Received 6 November 2010
Accepted 8 February 2011

Forecasting future trends in Dubai housing market by using Box-Jenkins autoregressive integrated moving average

Ali Hepşen
Department of Finance, Faculty of Business Administration, Istanbul University, Istanbul, Turkey, and
Metin Vatansever
Department of Mathematics and Statistics, Faculty of Arts and Science, Yıldız Technical University, Istanbul, Turkey

Abstract

Purpose – It is important to forecast index series to identify future rises, falls, and turning points in the property market. From the point of this necessity and importance, the main purpose of this paper is to forecast the future trends in Dubai housing market.

Design/methodology/approach – This paper uses the monthly time series of Reidin.com Dubai Residential Property Price Index (DRPPI) data. In order to forecast the future trends in Dubai housing market, Box-Jenkins autoregressive integrated moving average (ARIMA) forecasting method is utilized.

Findings – The results of the ARIMA modeling clearly indicate that average monthly percentage increase in the Reidin.com DRPPI will be 0.23 percent during the period January 2011-December 2011. That is a 2.44 percent increase in the index for the same period.

Practical implications – Reidin.com residential property price index is a crucial tool to measure Dubai's real estate market. Based on the current index values or past trend, real estate investors (i.e. developers and constructors) decide to start new projects. Attempts have also been made in the past to forecast index series to identify future rises, falls, and turning points in the property market. The results of this paper would also help government and property investors for creating more effective property management strategies in Dubai.

Originality/value – There is no previous study analyzing the future trends in Dubai housing market. At this point, the paper is the first academic study that identifies this relationship.

Keywords Dubai, Real estate, Residential property, Price index, Forecasting, Box-Jenkins, Autoregressive integrated moving average (ARIMA)

Paper type Research paper

Introduction

There are a lot of individuals or organizations using residential property price indices directly or indirectly either to influence practical decision making and conduct of economic policy. Analysts, policymakers, investors, and financial institutions follow trends in house prices to expand their understanding of real estate and credit market conditions, as well as their impact on economic activity, and financial stability and soundness (Case and Wachter, 2005). For instance, mortgage lenders use this information to gauge default risk and central banks often rely on movements in house price indices to track households borrowing capacity (Finocchiaro and Queijo, 2009)

International Journal of Housing Markets and Analysis
Vol. 4 No. 3, 2011
pp. 210-223

1753-8270
DOI 10.1108/17538271111153004

and aggregate consumption (Belsky and Prakken, 2004). Thus, it is important to forecast index series to identify future rises, falls, and turning points in the real estate market.

Real estate has been a key driver of growth and has been a steady and robust performer over the years. For instance, the real estate sector contributed 14.73 percent of Dubai's total gross domestic product (GDP) in 2008 (Dubai Statistics Center, 2010). Factors that have been critical to the strong development of this sector include property rights, transaction costs, and capital gains taxes. With the opening up of the market to allow freehold ownership of properties to foreigners in Dubai, international investors have driven a huge demand for properties. At that point, monitoring the evolution of real estate prices and the market trend over time is obviously a necessary requirement in Dubai. This paper focuses forecasting future trends in Dubai housing market by using the Reidin.com (2010) Dubai Residential Property Price Index (DRPPI). The results of this paper would help government and property investors for creating more effective property management strategies in Dubai.

The layout of the remainder of this article is as follows. The next section presents a literature review in real estate market forecasting models. The following section provides overview of autoregressive integrated moving average (ARIMA) models and Box-Jenkins procedure, and the data. Section four describes the development of the forecasting model and the results of this study. The final section is the conclusion.

Literature review

Forecasting models can be broadly categorized into two groups: multivariate forecasting models and univariate forecasting models. Multivariate forecasting models attempt to explain changes in a variable by references to the movements in the current or past values of other (explanatory) variables. Whereas, univariate forecasting models constitute a class of specifications in which one attempts to model and to predict time series variables using only information contained in their own past values and current and, possibly, past values of an error term (Brooks and Tsolacos, 2010; Bönner, 2009).

Many studies have been focused on the multivariate forecasting in the office market but few are focused on residential. Grebler (1979) points out that income, the consumer price index (CPI), seasonal factors, vacancy rate, and previous housing prices can be used to predict the housing prices. Rosen (1984) develops a multi-equation model including seven supply and demand equations. Within his setting, he forecasts the stock of office space, the flow of new office construction, the vacancy rate, and eventually the rent for office space by using historical San Francisco data from the period of 1961-1983. Hekman (1985) examines the rental price adjustment mechanism and the investment response in the US office construction market, using data from 14 cities over the period of 1979-1983. Case (1990) forecast the housing price and excess returns in the housing market by using the percentage change in real per capita income, real construction costs, adult population, marginal tax rate, and housing starts for four American cities (Atlanta, Chicago, Dallas, and San Francisco). Giussani and Tsolacos (1993) analyze the determination of UK real office rents by incorporating cointegration analysis. Tsolacos *et al.* (1998) apply this US-literature-based model to the UK office market. However, they use lagged changes in new office building, in GDP, and in employment in the banking, finance, and insurance sector instead of vacancy rates to determine the rental adjustment model. Seko (2003) examines the correlation between Japanese dwelling prices and economic fundamentals in some areas. He uses a time series model to forecast housing prices with some indicators, including average sales price of private ownership dwelling, annual household income, population, new-started dwelling, the CPI, and the vacancy rate.

