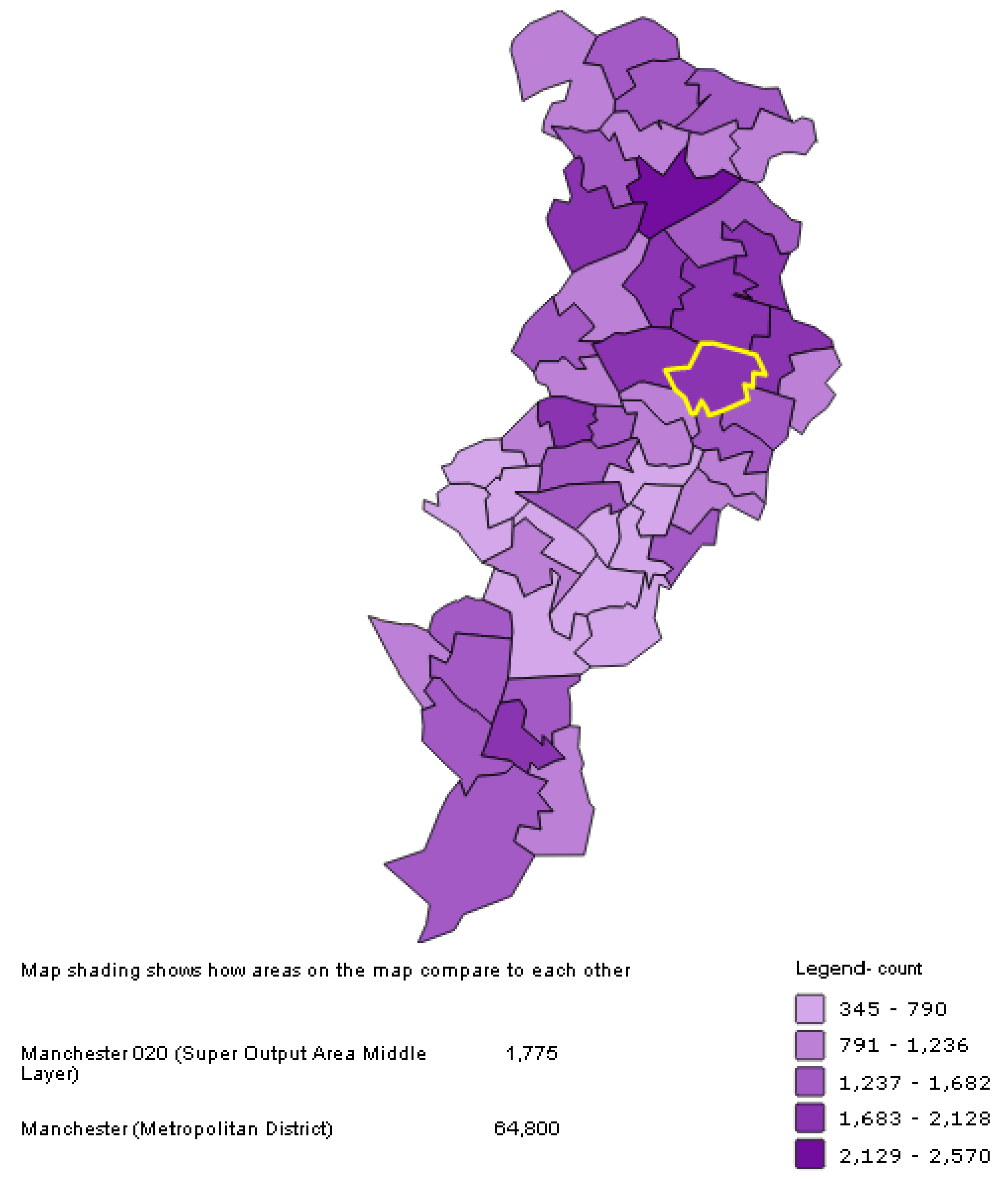

Figure: 7.57: Map Showing People Living on Benefit (Economically Deprived People)

