Modelling and Forecasting Property Types’ Price Changes and Correlations within the City of Manchester, UK

Hamed Ahmed Al-Marwani
Qatar Fertiliser Company, QAFCO, P.O Box 50001, Mesaieed, Qatar

Abstract: Most of the research done on real estate markets to date has concentrated on aggregate real estate price indices and correlations between regional properties assets. Previous 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 paper covers residential real estate markets by property types (flats, terraced, semi-detached, and detached) within the city of Manchester, UK. The paper covers their time series properties as well as their correlations. The data period is divided into estimation sample from 1995 to 2011 and forecasting sample from 2011 to 2013. 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. Property types correlations show that the highest correlation is between the most expensive properties, detached and semi-detached and the next highest correlations are between the less expensive, terraced and flats due to the close substitution of those property types. 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.

Keywords: Estate Prices, Investment by Property Types, Forecasting Real Estate.

I. INTRODUCTION
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. A gap has clearly been identified as 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. 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.

The paper applies the Box-Jenkins (1974) methodology to the different property types (flats, terraced, semi-detached, and detached) in Manchester during the period from 1995 (quarter one) to 2011 (quarter one). The Box
Jenkins methodology has the advantage over automatic model selection criterion in that it does not tend to overfit the model in addition to its ability to explain the variables dynamics (see Brocks and Tsolacos (1999)). Previous have considered the UK regional house price index of Nationwide Building Society and the correlations within properties within the regions of the UK (Worthington and Higgs, 2003). There is a room for further study of correlations between property types within a region or a city and whether they will provide diversification benefits for real estate investors. This paper is the first to examine residential properties within a region or a city in addition to the UK All house price index. 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. Manchester was chosen since London house prices are more influenced by international factors than Manchester due to massive foreign investments. Most research on real estate in the UK concentrated on correlations between regional properties and very few examined residential property investments. This paper is the first to examine residential properties within a region or a city outside of the Greater London area. The paper is organized in three sections where section one is the introduction, section 2 is the literature review, section 3 has the data, methodology and empirical analysis and section 4 is a summary and conclusions.

II. LITERATURE REVIEW

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 which need worth mentioning 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) and Glaeser and Gyourko (2006)) and the housing market is prone to large time frame of rise in price followed by time periods of shocking price decrease (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 stepahead forecast errors.

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. 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.

### III. DATA, METHODOLOGY AND EMPIRICAL RESULTS

The paper examines price behavior by property types within Greater Manchester, the second largest city in the United Kingdom outside of Greater London. There are sixty five quarter observations available from quarter one
1965 to quarter one 2011 that were obtained from the Land Registry of the UK (www.landregistry.gov.uk). The Land Registry Database covers all the house prices of Greater Manchester. It is hard to define the boundaries of Greater Manchester, and all other cities in the UK present the same problem of overlapping boundaries and boundaries not being defined (Manchester.gov.uk). 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. The quarterly data for property types within Greater Manchester (UK) during the period from quarter 1 of 1995 to quarter 1 of 2011 were obtained from the UK Land Registry (www.landregistry.gov.uk). The city of Manchester was chosen on the basis that it is the second most populous conurbation as per office for national statistics ONS. 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 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. The Land Registry Database covers all the house prices of Greater Manchester. As was highlighted in the research literature, no study has concentrated on property type within a region or city in the UK.

The methodology adopted in this paper is based on Box-Jenkins (1974) approach to modelling univariate time series. The simplest class of time series models that one could entertain is that of the moving average process. Let $u_t$ ($t = 1, 2, 3, \ldots, n$) be a white noise process with Expected ($u_t$) = 0 and variance ($u_t^2$) = $\sigma^2$. Then

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

Equation 1 is a q order moving average model, denoted MA(q), and $\theta$s are 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, denoted an AR(p), can be expressed as

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + u_t \quad (2)$$

Where $\mu$ is the constant, $\phi$s are the parameter estimates for each lag up to lag p, and $u_t$ is white noise disturbance term. Combining the AR(p) and MA(q) models, an ARMA(p,q) model is obtained as in equation 3.

$$y_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \phi_1 u_{t-1} + \ldots + \phi_q u_{t-q} + u_t \quad (3)$$

### Table 1. Key statistics for real estate properties by type for Manchester UK during the period from 1995 Quarter one to 2011 Quarter one.