Besides the application of econometric multivariate forecasting models, univariate forecasting models, such as ARIMA and generalized AR conditional heteroskedasticity (GARCH) models can also be implemented to predict future trends in real estate markets. McGough and Tsolacos (1995) obtain a dynamic structure of the UK commercial property sector real rent over the period from the second quarter 1977 to the second quarter 1993 to model ARIMA processes for the retail, office, and industrial sectors. Tse (1997) determines the direction of the office and industrial property market in Hong Kong by using ARIMA models. Wilson *et al.* (2000) compares univariate forecasting techniques in the USA, the UK, and Australian property markets. The article shows that both the ARIMA and spectral regression modeling processes are capable of predicting turning points in property markets. Crawford and Fratantoni (2003) compare the forecasting performance of three types of univariate time series models: ARIMA, GARCH and regime-switching by applying quarterly annualized growth rate data over the period from 1979 to 2001. Stevenson (2007) studies the relative accuracy of ARIMA-based forecasts using quarterly UK office rent data over rolling windows. The author conducts forecasts for a variety of alternative specifications to evaluate the best fitting in sample model relative to other alternatives.

Theoretical framework and the data

We first start with a brief explanation of the theoretical framework of ARIMA models and Box-Jenkins procedure which are important for understanding the methodology of this study. In this section, we also provide an overview of data used in the estimation and forecasting process.

ARIMA models

ARIMA models consist of three components: lagged values of the variable of interest (the AR component), lagged values of the error term (the MA component) and the degree of integration (the number of differences required to make a series stationary) (Crawford and Fratantoni, 2003). AR models were first introduced by Yule in 1926. They were subsequently supplemented by Slutsky who in 1937 presented MA schemes. The ARMA models have been originated from the combination of the AR and MA by Wold (1938). ARMA process can be used to model a large class of stationary time series as long as the appropriate order of p, the number of AR terms and q, the number of MA terms was appropriately specified (Makridakis and Hibon, 1997). Such a model 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 an error term. The ARMA (p, q) can be written:

$$y_t = \mu + \O_1 y_{t-1} + \O_2 y_{t-2} + \cdots + \O_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (1)$$

with:

$$E(\varepsilon_t) = 0; \quad E\left(\varepsilon_t^2\right) = \sigma^2; \quad E(\varepsilon_t, \varepsilon_s) = 0, \;\; t \neq s \quad (2)$$

where $\{\text{Ø}_t\}$ and $\{\theta_t\}$ are the coefficients, p and q are the orders of AR and MA polynomials, respectively. y_t is the value of the time series observation at time t, ε_t is a series of random shocks which are assumed to be independently, normally distributed with constant mean 0, variance and uncorrelated with each other (equation (2)). Similarly, a non-stationary model can be repressed as ARIMA (p, q, d), where the letter "I", between AR and MA, stood for the "Integrated" and reflected the need for differencing to make the series stationary (Makridakis and Hibon, 1997; Johnston and Dinardo, 1997; Brooks and Tsolacos, 2010). The ARIMA (p, d, q) model can be written:

$$\text{Ø}(B)(I - B)^d y_t = \theta(B)\varepsilon_t \tag{3}$$

with:

$$E(\varepsilon_t) = 0; \quad E\left(\varepsilon_t^2\right) = \sigma^2; \quad E(\varepsilon_t, \varepsilon_s) = 0, \ \ t \neq s \tag{4}$$

where B represents the backshift operator such that By $= y_{t-1}$, d represents the order of difference. If a series is stationary, then d = 0. In equation (3), *Ø(B)* is a polynomial of order p in the backshift operator B, which is defined as (Tse, 1997):

$$\text{Ø}(B) = 1 - \sum_{i=1}^{p} \text{Ø}_i B^i \tag{5}$$

Similarly, $\phi(B)$ is defined to be a polynomial of order q in B, such that:

$$\theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i \tag{6}$$

The Box-Jenkins procedure

Box and Jenkins (1976) popularized the use of ARIMA models through the following three steps (Makridakis and Hibon, 1997):

(1) *Model identification.* The step involves determining the order of the model required to capture the dynamic features of the data. Graphical procedures are used (plotting the autocorrelation function (ACF) and partial ACF (PACF) of the time series) to decide which (if any) AR or MA component should be used in the model (Makridakis and Hibon, 1997; Johnston and Dinardo, 1997; Brooks and Tsolacos, 2010). To achieve this, first ARIMA needs to be stationary, that is, it should have a constant mean, variance and autocorrelation through time. Since the data are non-stationary, we have to transform the series to induce stationarity.