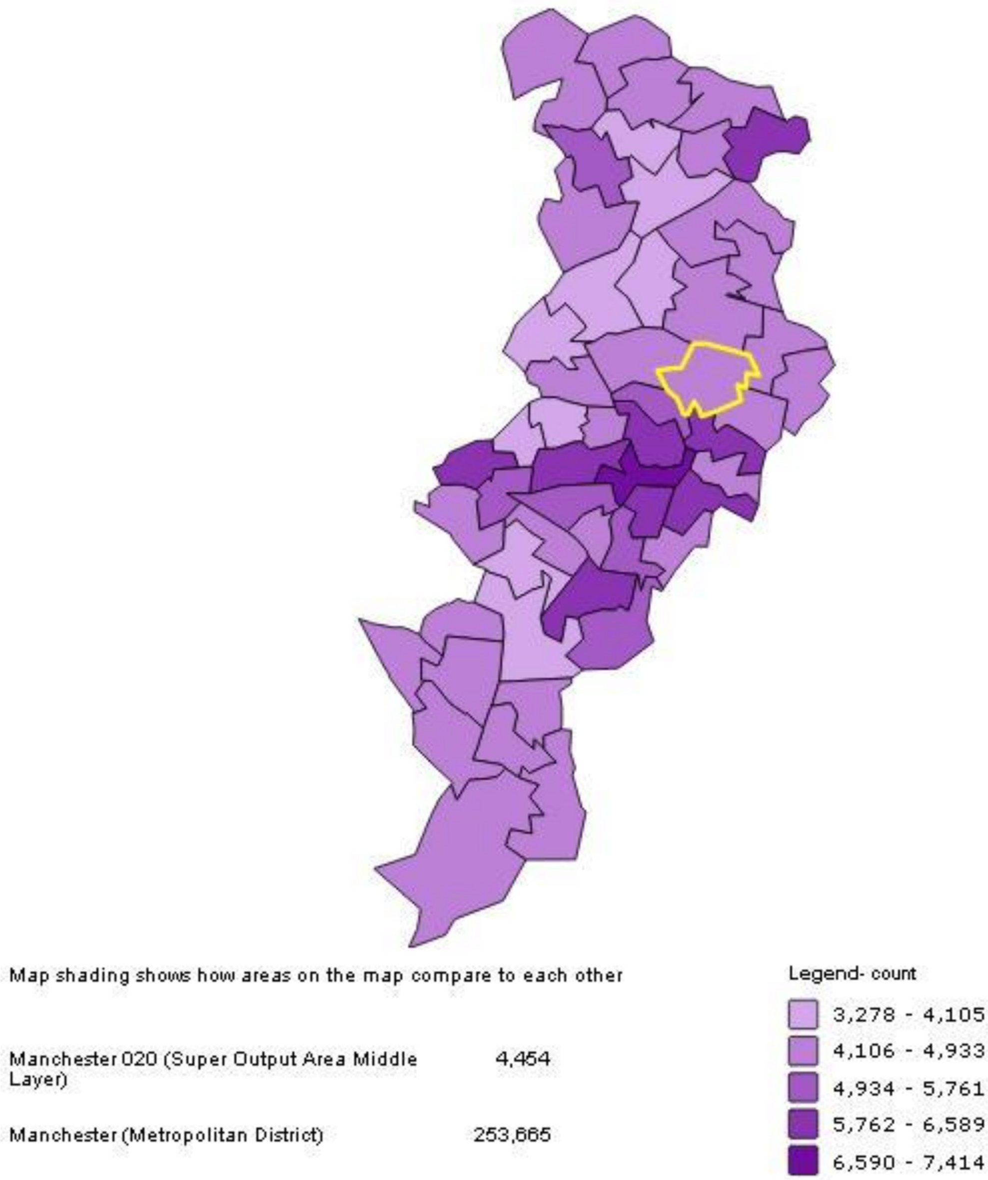

Figure 7.58: Map Showing People with Good Health

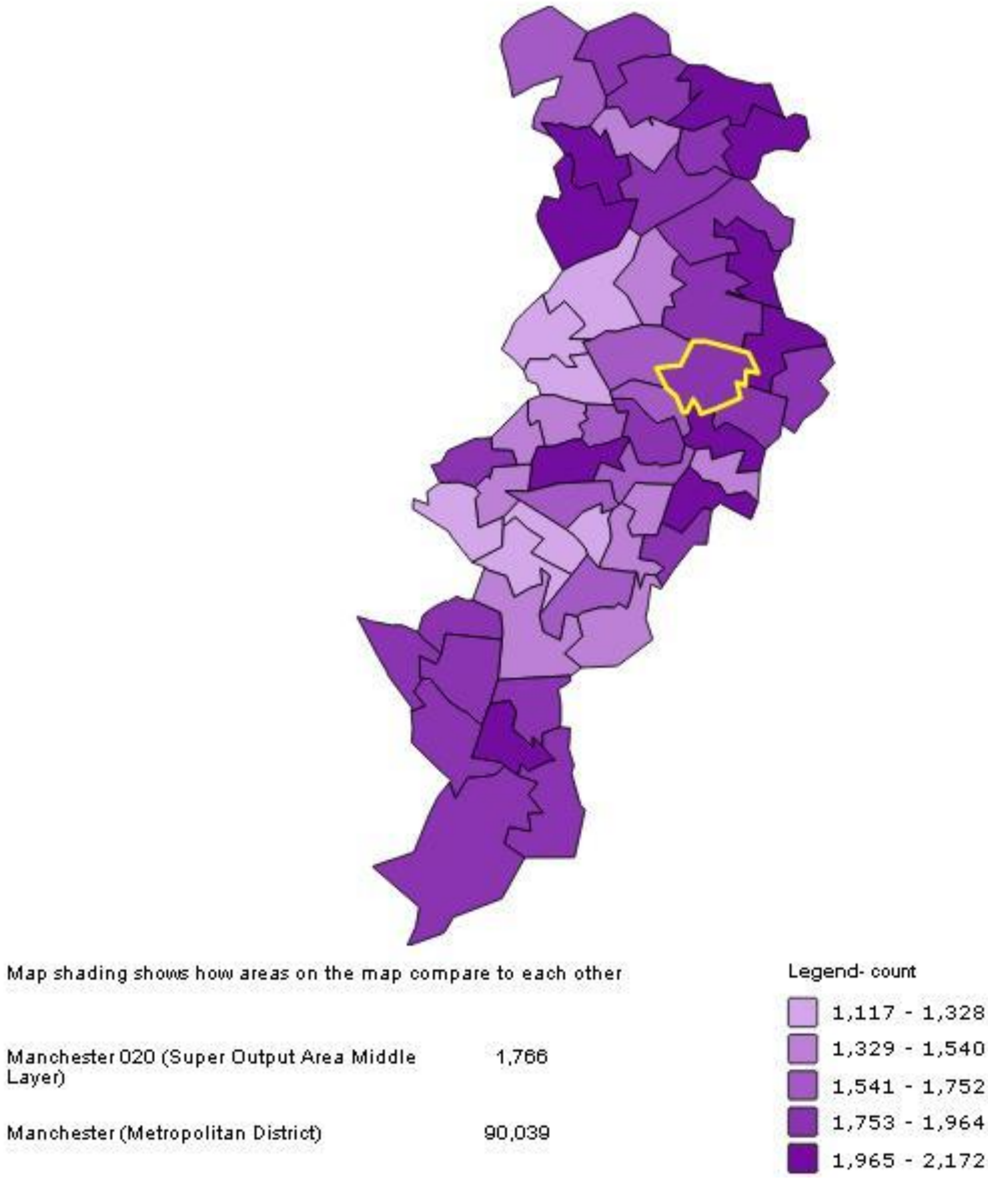

Figure 7.59: Map Showing People with Fairly Good Health

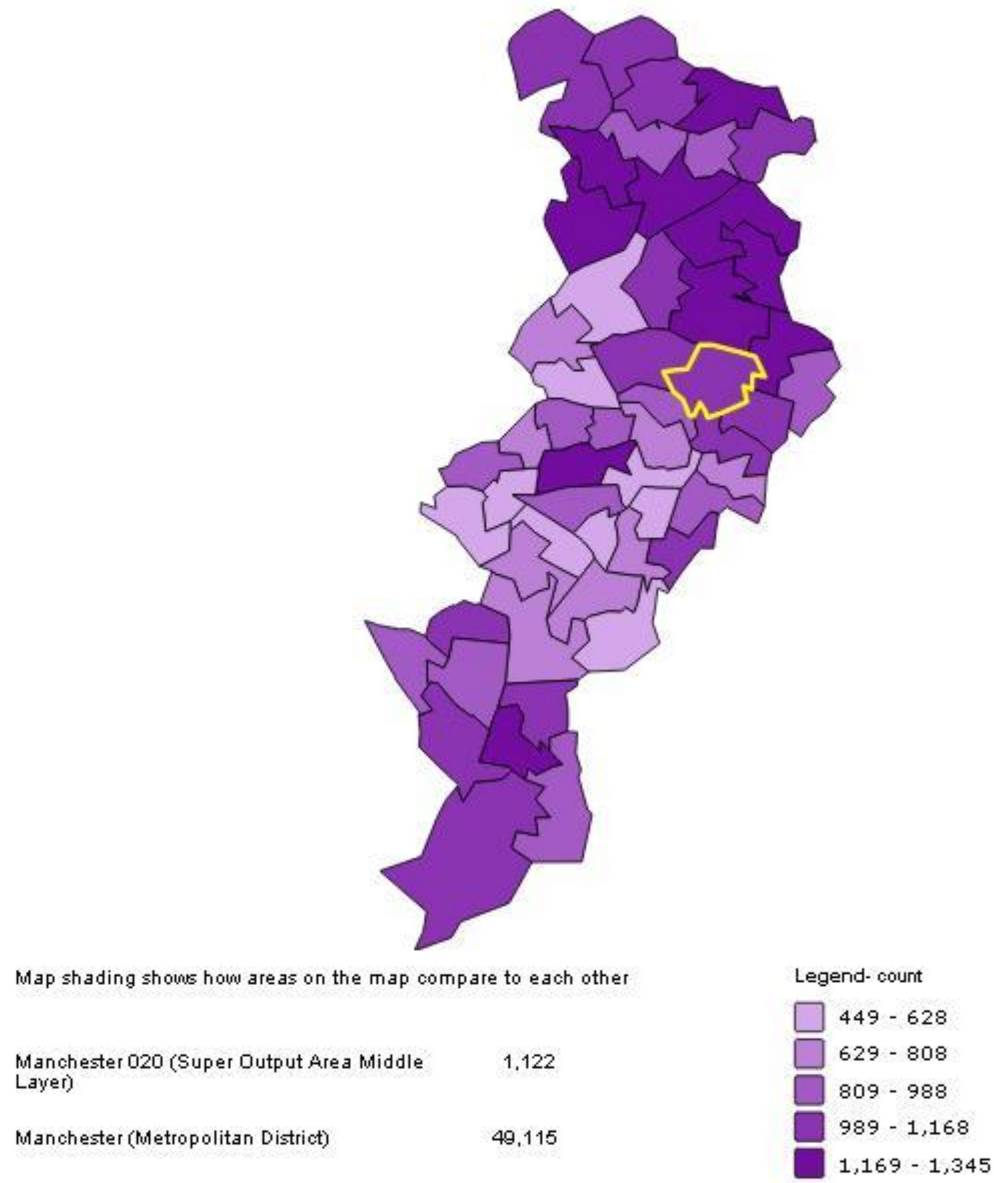

Figure 7.60: Map Showing People with Not Good Health

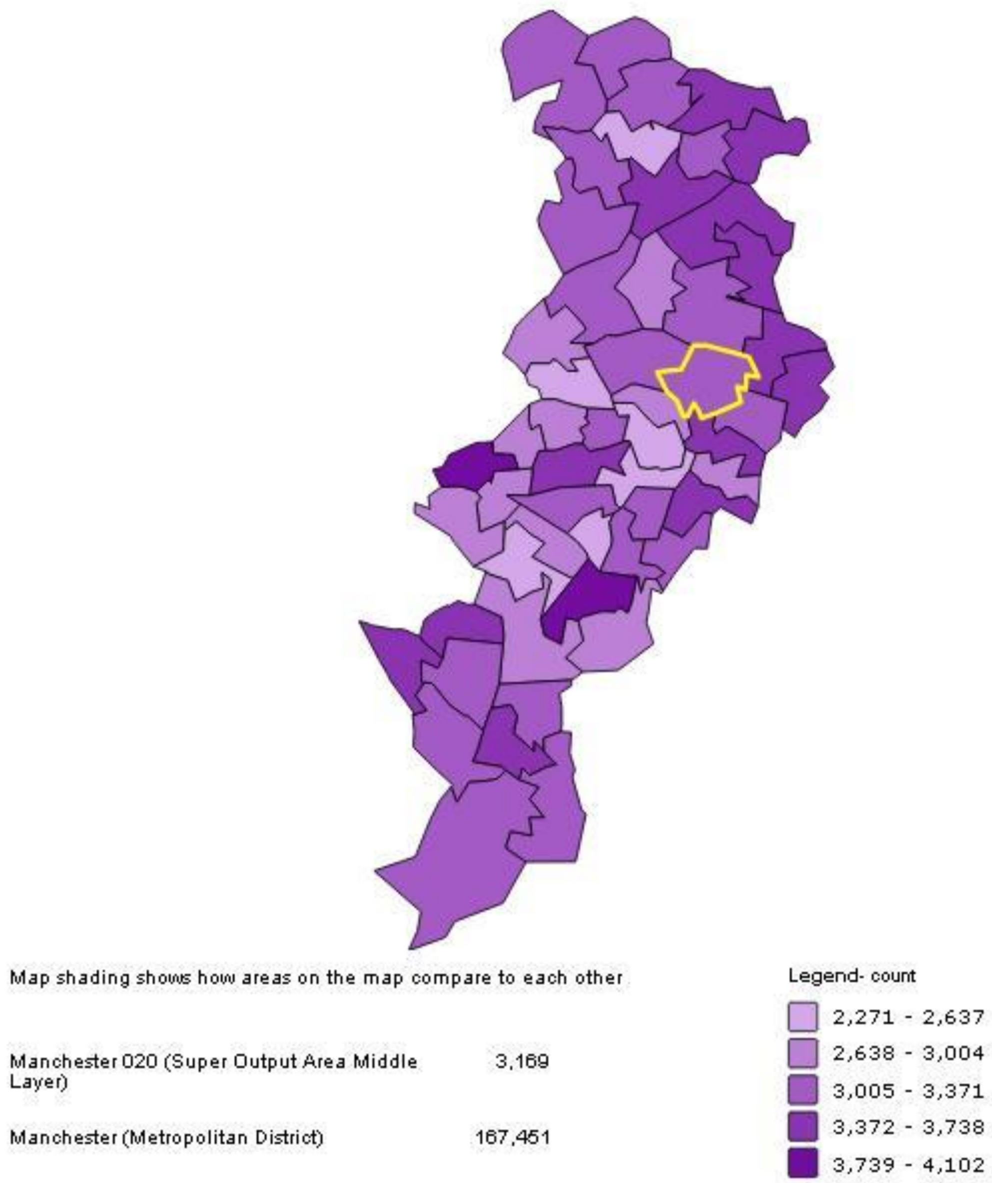