<table>
<thead>
<tr>
<th></th>
<th>Terraced</th>
<th>Flats</th>
<th>Semi-Detached</th>
<th>Detached</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>70,377.86</td>
<td>108,076.04</td>
<td>107,937.29</td>
<td>175,638.70</td>
</tr>
<tr>
<td>10th Percentile</td>
<td>31,549.06</td>
<td>47,857.50</td>
<td>50,983.18</td>
<td>86,323.93</td>
</tr>
<tr>
<td>25th Percentile</td>
<td>37,173.59</td>
<td>64,263.34</td>
<td>57,742.00</td>
<td>98,806.98</td>
</tr>
<tr>
<td>Median</td>
<td>56,960.11</td>
<td>123,237.57</td>
<td>98,218.23</td>
<td>168,687.54</td>
</tr>
<tr>
<td>75th Percentile</td>
<td>111,738.50</td>
<td>143,024.33</td>
<td>161,041.42</td>
<td>242,245.80</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>122,187.57</td>
<td>152,940.80</td>
<td>170,265.60</td>
<td>281,729.19</td>
</tr>
<tr>
<td>SD</td>
<td>36,394.57</td>
<td>40,634.83</td>
<td>49,305.05</td>
<td>76,278.07</td>
</tr>
</tbody>
</table>
The ARMA(p,q) in equation 3 states that the current value of some series y depends linearly on its own previous values plus a combination of the current and previous 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 (1974) methodology. The automatic order selections are dominated in the literature by the Akaike Information criterion (Hirotugu (1974), AIC). AIC criteria is 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. 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 (1974) which is based on model identification by examining the autocorrelations functions (ACF) and partial autocorrelations functions (PACF). 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 correlation between detached and semidetached is 0.34 which is significant at the 1% level indicating close positive association between price rises and declines for the two property types. The result is expected as clients for detached will start targeting the semi in case of price rise which will push the prices up for the semi-detached. Price declines will also follow similar pattern as a reduction in the price of detached will force the holders of semi-detached properties to reduce their offering prices. The correlation between detached and flat is -0.025 which is small and insignificant at even the 10% indicating no association between price changes for the two properties types. The result is not surprising given that the clients for the properties types are very different with the detached clients more inelastic to price changes than for the flats properties. The correlation between detached and terraced is 0.22 which is significant at the 68%. It indicates prices will tend to rise and fall in the same pattern but correlation is less than for detached and semidetached as detached and semi-detached could be seen as a close substitute for each other while detached and terraced could be less of a substitute for each other. The correlation between semi-detached and flats is 0.12 which is on the margin of 10% significance indicating low correlations. The correlation between semi-detached and terraced is 0.23 which is significant at 5% level. The highest correlation observed between a property type and flat is observed for terraced and flats which is 0.28 and signifyficate at the 5% level. Terraced and flats the lowest categories in prices tend to be more correlated as they are close substitute for real estate buyers.

In summary, 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. Table 2 shows the quarterly and annual returns by property type for the data from quarter one 1995 to quarter 1 2011 (65 observations).

Quarterly average price changes are average price changes over the 65 quarters by property type. Quarterly risk is the standard deviation of quarterly average price changes by property type. Coefficient of variation is calculated by dividing quarterly risk (standard deviation) by quarterly average price change. From Table 2, it can be said that highest quarterly average price changes are for terraced and detached properties (2.63% and 2.5%) followed by flats and semi-detached properties (2.05% and 2%) respectively. The quarterly risk is highest for detached properties followed by terraced, flat
Table 2. Quarterly Average Price Changes, Quarterly Risk and Quarterly Co-efficient of Variation of the Property Types

<table>
<thead>
<tr>
<th>Variables</th>
<th>Quarterly Average Price Changes</th>
<th>Quarterly Risk</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi</td>
<td>0.0205</td>
<td>0.0671</td>
<td>3.28</td>
</tr>
<tr>
<td>Detached</td>
<td>0.0255</td>
<td>0.1312</td>
<td>5.13</td>
</tr>
<tr>
<td>Flat</td>
<td>0.0206</td>
<td>0.0823</td>
<td>4.04</td>
</tr>
<tr>
<td>Terraced</td>
<td>0.0263</td>
<td>0.1129</td>
<td>4.34</td>
</tr>
</tbody>
</table>

Figure 1: Price Change for Flats in Manchester

Figure 2: Terraced Price Change
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.