(2) *Parameter estimation.* This step involves estimating the parameters of the models specified in model identification step. Computational algorithms (least squares or another technique, known as maximum likelihood) are used to arrive at coefficients which best fit the selected ARIMA model (Makridakis and Hibon, 1997; Johnston and Dinardo, 1997; Brooks and Tsolacos, 2010).

(3) *Diagnostic checking.* This step is to test whether the model specified and estimated is adequate. Box and Jenkins suggest two methods: over fitting

and residual diagnostics. Over fitting involves deliberately fitting a larger model than that required to capture the dynamics of the data as identified in step 1. If the model specified at step 1 is adequate, any extra terms added to the ARIMA model would be insignificant. Residual diagnostics implies checking the residuals. The residuals should be white noise (or independent when their distributions are normal) drawings from a fixed distribution with a constant mean, variance and uncorrelated with each other. (Plotting autocorrelation and partial autocorrelation of the residuals are helpful to identify misspecification.) If the model is found to be inadequate, it is necessary to go back to step 2 and try to identify a better model (Makridakis and Hibon, 1997; Johnston and Dinardo, 1997; Brooks and Tsolacos, 2010). Figure 1 shows the three steps of the Box-Jenkins methodology.

Overview of data

The primary purpose of this research is to forecast the future trends by using Box-Jenkins ARIMA in Dubai housing market. Monthly time series of Reidin.com DRPPI data are used for the period January 2003-March 2010, a total of 87 observations.

Reidin.com DRPPIs are designed to be a reliable and consistent benchmark of housing sales prices in Dubai. Index series are calculated monthly and quarterly, and cover seven city-wide, 11 main districts, and seven major communities/projects. They are powered by transaction data made available through the Dubai Land Department. Index series employ arithmetic average of the median prices of districts for constructing index series (the unit method). All indices are also calculated by using an MA algorithm. An MA is commonly used with time series data to smooth out short-term fluctuations and highlight longer term trends or cycles. Outliers and extreme values (as a result

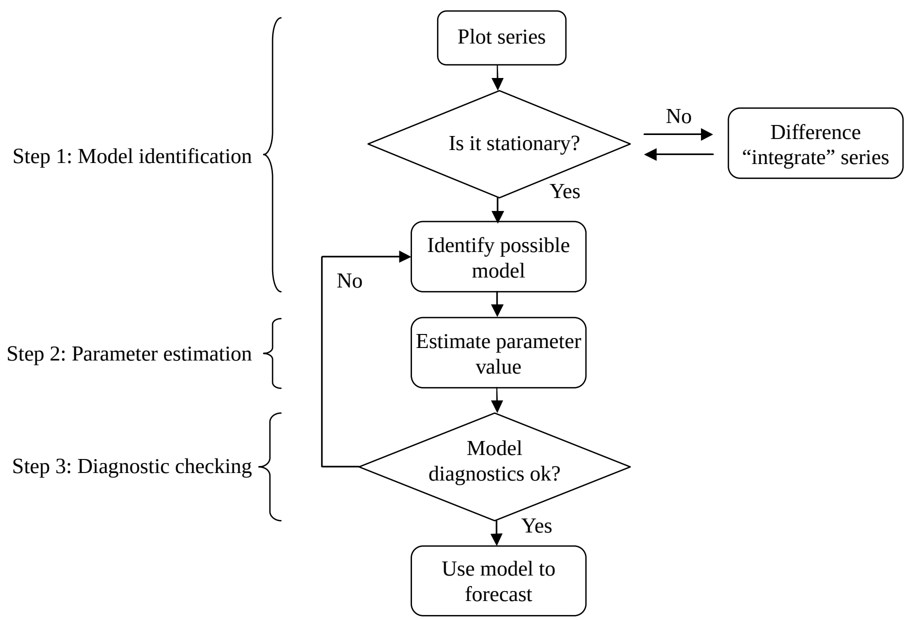


Figure 1.
Box-Jenkins procedure

of incomplete, inconsistent or erroneous data) are excluded by the outlier detection procedure of the inter-quartile range (IQR) based on the calculated price per square meter of each property. This commonly used methodology considers any data that is more than 1.5 times the IQR from the upper or lower quartile to be an outlier. They are calculated by using the Dutot price index formula.