Figure 7.61: Map Showing Occupied Households

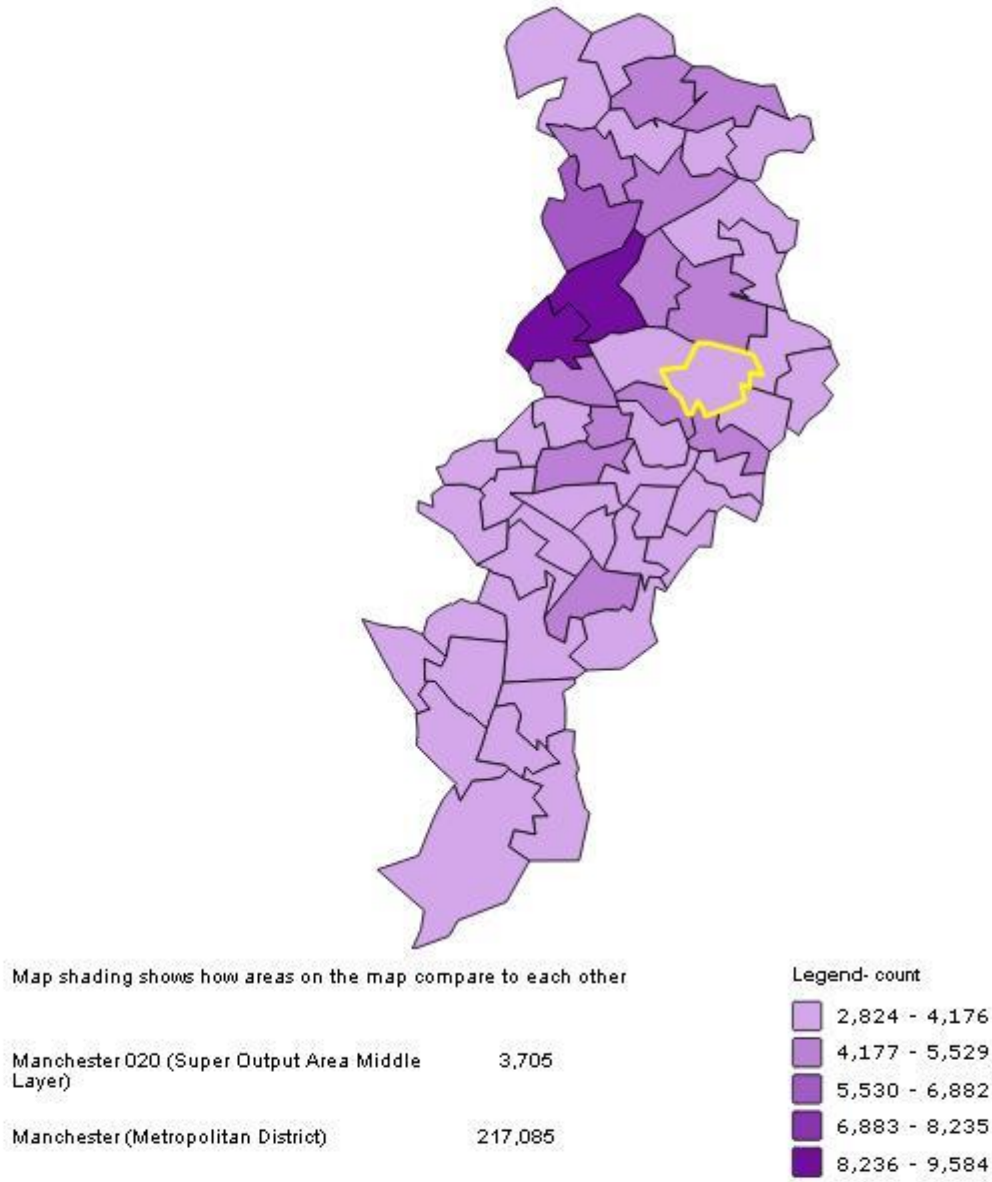

Figure 7.62: Map Showing Dwelling Stock

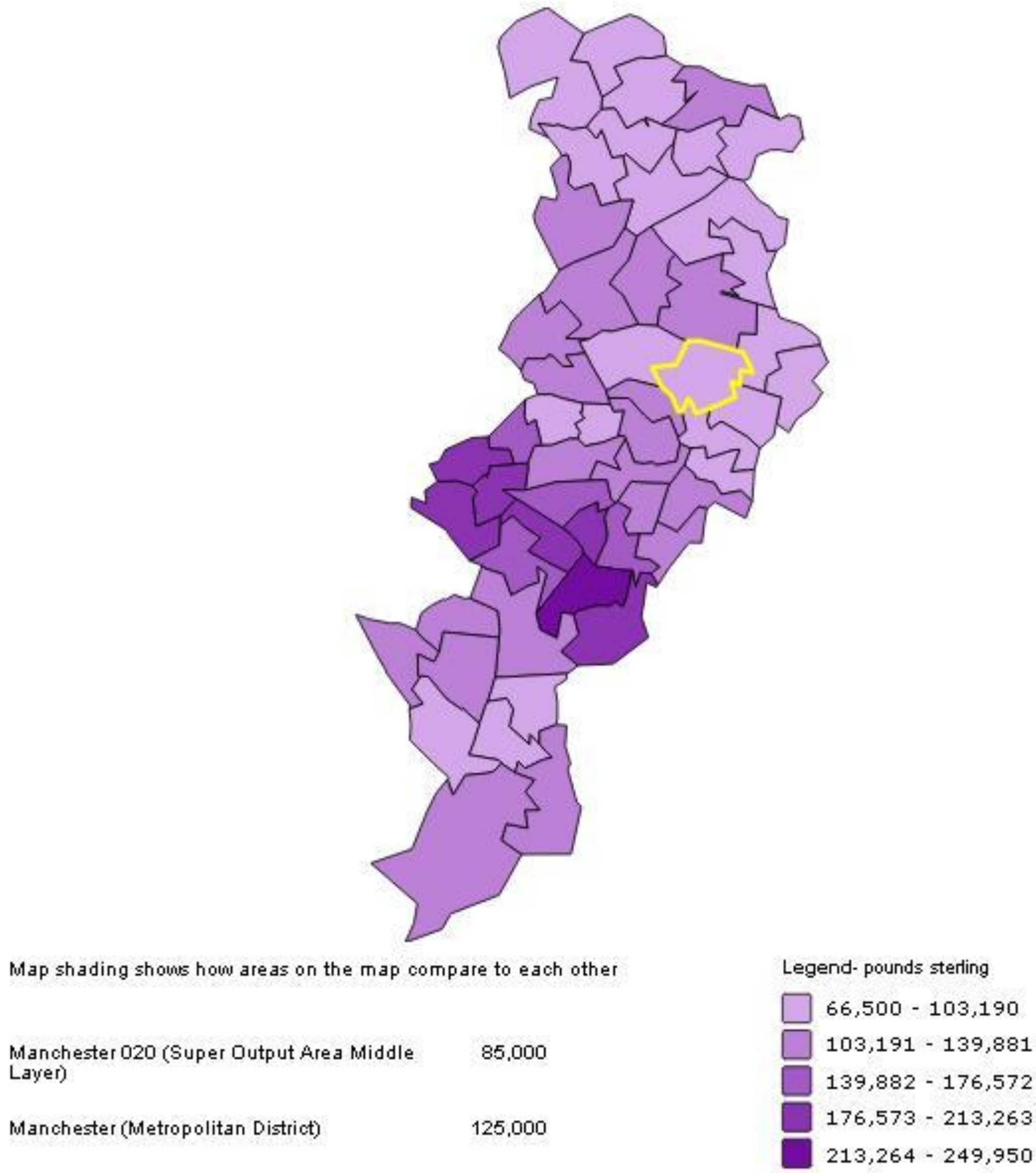

Figure 7.63: Map Showing Dwelling Price

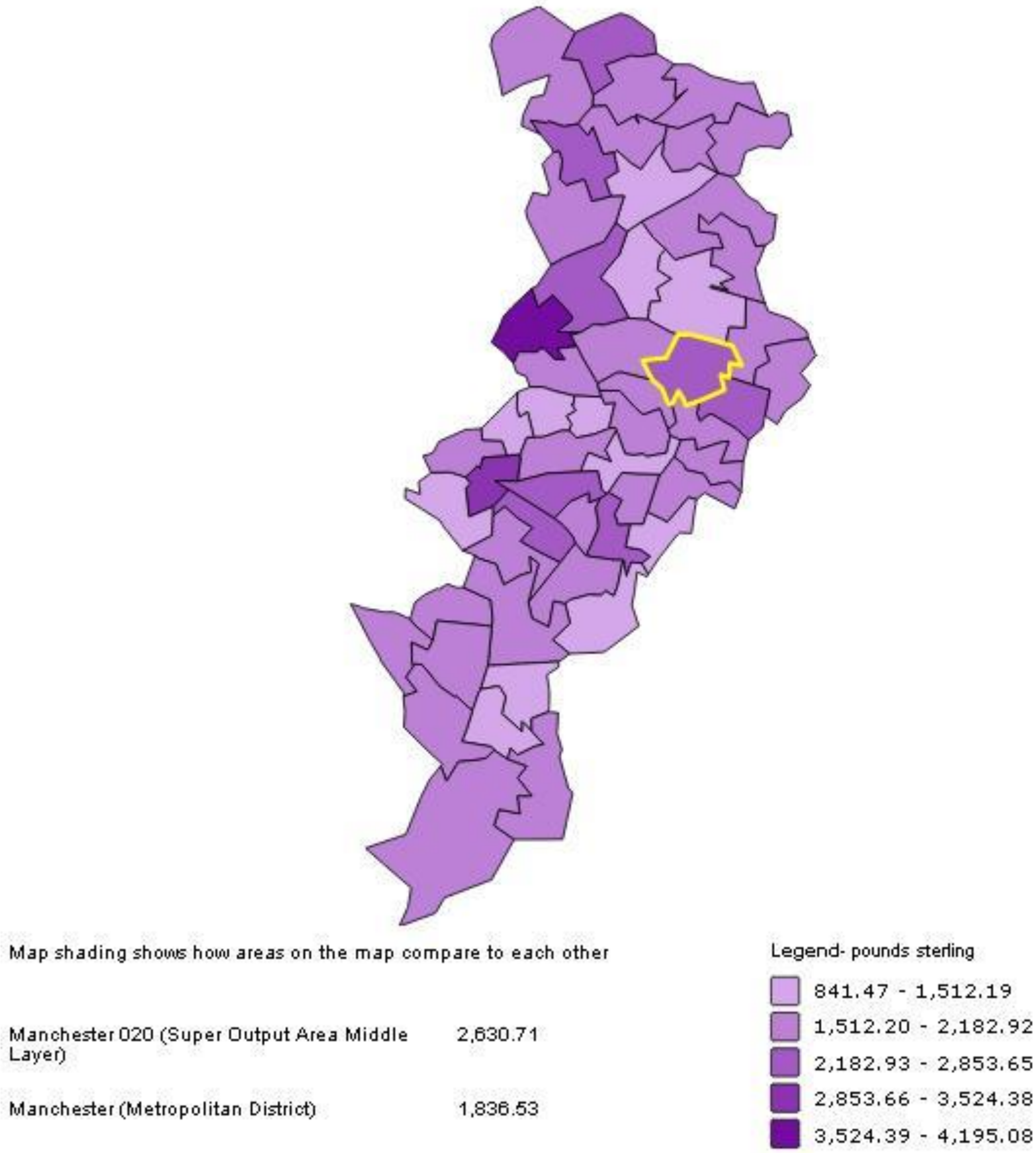

Figure 7.64: Map Showing Personal Debt

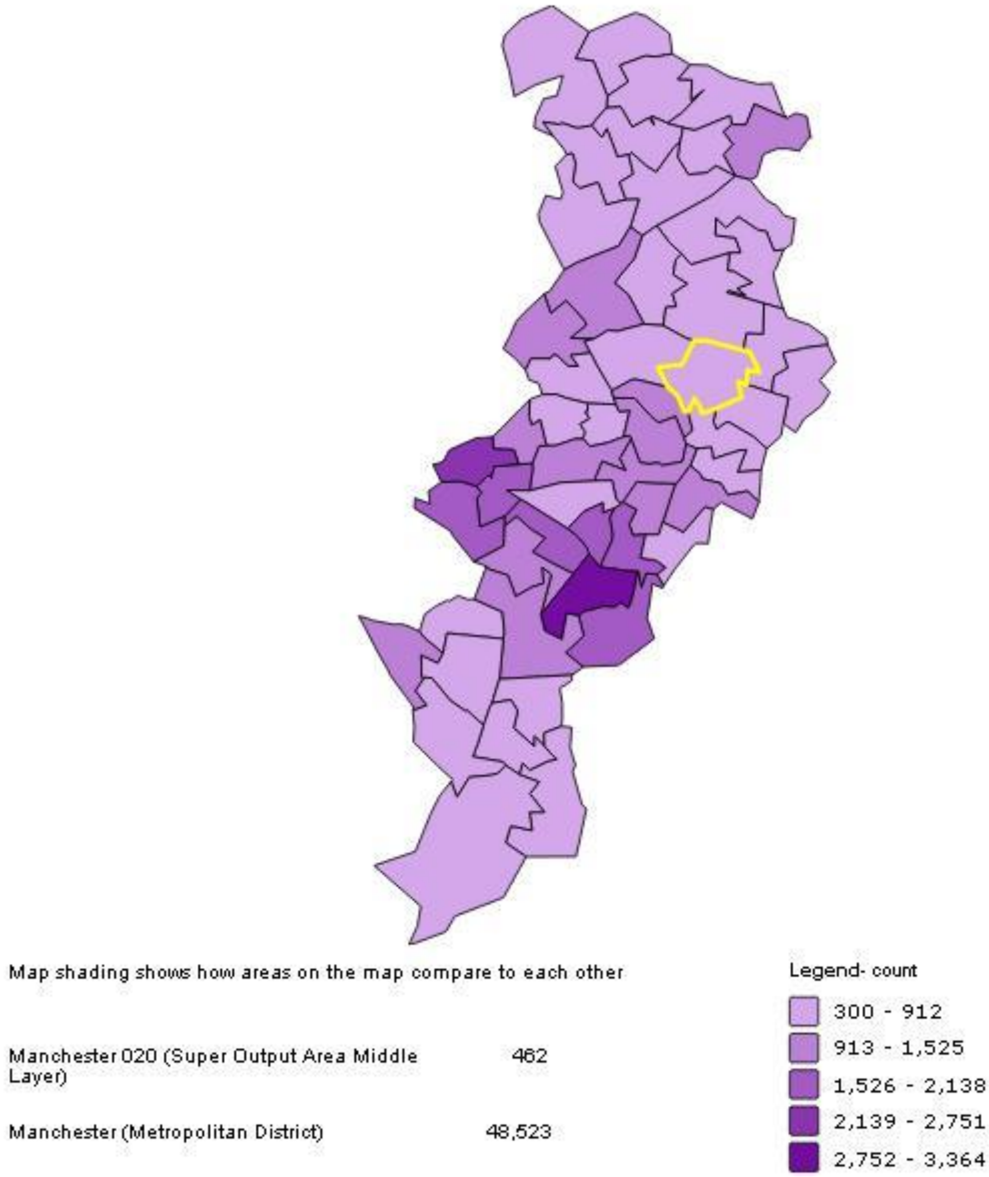