Flats price changes between successive quarters have a mean of 0.021 and standard deviation of 0.083. Figure 1 has the price changes for Flats in Manchester from quarter two 1995 to quarter one 2011.

From figure 1, price changes can be categorized into three scales: low volatility, medium volatility and high volatility. In the above figure, flats price changes had low volatility between the periods 2002 Q2 to 2007 Q2; medium volatility between 1995 Q2 to 2001 Q2; and high volatility, as a result of the credit crunch, between 2008 Q1 to 2011 Q1. Figure 3 plots the terraced price changes for Manchester from quarter two 1995 to quarter one 2011.

The price changes for terraced has a mean of 0.026 with a standard deviation is 0.113. Figure 2 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 3 has the price changes for the Semi- Detached versus the UK all house prices quarterly from quarter two 1995 to quarter one 2011. Price changes for semi-detached have a mean of 0.020 and standard deviation of 0.07. Figure 3 shows that price changes can be categorized into three scales: low volatility between the period 1995 Q2 and 2001 Q2, medium volatility between 2002 Q2 and 2007 Q2 and high volatility, as a result of the credit crunch of 2008, between 2008 Q1 and 2011 Q1. Figure 4 plots the price changes for the Detached for the period from
quarter two 1995 to quarter one 2011. The price changes for the detached have a mean of 0.03 and a standard deviation of 0.131. Figure 4 indicates that detached price change can be categorized into three periods: low volatility, medium volatility and high volatility. In figure-4, low volatility between the period from 1995 Q2 to 2001 Q2; medium volatility between 2002 Q2 to 2007 Q2 and high volatility between 2008 Q1 to 2011 Q1. The Box-Jenkins approach to modelling time series models using autoregressive moving average (ARMA) of time series variables, as explained in methodology section, is applied to the four property price changes. The results are summarized in table 3.

The results show that Manchester’s most expensive housing type, detached and semi-detached, have longer memory as indicated by more lags in the autoregressive parts in contrast to the less expensive property types, terraced and detached, which have a short memory of one quarter. The detached and semi-detached properties 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 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.
higher adjusted R-squared for detached and semi-detached than for flats and terraced. Accordingly, detached and semi-detached should have outperformed out of sample in comparison to terraced and flats. However, the results indicate that out of sample forecasts outperformed better for terraced and flats in contrast to in sample expectation.

IV. SUMMARY

Previous research has considered the UK regional house price index of Nationwide Building Society and the correlations within properties within the regions of the UK (Worthington and Higgs, 2003). There is a room for further study of correlations between property types within a region or a city and whether they will provide diversification benefits for real estate investors. The paper examines residential properties within a region or a city. 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 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.

The paper’s limitations are related to the utilized data. The models developed in this paper 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.

REFERENCES


Hamed Ahmed Al-Marwani has earned his Ph.D. degree from Brunel University London, United Kingdom in the field of Management Studies Research, and completed the AMP-Advance Management Program from IESE Business School – Barcelona, Spain. He has his Bachelor’s degree in Computer Science from East Carolina University, USA. He is a Certified Information Systems Auditor (CISA) He is working at the top management level in QAFCO in the position of Chief Administrative Officer. He has more than 30 years’ in-depth experience in various fields such as Information Technology, Industrial Relations, Public Relations, Human Resources, General Services and numerous projects. His analytical & management experiences are gained through working at the grass root and executive level at QAFCO.

In summary, he has the Executive Management experience on management issues involving people, technology and process, negotiations management both at macro and micro level.