Figure 2 shows time series plot of Reidin.com DRPPI.

A close look at the graph in Figure 2, the plotted data indicates that DRPPI has non-constant mean and non-constant variances, i.e. it seems to be non-stationary.

Development of the forecasting model and evaluation of study results

As mentioned above, ARIMA model or Box-Jenkins procedure consists of three steps. In the first step, namely model identification step, we need to identify the appropriate parameters in our ARIMA (p, d, q) model.

EViews 5.0 statistical software package is used to establish the ARIMA model. To start applying ARIMA model, ACF and PACF should be determined. Autocorrelation and partial autocorrelation graphics, which provides information about the AR and MA orders, are then drawn based on the specified lag numbers. If ACF will decline steadily, or follow a damped cycle and PACF will cut off suddenly after p lags, the data are AR (p). If ACF will cut off suddenly after q lags and PACF will decline steadily, the data are MA (q).

Figure 3 graph shows us with the correlogram up to 30 lags. The column labeled autocorrelation and partial autocorrelation are the sample ACF and the sample PACF, respectively. Also the diagrams of autocorrelation and partial autocorrelation are provided on the left. The solid and dashed vertical lines in the diagram represent the zero axes and 95 percent confidence interval, respectively. From this figure, two facts stand out: first, the autocorrelation coefficient starts at a very high value at lag 1 (0.981) and declines very slowly, and ACF up to 22 lags are individually statistically significant different from zero as they are all outside the 95 percent confidence bounds. Second, after the lag 1, the PACF drops dramatically, and all PACFs after lag 2 are statistically insignificant. These two facts strongly support the idea that the DRPPI time series is non-stationary. It may be non-stationary in mean or variance, or both.

Since the data are non-stationary, we have to make it stationary by differencing non-stationary series one or more times to achieve stationarity. Figure 4 shows the correlograms of the first-differenced data up to 30 lags. Figure 4 shows that the ACF

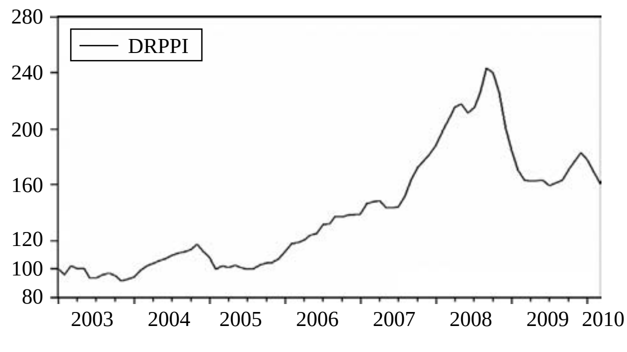
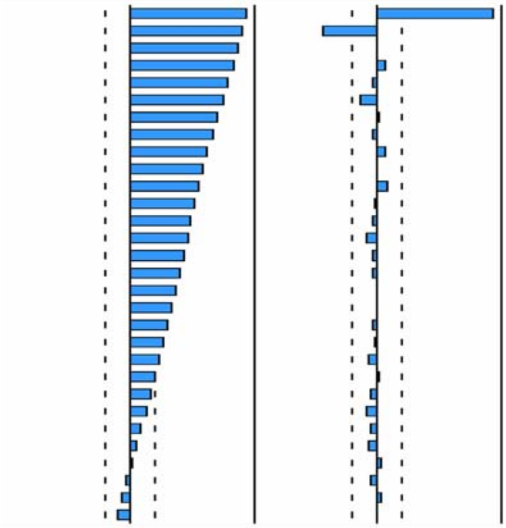


Figure 2.
Reidin.com DRPPI
(January 2003 = 100)

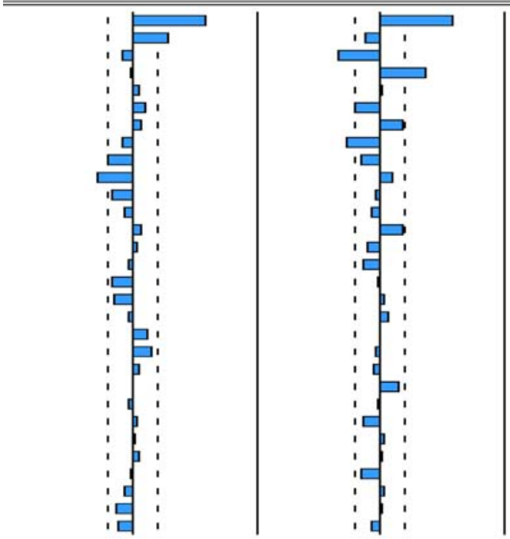
Sample: 2003:01 2010:03
Included observations: 87