Figure 7.65: Map Showing Number of Higher Professional

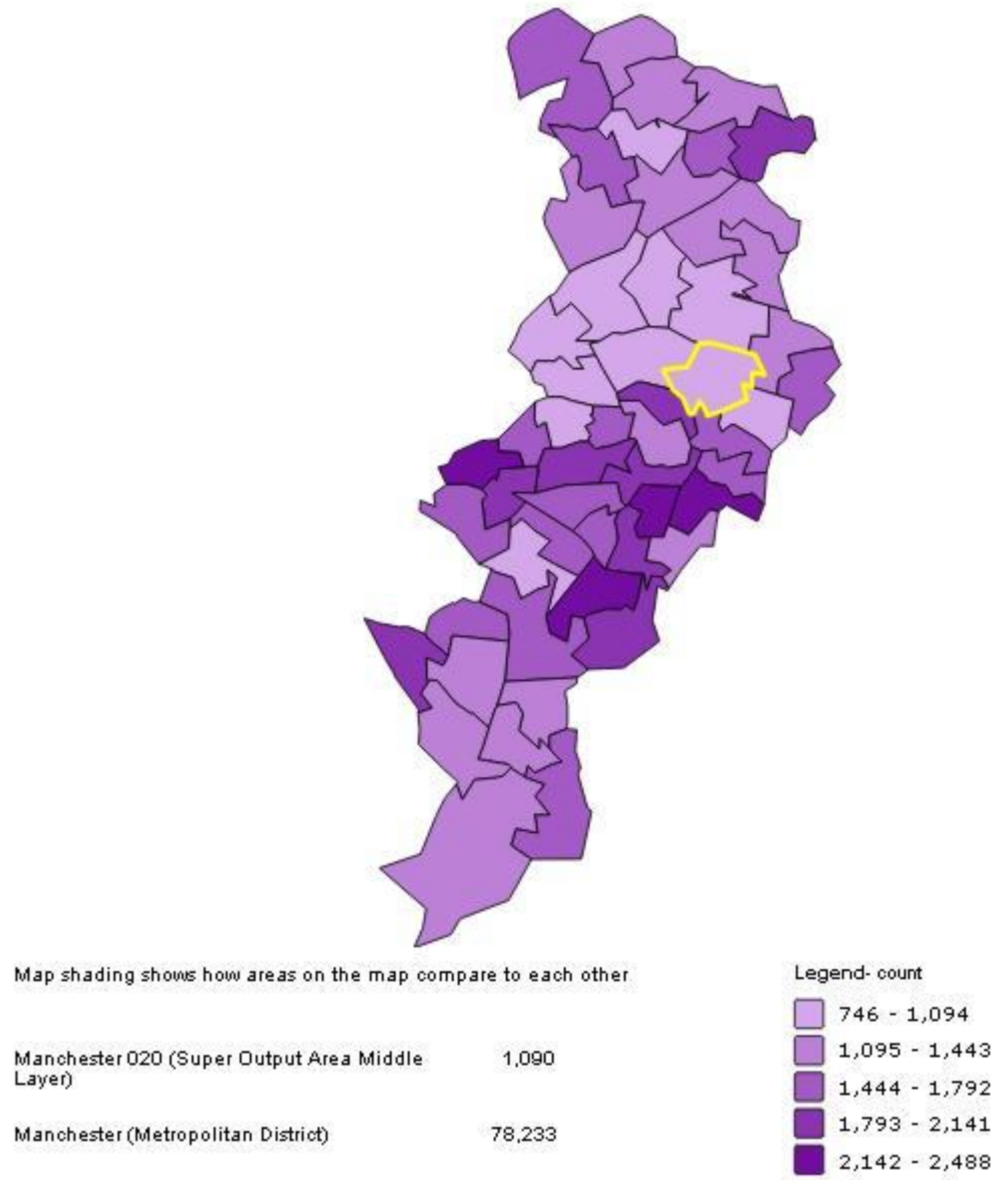

Figure 7.66: Map Showing Number of Supervisory Professional

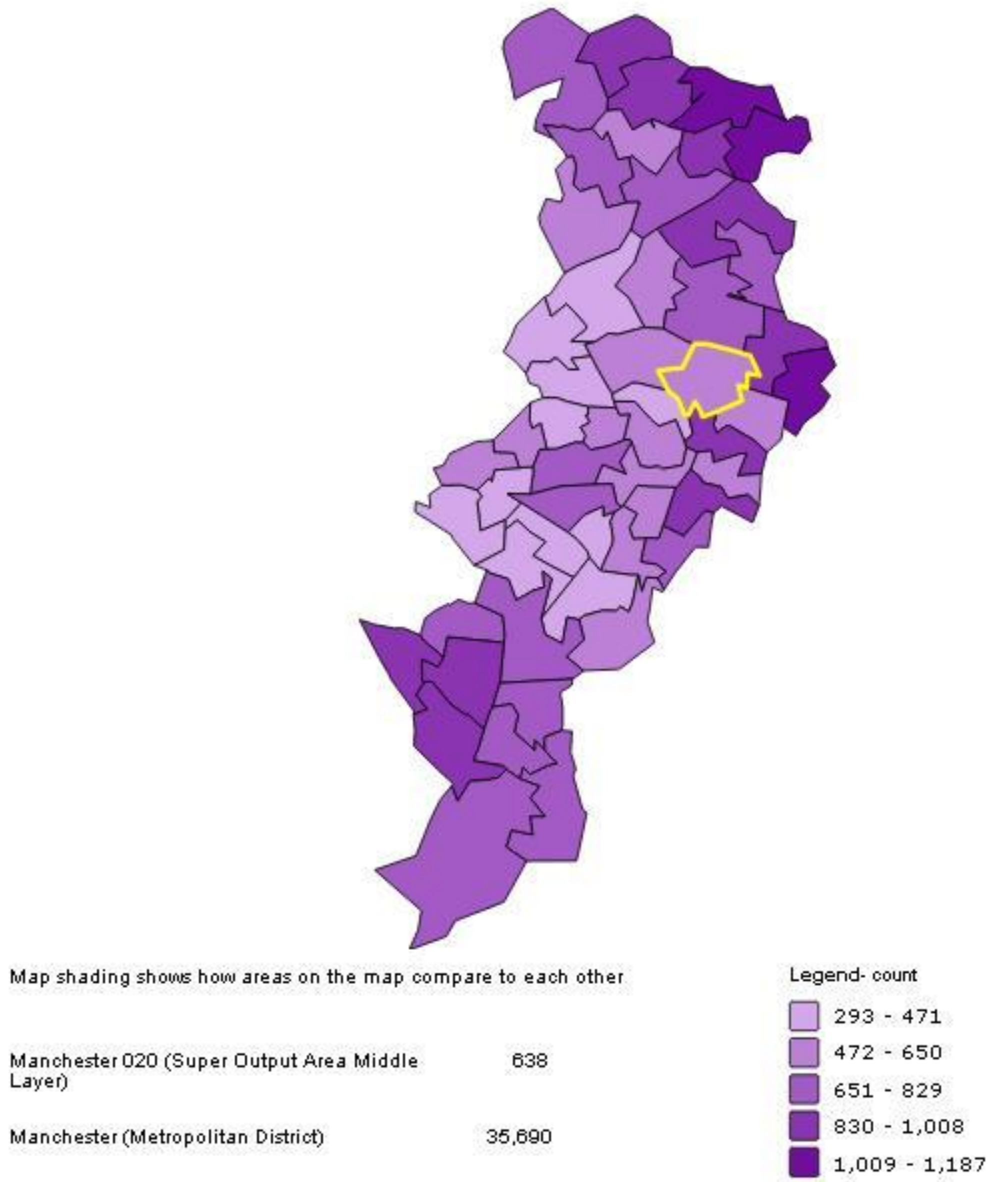

Figure 7.67: Map Showing Number of Skilled Professional

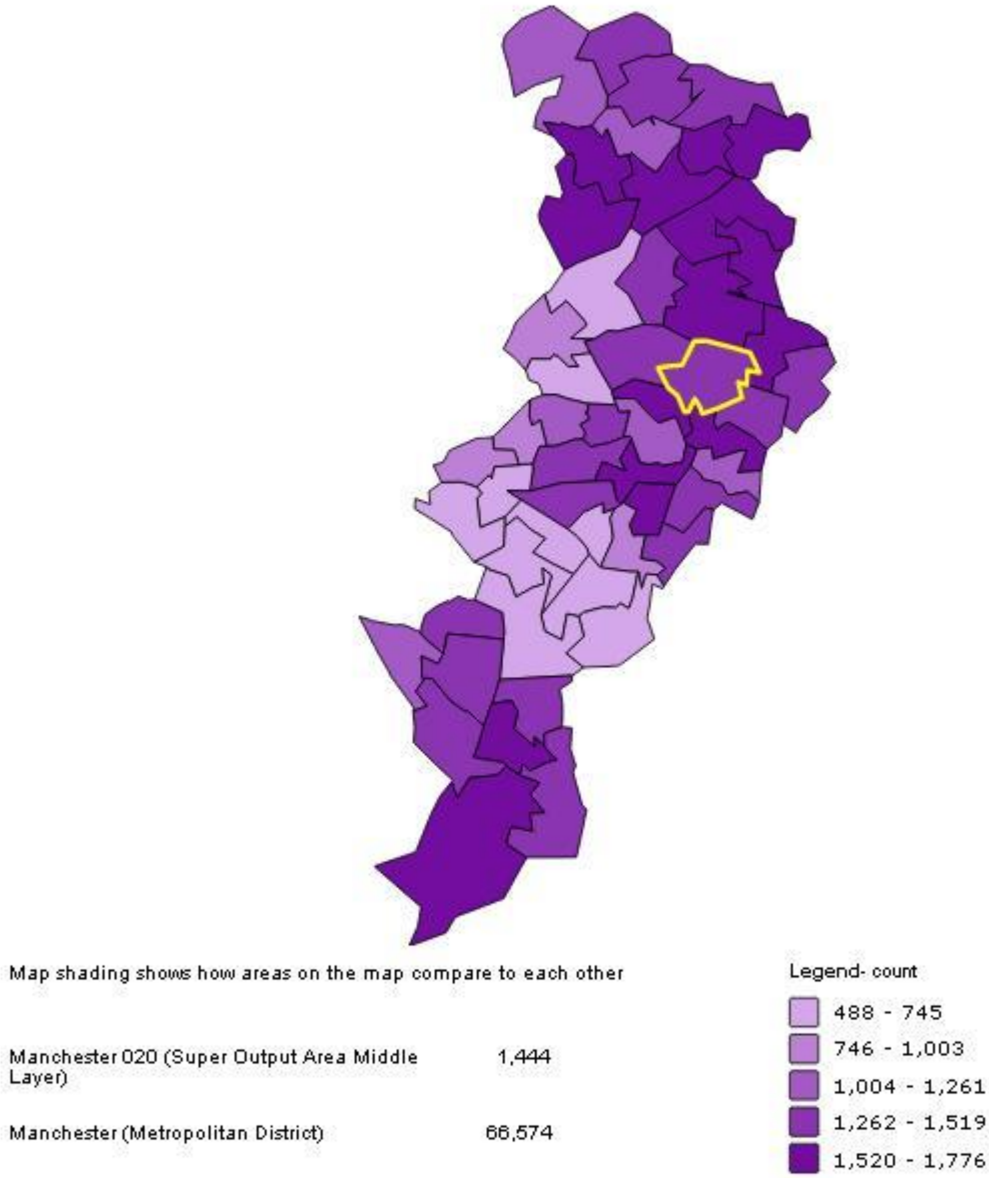