	AC	PAC	Q-stat.	Prob.
1	0.981	0.981	86.718	0.000
2	0.946	–0.456	168.30	0.000
3	0.905	–0.004	243.74	0.000
4	0.863	0.074	313.24	0.000
5	0.823	–0.034	377.16	0.000
6	0.780	–0.138	435.30	0.000
7	0.735	0.010	487.64	0.000
8	0.690	–0.026	534.27	0.000
9	0.646	0.060	575.75	0.000
10	0.607	0.005	612.74	0.000
11	0.572	0.084	646.08	0.000
12	0.541	–0.020	676.33	0.000
13	0.512	–0.032	703.76	0.000
14	0.481	–0.092	728.32	0.000
15	0.448	–0.041	749.90	0.000
16	0.413	–0.041	768.50	0.000
17	0.378	0.004	784.34	0.000
18	0.345	–0.008	797.71	0.000
19	0.312	–0.031	808.81	0.000
20	0.279	–0.018	817.79	0.000
21	0.243	–0.076	824.71	0.000
22	0.206	0.010	829.78	0.000
23	0.169	–0.042	833.25	0.000
24	0.131	–0.096	835.37	0.000
25	0.093	–0.050	836.44	0.000
26	0.051	–0.073	836.78	0.000
27	0.011	0.034	836.79	0.000
28	–0.029	–0.046	836.90	0.000
29	–0.067	0.026	837.50	0.000
30	–0.102	0.007	838.91	0.000

Figure 3.
The correlogram of DRPPI data up to 30 lags

Sample: 2003:01 2010:03
Included observations: 86



	AC	PAC	Q-stat.	Prob.
1	0.609	0.609	33.027	0.000
2	0.298	–0.117	41.007	0.000
3	–0.085	–0.348	41.667	0.000
4	–0.023	0.380	41.710	0.000
5	0.041	0.021	41.868	0.000
6	0.104	–0.208	42.884	0.000
7	0.078	0.181	43.468	0.000
8	–0.095	–0.282	44.344	0.000
9	–0.211	–0.152	48.733	0.000
10	–0.295	0.109	57.411	0.000
11	–0.166	–0.027	60.205	0.000
12	–0.067	–0.068	60.664	0.000
13	0.069	0.190	61.156	0.000
14	0.036	–0.113	61.289	0.000
15	–0.028	–0.134	61.375	0.000
16	–0.169	–0.011	64.471	0.000
17	–0.159	0.029	67.232	0.000
18	–0.033	0.072	67.352	0.000
19	0.118	–0.005	68.929	0.000
20	0.160	–0.034	71.870	0.000
21	0.145	–0.043	72.107	0.000
22	–0.001	0.149	72.107	0.000
23	–0.035	–0.013	72.253	0.000
24	0.026	–0.140	72.338	0.000
25	0.017	0.028	72.372	0.000
26	0.043	0.025	72.604	0.000
27	–0.015	–0.154	72.633	0.000
28	–0.070	0.026	73.274	0.000
29	–0.138	0.015	75.806	0.000
30	–0.123	–0.074	77.856	0.000

Figure 4.
The correlogram of the first-differenced DRPPI data up to 30 lags

and PACF start at a high value at lag 1 (0.609) and drop dramatically. These can support the idea that the DRPPI time series is non-stationary. Perhaps, the second-differenced data are stationary.

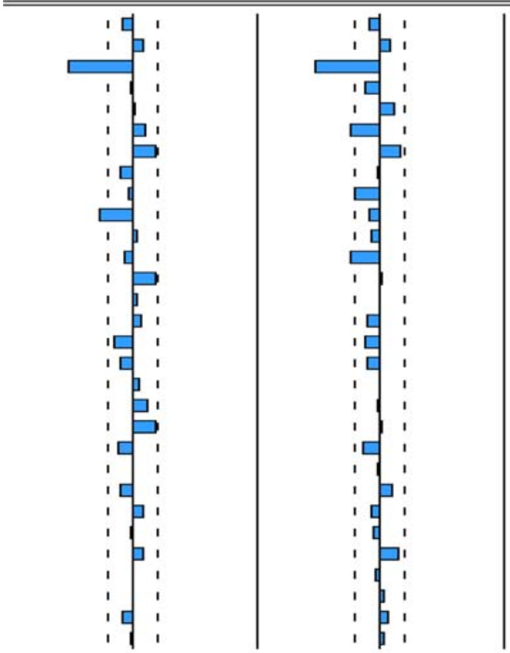
Figure 5 shows the correlograms of the second-differenced data up to 30 lags. In Figure 5, we have a much different pattern of ACF and PACF. The ACFs at lags 3, 7, 10, 13 and 20; and PACFs at 3, 6, 9 and 12 seem statistically different from zero. But at all other lags, they are not statistically different from zero.

Visual inspection, or correlograms, does not give any strong indication of non-stationary. This is confirmed by formal tests such as the augmented Dickey-Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) unit root tests. Tables I-III give the results of unit root tests[1]. The null hypothesis of the test is that there is a unit root against the alternative one that there is no unit root in the variables.

The ADF, PP, and KPSS statistics for the DRPPI is insignificant at 1, 5 and 10 percent levels of significant, which leads to no rejection of the null hypothesis that there is a unit root problem in the DRPPI. Based on ADF, PP, and KPSS tests, it is obvious that the DRPPI is non-stationary.

Differencing has the effect of making the variables stationary. Tables II and III summarize the results of unit root tests for the differenced variables. The ADF and KPSS statistics for the first-difference DRPPI is insignificant at 1, 5 and 10 percent levels of significant and the PP statistics is just insignificant at 1 percent level of significant,

Sample: 2003:01 2010:03
Included observations: 85



	AC	PAC	Q-stat.	Prob.
1	–0.090	–0.090	0.7121	0.399
2	0.091	0.084	1.4523	0.484
3	–0.544	–0.537	28.152	0.000
4	–0.017	–0.127	28.178	0.000
5	0.021	0.118	28.218	0.000
6	0.100	–0.254	29.161	0.000
7	0.195	0.171	32.755	0.000
8	–0.102	–0.014	33.757	0.000
9	–0.032	–0.204	33.855	0.000
10	–0.273	–0.091	41.204	0.000
11	0.034	–0.068	41.322	0.000
12	–0.065	–0.251	41.750	0.000
13	0.196	0.017	45.679	0.000
14	0.035	–0.001	45.806	0.000
15	0.065	–0.107	46.256	0.000
16	–0.164	–0.115	49.123	0.000
17	–0.111	–0.104	50.467	0.000
18	0.042	0.002	50.658	0.000
19	0.118	–0.012	52.208	0.000
20	0.183	0.014	56.026	0.000
21	–0.120	–0.141	57.699	0.000
22	–0.004	–0.012	57.701	0.000
23	–0.101	0.112	58.927	0.000
24	0.087	–0.061	59.850	0.000
25	–0.023	–0.045	59.917	0.000
26	0.094	0.155	61.021	0.000
27	–0.003	–0.040	61.022	0.000
28	0.007	0.025	61.029	0.000
29	–0.092	0.078	62.147	0.000
30	–0.010	0.031	62.162	0.001

Figure 5.
The correlogram of the second-differenced DRPPI data up to 30 lags

but the ADF, PP, and KPSS statistics for the second-difference DRPPI is significant at 1, 5 and 10 percent levels of significant, which leads to rejection of the null hypothesis that there is a unit root problem in the DRPPI. Based on unit root tests, it is apparent that the second-difference DRPPI is stationary, which implies that the DRPPI is integrated of order one, I (2). Tables I-III show the correlograms, which tell a similar story.

The series correlogram has allowed us to choose appropriate p and q for the data series. We have estimated more models in order to determine the right specification, by choosing from both the different models estimated on adjusted R^2, sum of squared resid, log likelihood, the Akaike, Schwarz's Bayesian information criteria and by generating predictions on the basis of estimated models. The series correlogram suggests the necessity of introduction in the process estimation of both the analyzed variable lags and the lags of the error. We have started with an AR (1) process and further analyzed the residual correlogram in order to catch the correlations and autocorrelations from lags bigger than 1. From adjusted R^2, sum of squared resid, log likelihood, the Akaike and Schwarz criterias' point of view, the proper model to best adjust the data is ARIMA (3, 2, 6).

Table I.
Summary of unit root tests in the variables (in level from with a trend and intercept)

Unit root tests	Test statistics	Probability	Test critical value (1%)	Test critical value (5%)	Test critical value (10%)	Results
ADF	−2.321	0.418	−4.074	−3.466	3.159	Fail to reject the null
PP	−1.654	0.7628	−4.068	−3.463	−3.158	Fail to reject the null
KPSS	0.116		0.216	0.146	0.119	Fail to reject the null

Table II.
Summary of unit root tests in the variables (in first difference from with a trend and intercept)

Unit root tests	Test statistics	Probability	Test critical value (1%)	Test critical value (5%)	Test critical value (10%)	Results
ADF	−3.051	0.125	−4.074	−3.466	−3.159	Fail to reject the null
PP	−3.912	0.016	−4.070	−3.464	−3.158	Fail to reject the null
KPSS	0.107		0.216	0.146	0.119	Fail to reject the null

Table III.
Summary of unit root tests in the variables (in second difference from with a trend and intercept)

Unit root tests	Test statistics	Probability	Test critical value (1%)	Test critical value (5%)	Test critical value (10%)	Results
ADF	−9.452	0.000	−4.074	−3.466	−3.159	Reject the null
PP	−23.992	0.000	−4.071	−3.464	−3.159	Reject the null
KPSS	0.299		0.216	0.146	0.119	Reject the null