Figure 7.68: Map Showing Number of Semi Skilled Professional

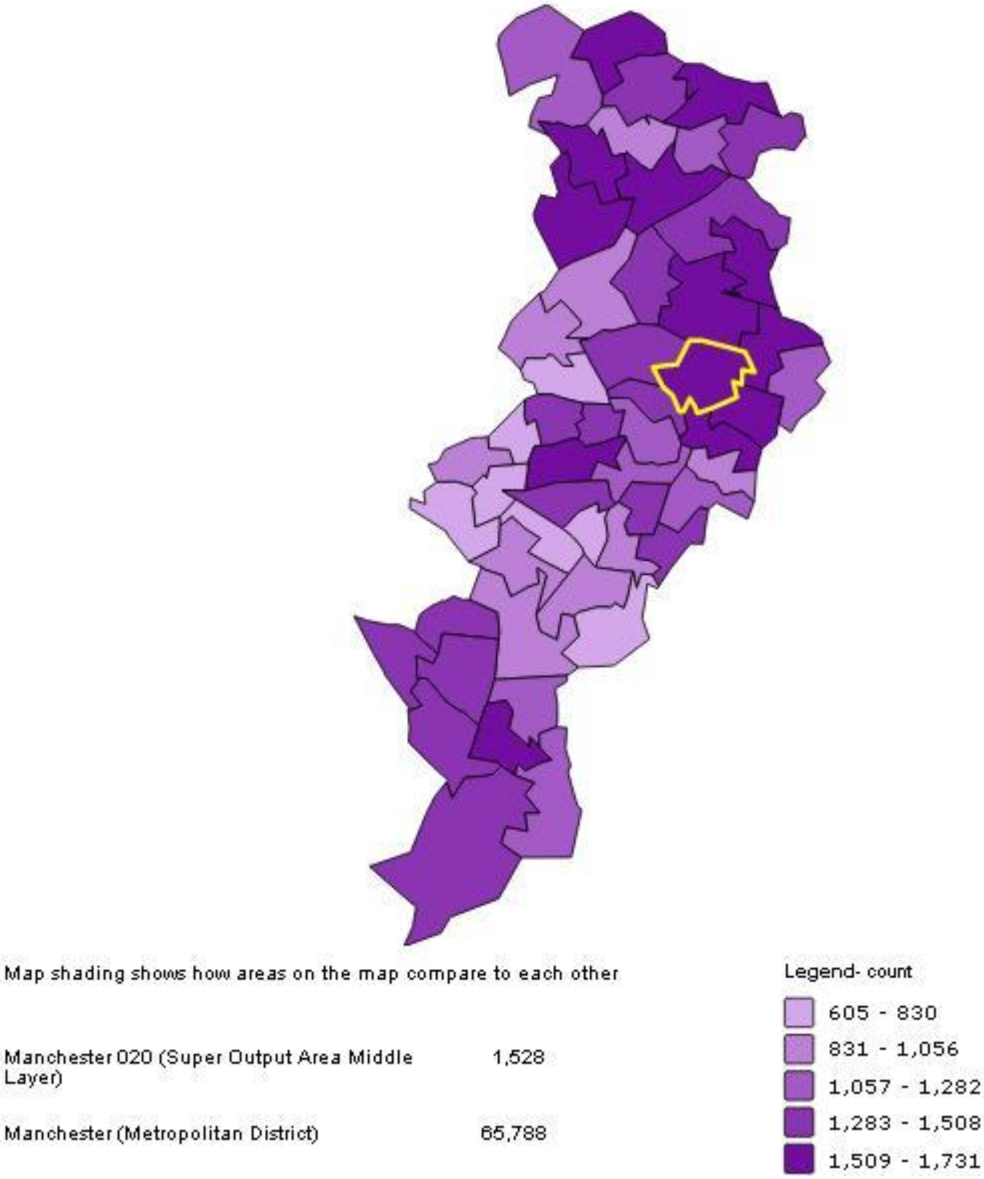

Figure 7.69: Map Showing Number of People Living on State Benefit

- M050 Neighbourhood Thematic Maps

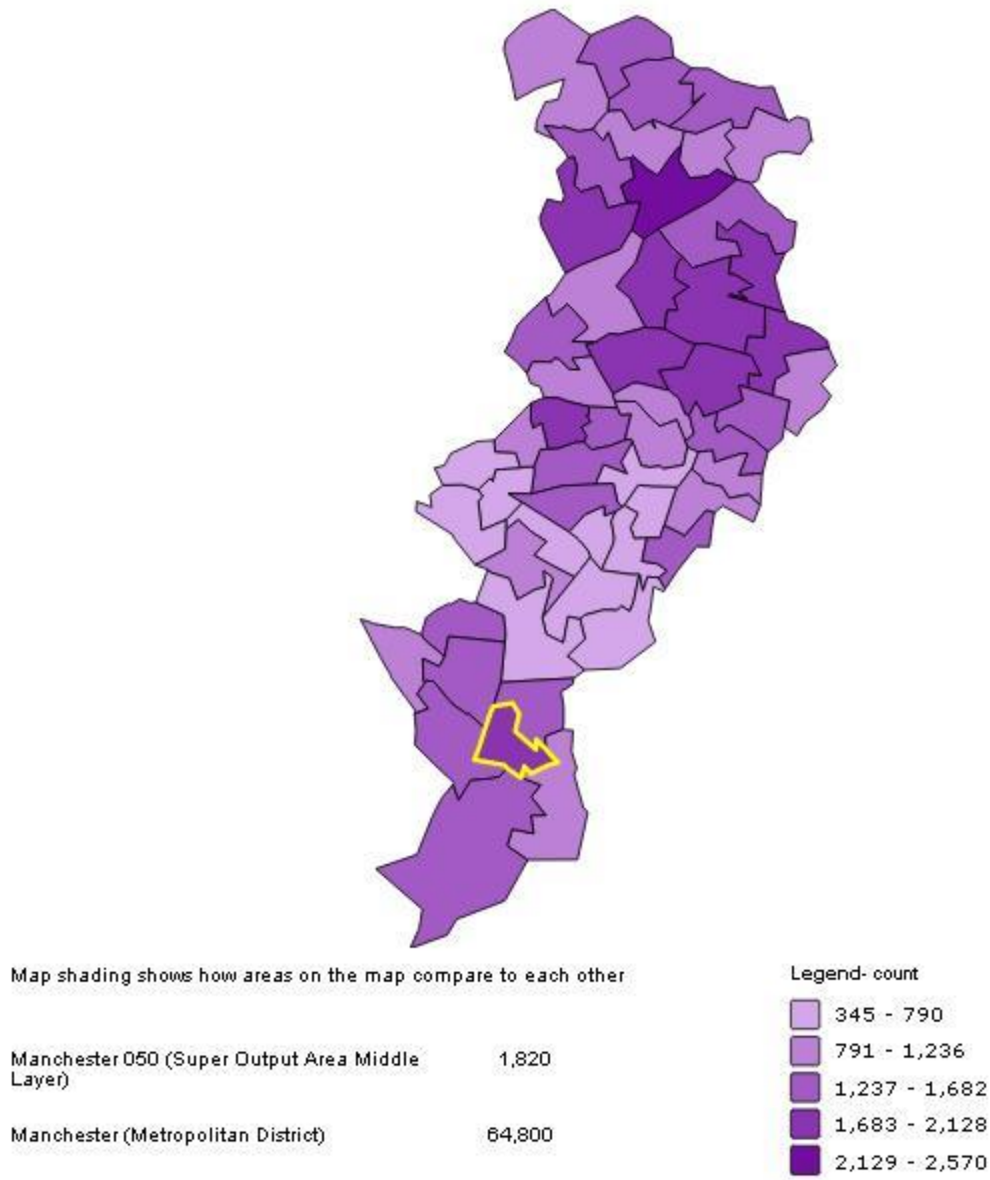

Figure: 7.71: Map Showing People Living on Benefit (Economically Deprived People)