The second step in ARIMA modeling is parameter identification. Let $\Delta^2 Y_t$ denote the second-differenced data. Then, in line with the conclusion in the first step, our model is:

$$\Delta^2 Y_t = \mu + \beta_1 \Delta^2 Y_{t-3} + \beta_2 \varepsilon_{t-6} + \varepsilon_t \tag{7}$$

Using EViews 5.0, we obtained the following estimates:

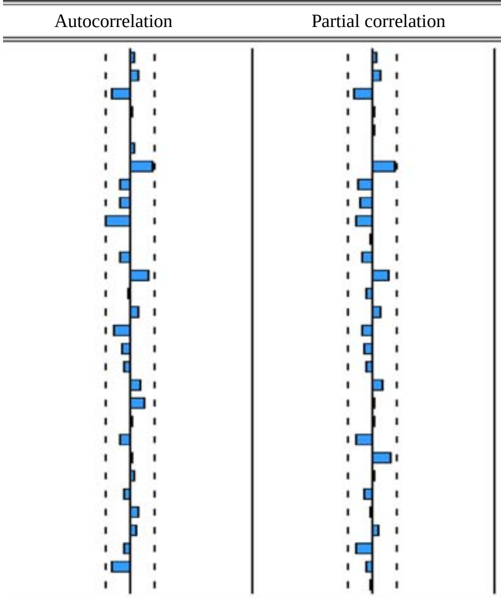
$$\Delta^2 \hat{Y}_t = -0.936 \Delta^2 Y_{t-3} - 0.915 \varepsilon_{t-6} \tag{8}$$

The coefficients of the model are significantly different of 0 (the *t*-test).

In the third step of ARIMA, a diagnosis check is used to validate the model assumptions. This diagnosis checks if the hypotheses made on the residuals (actual prices minus fitted prices, as estimated in second step) are true. Residuals must satisfy the requirements of a white noise process. We obtain residuals from equation (7) and get the ACF and PACF of these residuals up to lag 30 in order to check that the model represented by equation (7) is a reasonable fit to the data. The estimated ACF and PACF are shown in Figure 6. As can be seen in Figure 6, none of the autocorrelations and partial correlations is individually statistically significant. In other words, the correlograms of autocorrelation and partial autocorrelation give the impression that the residuals estimated from regression (8) are white noise. Hence, there is not any need to look for another ARIMA model.

The model from diagnostic checking step can be used to predict future value of DRPPI. However, we need to integrate the second-differenced series to obtain the forecast of DRPPI rather than its changes. We know that the following formula integrates data from second-differenced form into level form:

Sample: 2003:01 2010:03
Included observations: 83



	AC	PAC	Q-stat.	Prob.
1	0.038	0.038	0.1241	0.725
2	0.064	0.063	0.4812	0.786
3	–0.150	–0.155	2.4342	0.487
4	0.011	0.020	2.4446	0.655
5	–0.007	0.012	2.4493	0.784
6	0.029	0.004	2.5272	0.865
7	0.184	0.193	5.6481	0.581
8	–0.094	–0.122	6.4709	0.595
9	–0.095	–0.112	7.3298	0.603
10	–0.206	–0.134	11.405	0.327
11	0.000	–0.013	11.405	0.410
12	–0.091	–0.094	12.211	0.429
13	0.156	0.131	14.637	0.331
14	–0.013	–0.043	14.655	0.402
15	0.071	0.078	15.175	0.439
16	–0.141	–0.085	17.252	0.370
17	–0.079	–0.061	17.912	0.394
18	–0.047	–0.057	18.147	0.446
19	0.091	0.084	19.060	0.453
20	0.123	0.019	20.745	0.412
21	0.008	0.010	20.754	0.474
22	–0.095	–0.139	21.799	0.472
23	0.018	0.154	21.836	0.530
24	0.031	0.024	21.954	0.582
25	–0.054	–0.067	22.300	0.618
26	0.068	–0.016	22.861	0.641
27	0.051	0.057	23.187	0.675
28	–0.048	–0.134	23.481	0.709
29	–0.158	–0.042	26.745	0.585
30	–0.003	–0.016	26.746	0.637

Figure 6.
The correlogram of residuals from equation (8)

$$\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1} = Y_t - Y_{t-1} - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2} \quad (9)$$

If we transform all variables in equation (9) based on this formula and rearrange it, our model becomes:

$$Y_t = 2Y_{t-1} - Y_{t-2} + \beta_1 Y_{t-3} - 2\beta_1 Y_{t-4} + \beta_1 Y_{t-5} + \beta_2 \varepsilon_{t-6} + \varepsilon_t \quad (10)$$

The values of β_1 and β_2 are already known from the estimated regression (8) and ε_t is assumed to be zero, which enables us to convert equation (10) into equation (11). Using equation (11), we may easily obtain the forecast values for the period April 2010-December 2011:

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} - 0.936Y_{t-3} + 1.872Y_{t-4} - 0.936Y_{t-5} - 0.915\varepsilon_{t-6} \quad (11)$$

Using equation (11), Y_t can be estimated and compared with its actual values (Figure 7).