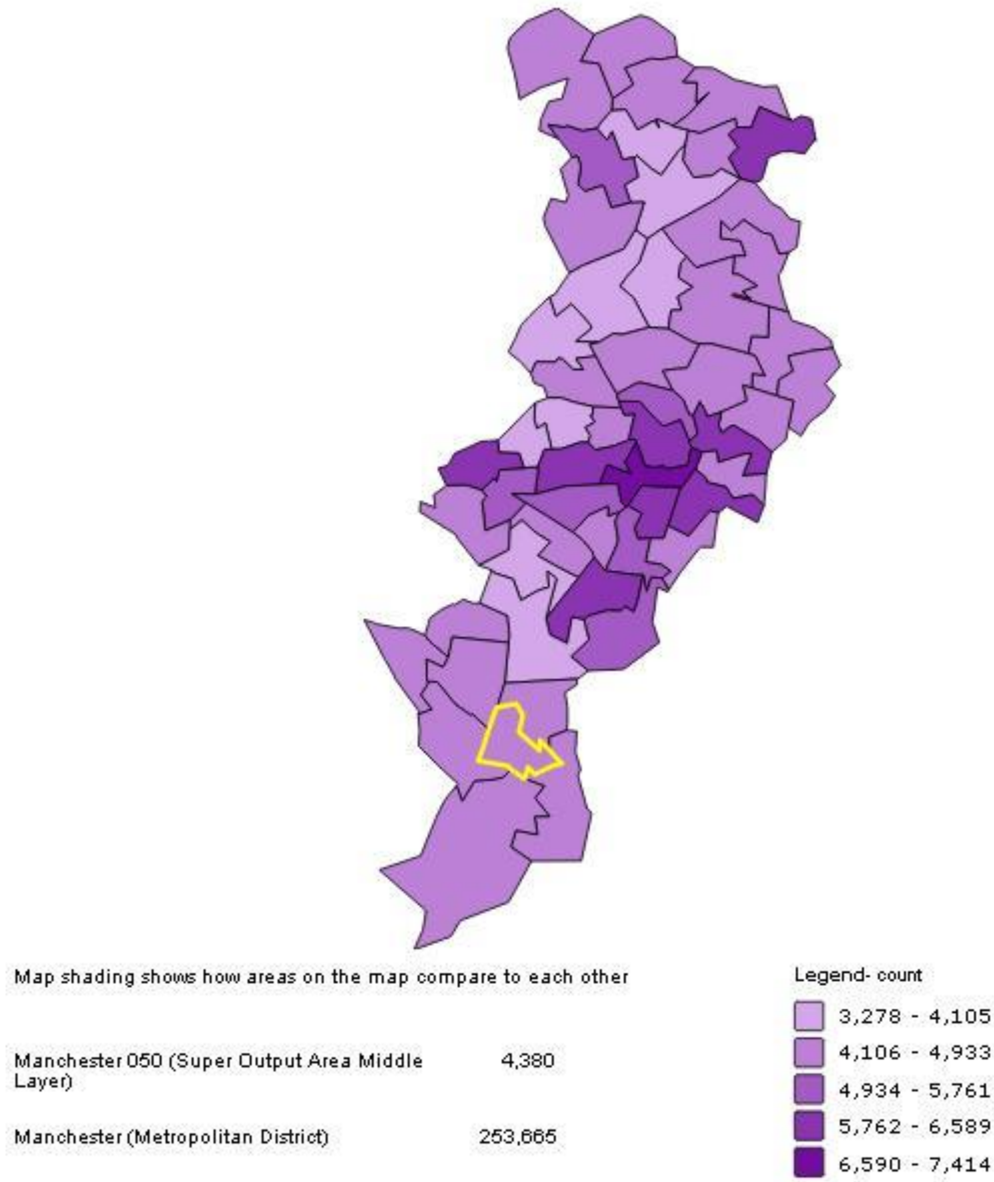

Figure 7.72: Map Showing People with Good Health

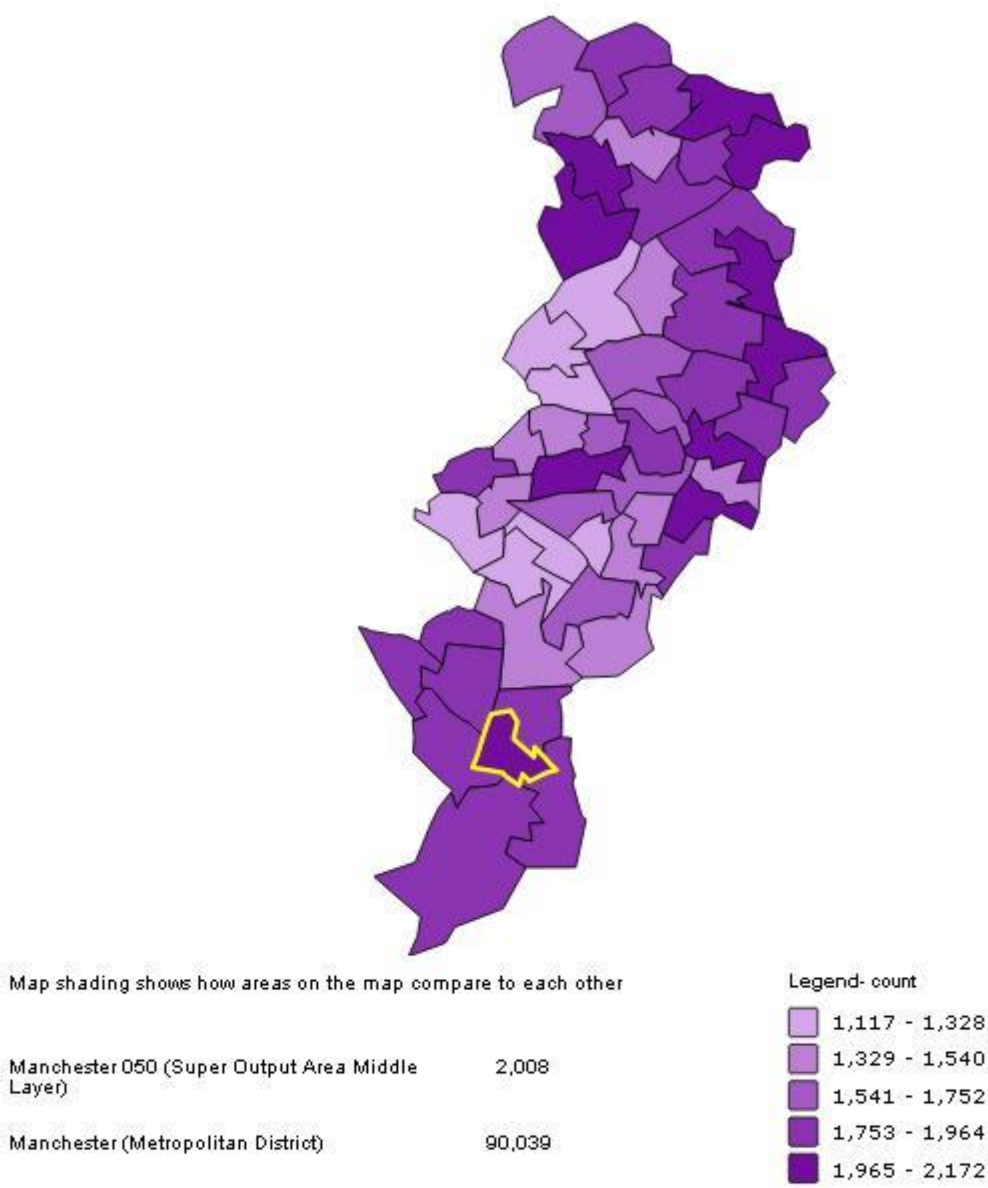

Figure 7.73: Map Showing People with Fairly Good Health

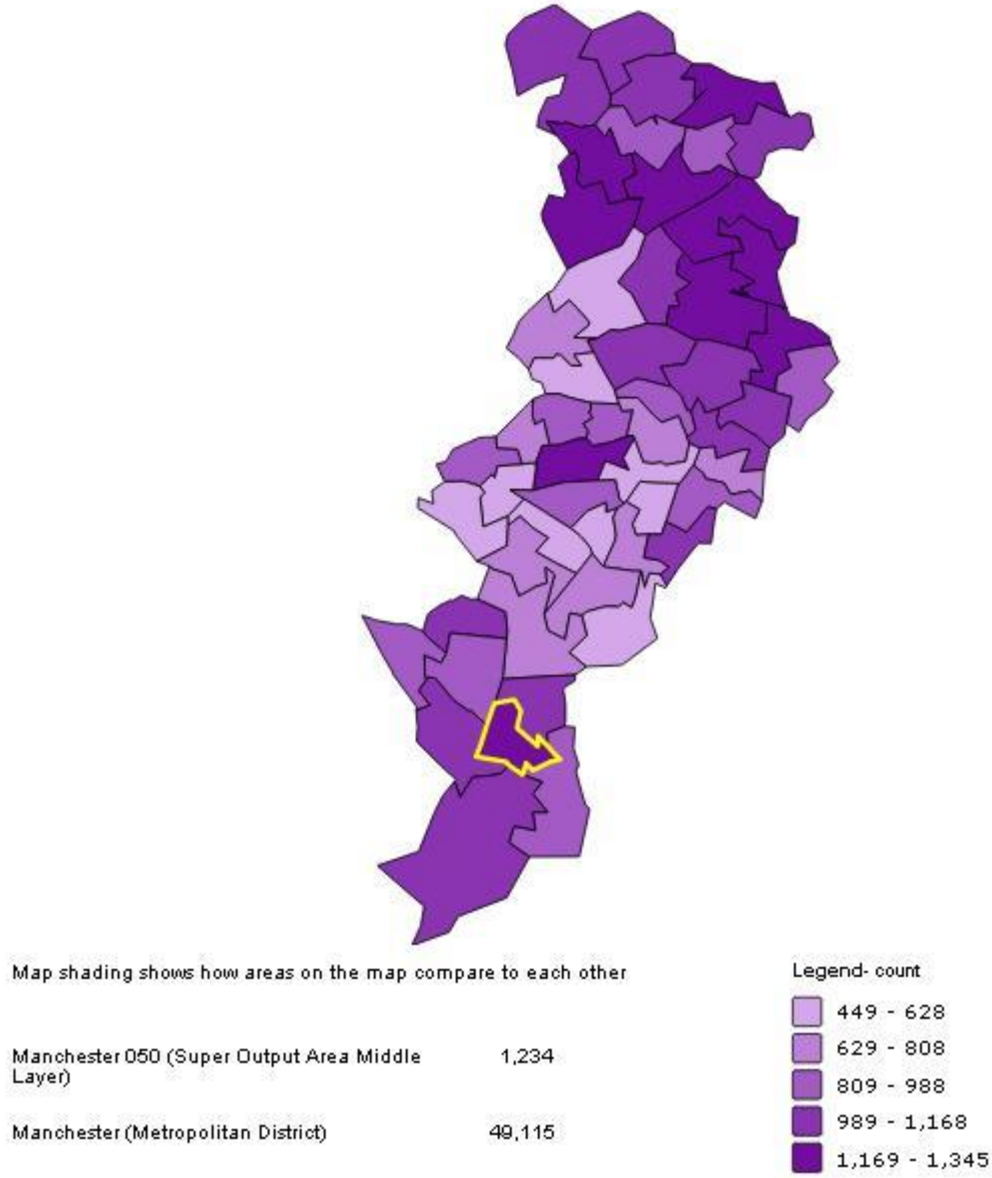

Figure 7.74: Map Showing People with Not Good Health

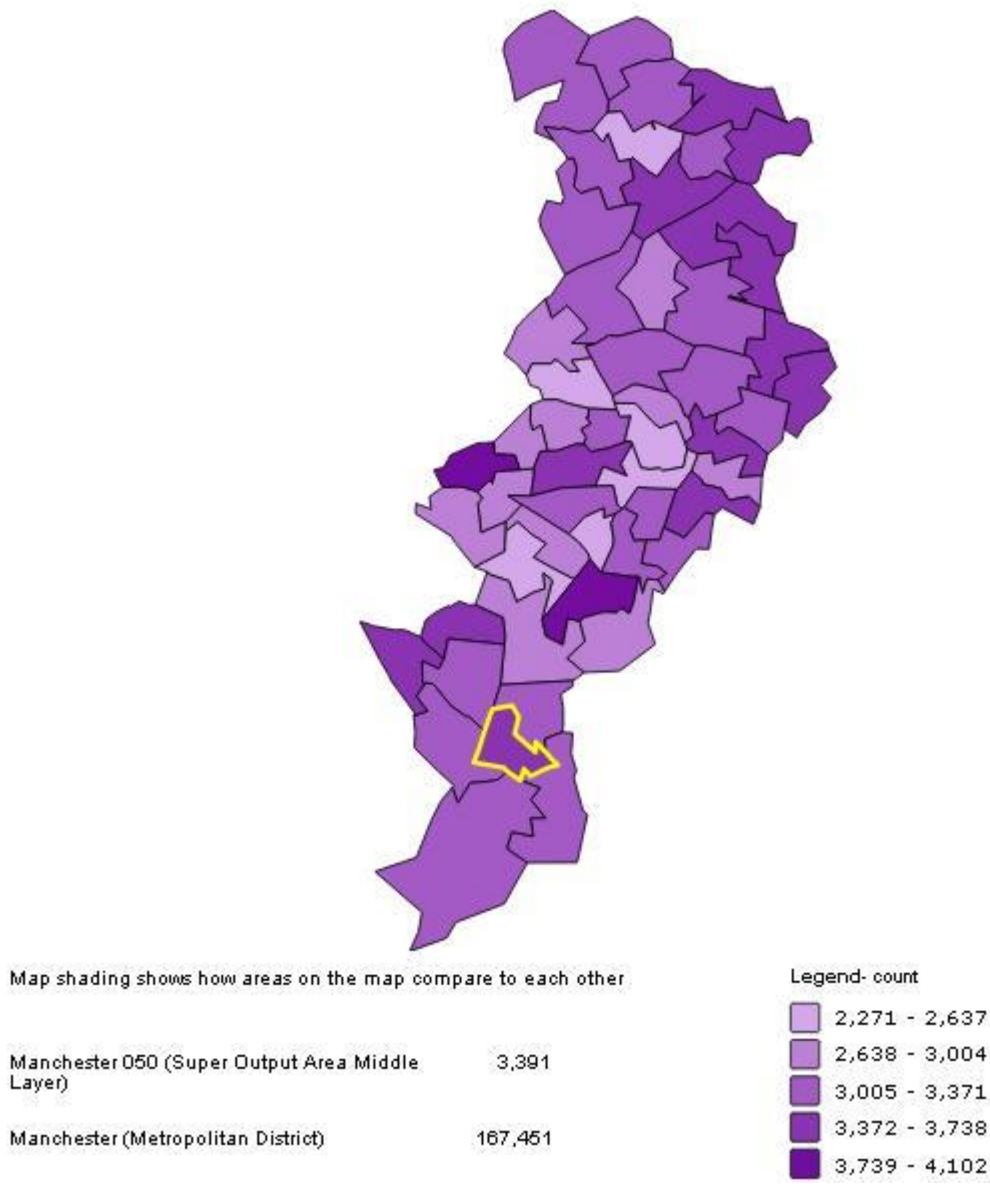

Figure 7.75: Map Showing Occupied Households

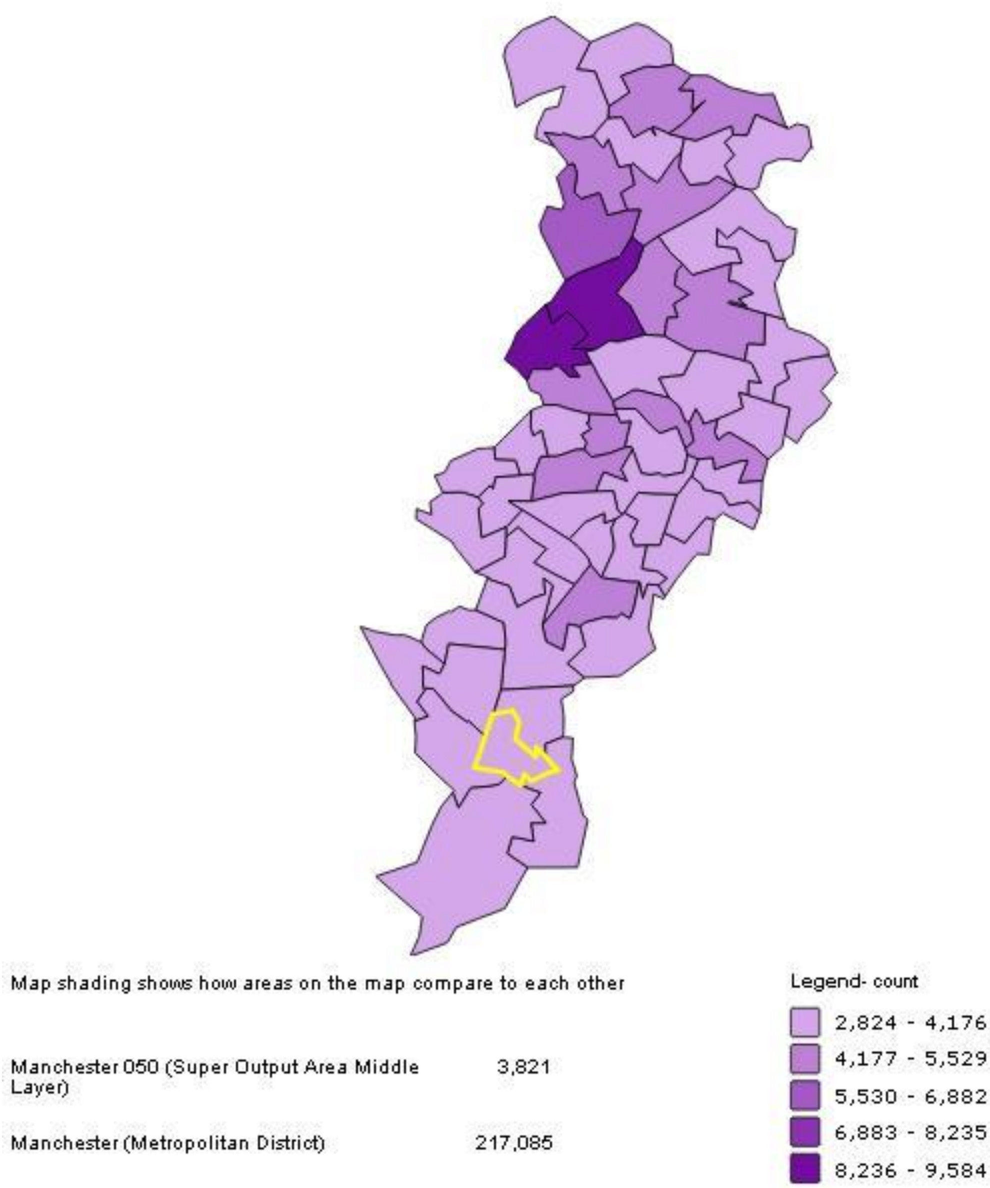

Figure 7.76: Map Showing Dwelling Stock

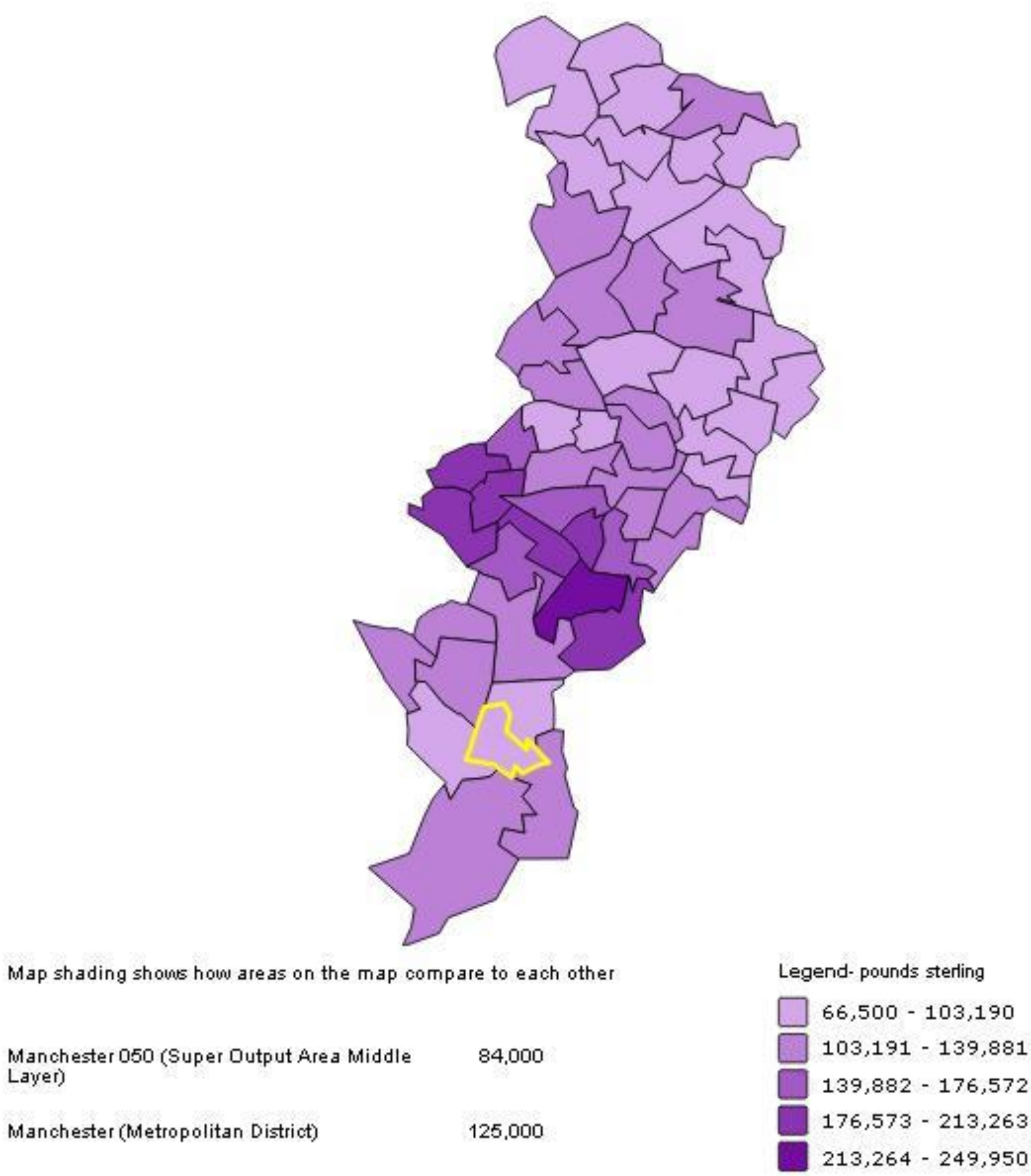

Figure 7.77: Map Showing Dwelling Price

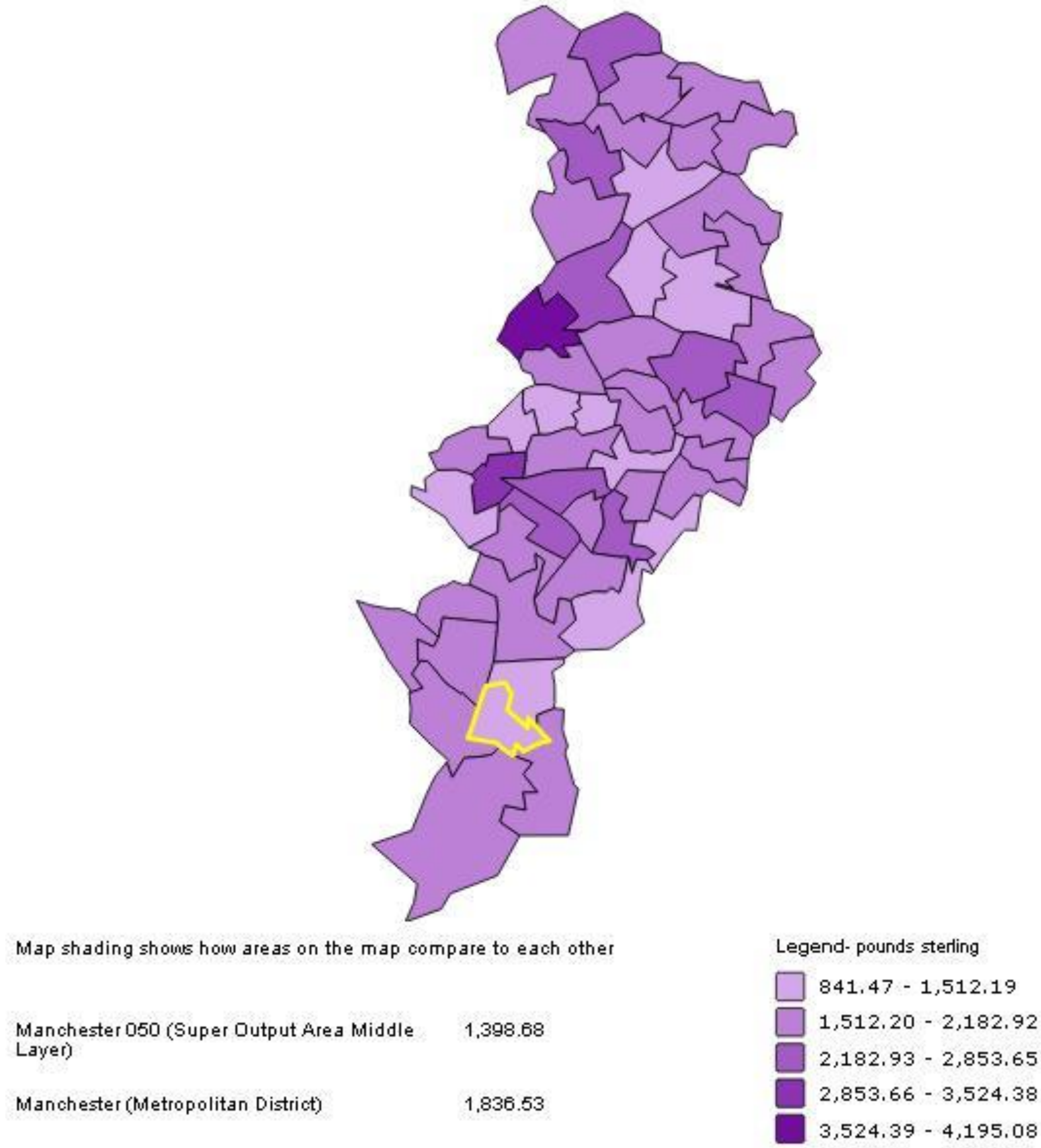

Figure 7.78: Map Showing Personal Debt

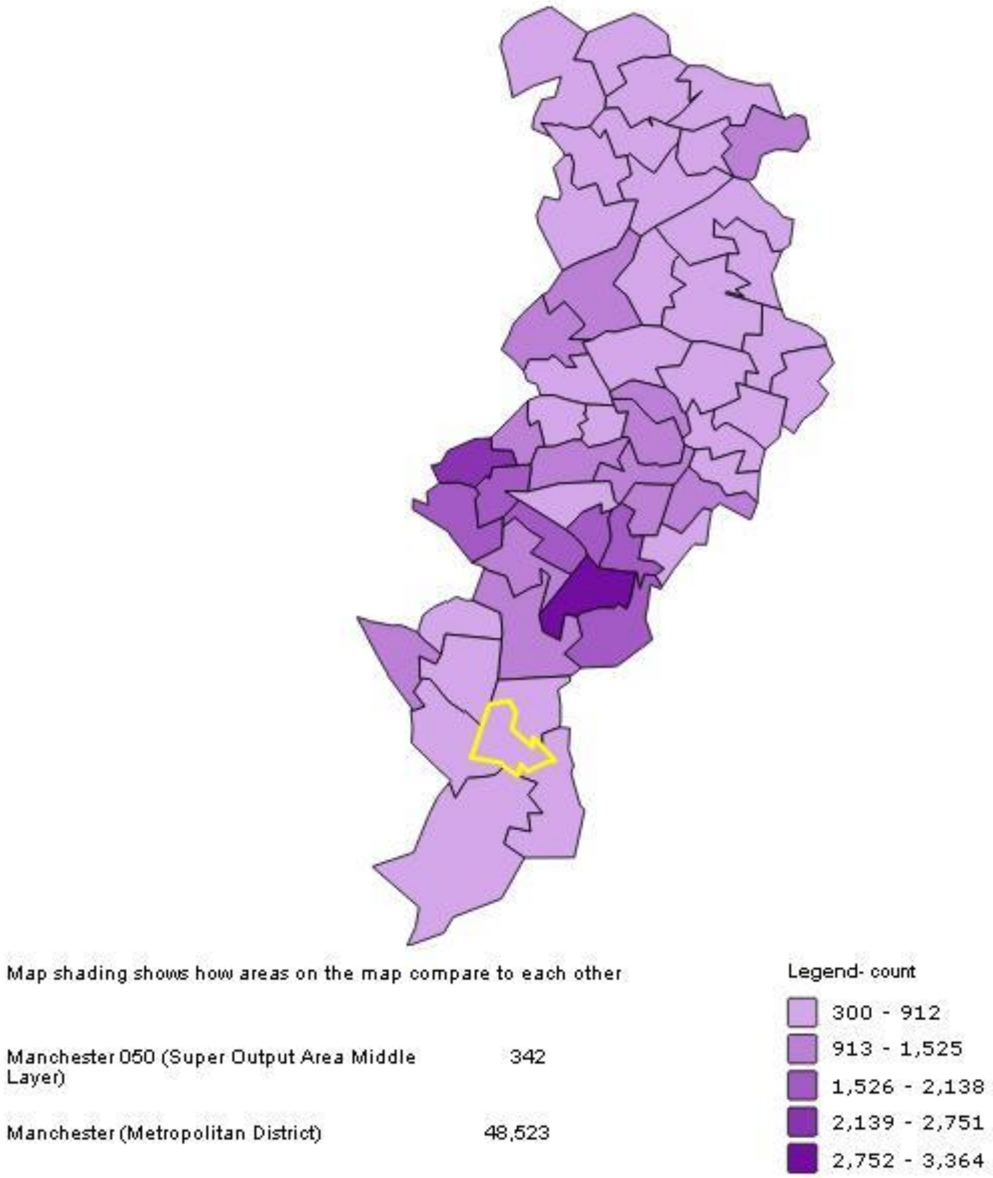

Figure 7.79: Map Showing Number of Higher Professional

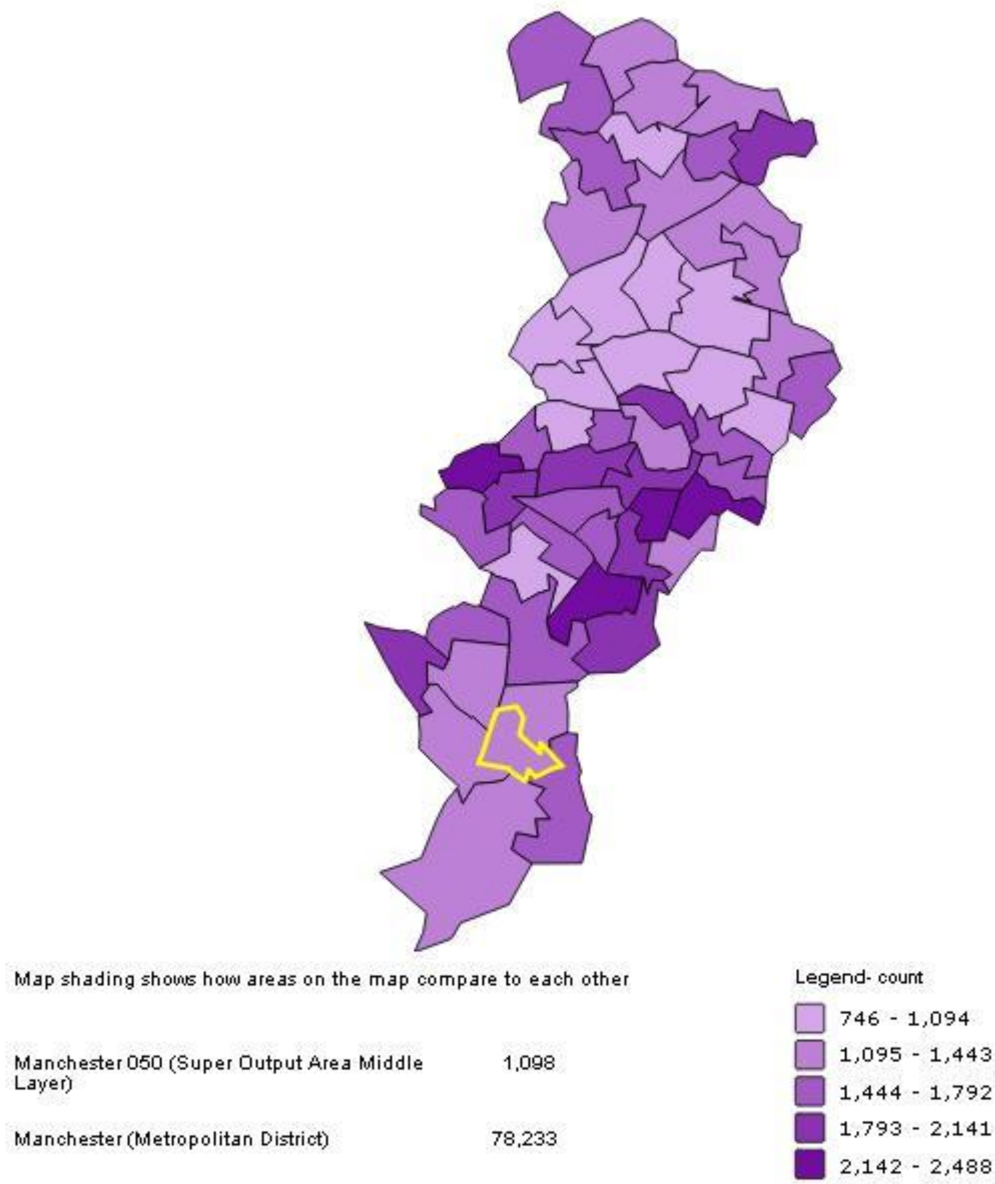

Figure 7.80: Map Showing Number of Supervisory Professional

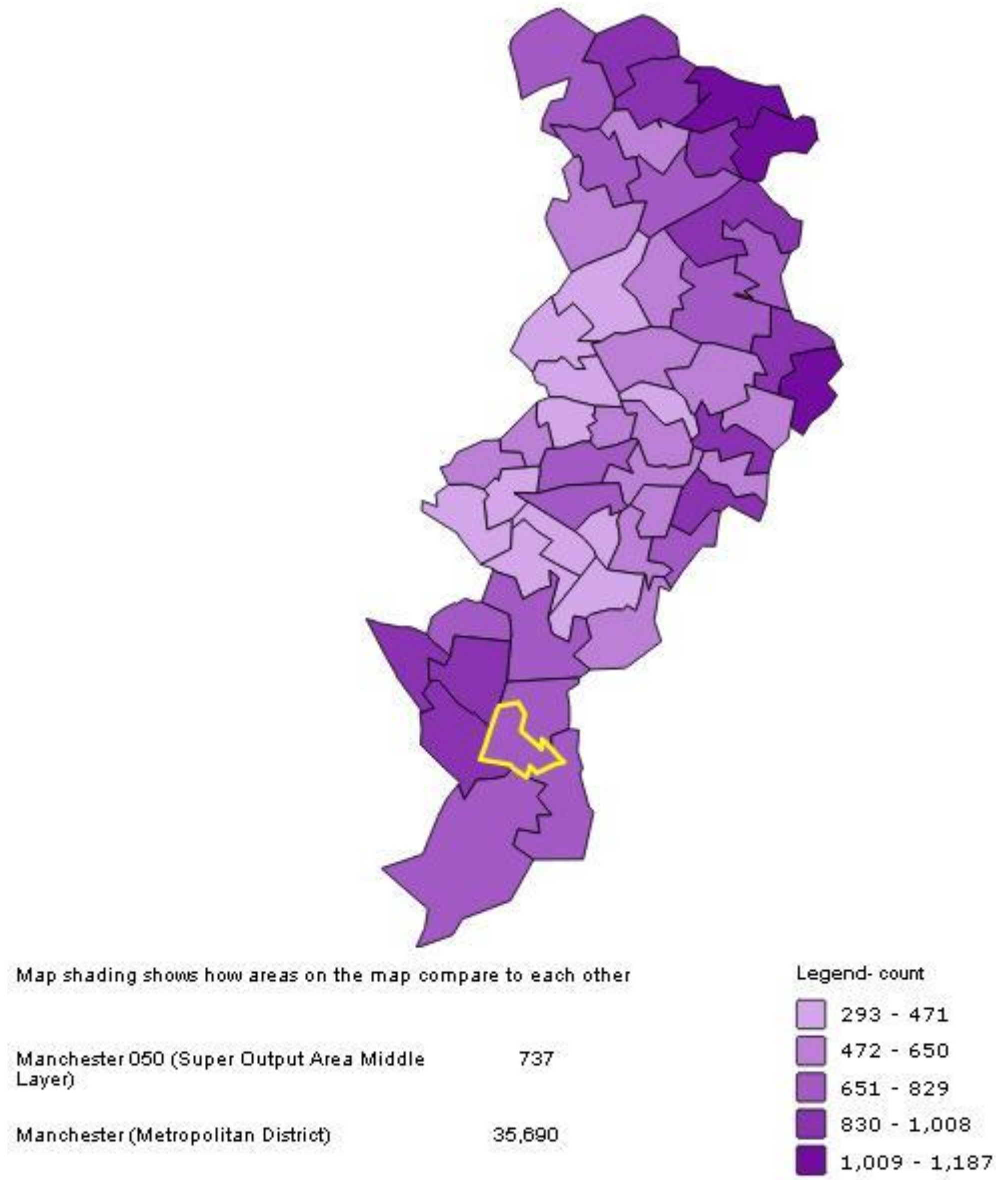

Figure 7.81: Map Showing Number of Skilled Professional

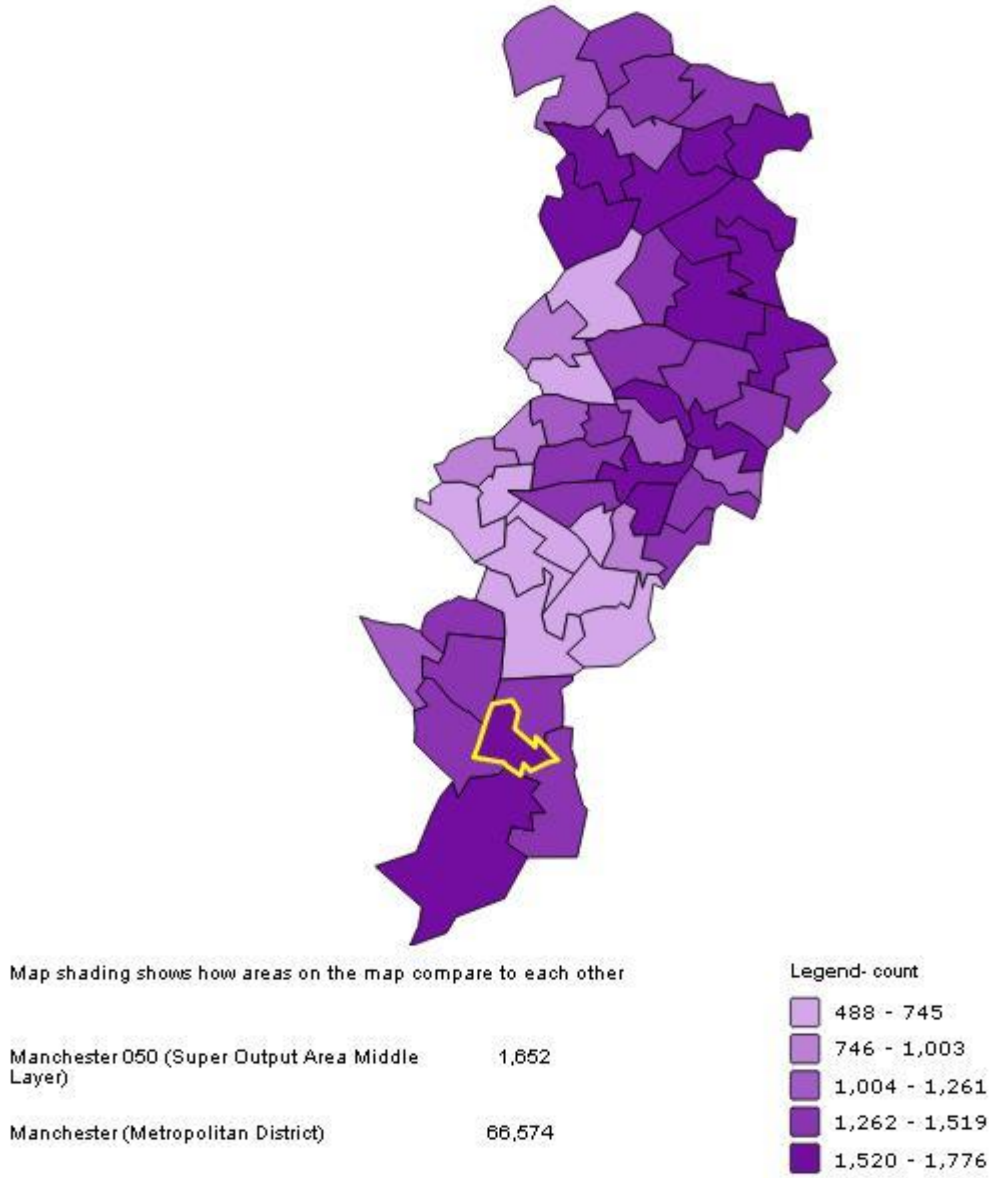

Figure 7.82: Map Showing Number of Semi Skilled Professional

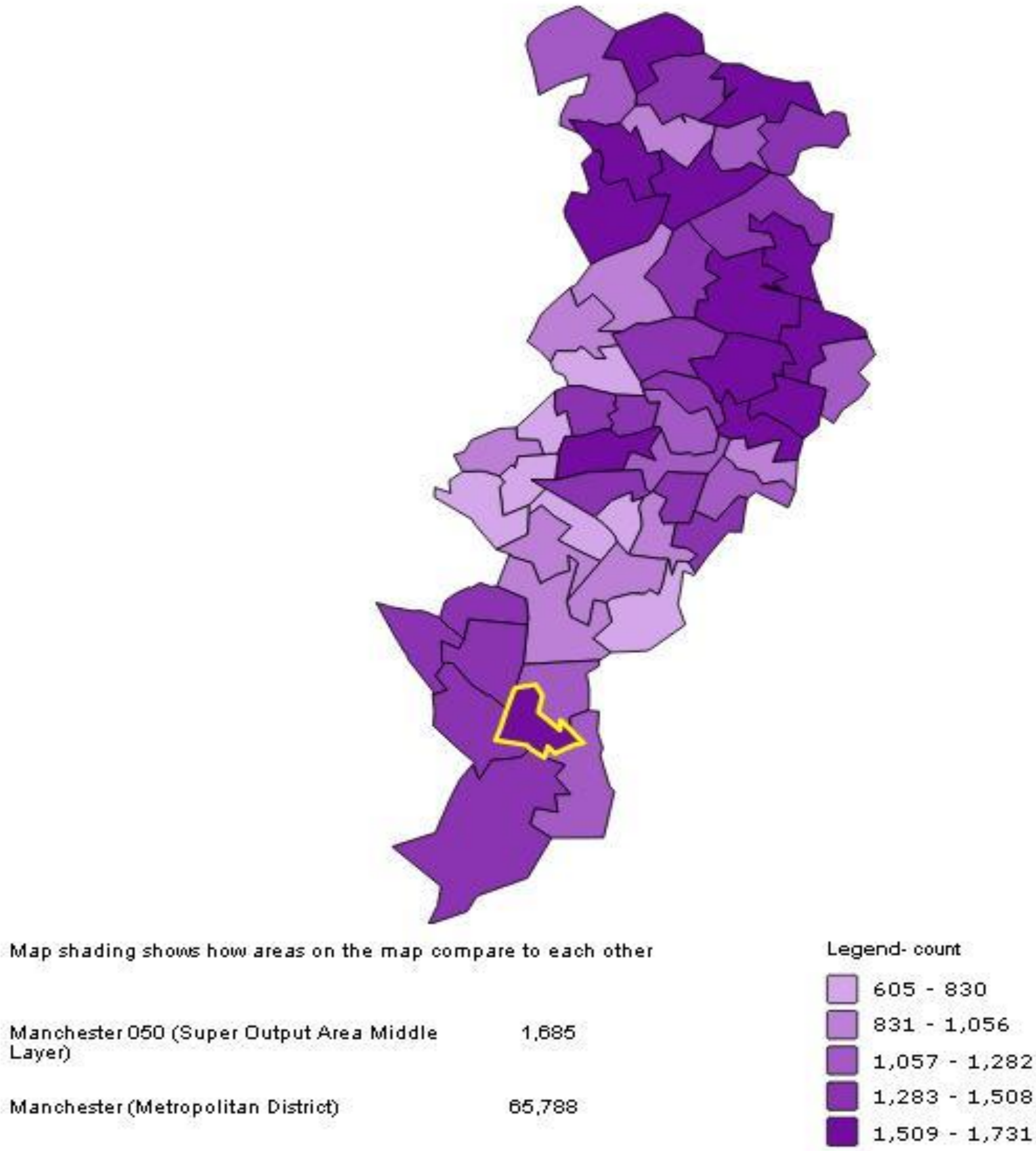

Figure 7.83: Map Showing Number of People Living on State Benefit