Before presenting the results, it is useful to validate the present model with observed data. In order to do this, DRPPI is calculated by equation (11) supposing that present month is April 2010, i.e. nine months observed data are used for validation. As can be seen in Table IV, mean forecast error, mean absolute error, mean absolute percentage error, mean squared error, root mean squared error, root mean squared percentage error, theil inequality coefficient, bias proportion, variance proportion, and covariance proportion forecast evaluation statistics are -2.37, 2.47, 1.58, 9.05, 3.00, 0.04, 0.01, 0.62, 0.03 and 0.35, respectively, which may definitely be regarded as within the acceptable range.

By using equation (11), DRPPI forecasts are obtained for the period January 2011-December 2011. As given in Table V, the results from ARIMA modeling clearly indicate that average monthly percentage increase in the Reidin.com DRPPI will be 0.23 percent during the period January 2011-December 2011. That is a 2.44 percent increase in the index for the same period.

Conclusion

It is important to forecast index series to identify future rises, falls, and turning points in the property market. From the point of this necessity and importance, the main

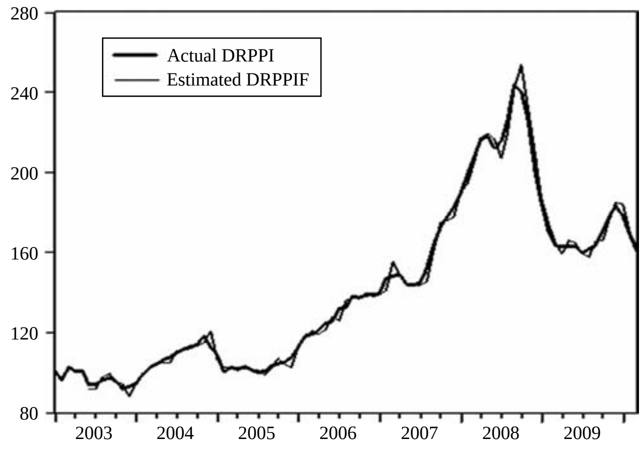


Figure 7.
Actual DRPPI and estimated DRPPIF

Date	Actual DRPPI	Monthly % change	Forecasted DRPPI	Monthly % change	Forecast errors
April 2010	159.20	−1.05	158.76	−2.22	0.44
May 2010	157.23	−1.23	158.18	−0.37	−0.95
June 2010	157.06	−0.11	158.66	−0.31	−1.59
July 2010	156.52	−0.34	159.13	0.30	−2.60
August 2010	156.27	−0.16	159.05	−0.05	−2.78
September 2010	158.82	1.63	159.31	0.17	−0.50
October 2010	156.64	−1.37	159.59	0.18	−2.95
November 2010	154.67	−1.25	160.38	0.50	−5.71
December 2010	156.19	0.98	160.85	0.29	−4.66
Forecast evaluation statistics					
Mean forecast error	−2.37				
Mean absolute error	2.47				
Mean absolute percentage error	1.58				
Mean squared error	9.05				
Root mean squared error	3.00				
Root mean squared percentage error	0.04				
Theil inequality coefficient	0.01				
Bias proportion	0.62				
Variance proportion	0.03				
Covariance proportion	0.35				

Table IV. Evaluation of forecasts for DRPPI, April 2010-December 2010

Date	Forecasted DRPPI	Monthly % change
January 2011	161.31	0.28
February 2011	161.28	−0.01
March 2011	161.56	0.17
April 2011	161.85	0.18
May 2011	162.59	0.46
June 2011	163.05	0.28
July 2011	163.49	0.27
August 2011	163.52	0.02
September 2011	163.81	0.18
October 2011	164.11	0.18
November 2011	164.80	0.42
December 2011	165.25	0.27

Table V. Forecast for DRPPI, January 2011-December 2011

purpose of this study is to forecast the future trends in Dubai housing market by using Box-Jenkins ARIMA. ARIMA modeling clearly indicates that average monthly percentage increase in the residential property price index will be 0.23 percent during the period January 2011-December 2011. That is a 2.44 percent increase in the index for the same period.

The findings of this paper would help Dubai Government and property investors for creating more effective property management strategies. For the government side, future rises, falls, and turning points of the property prices puts into perspective the effects of government policy created to deal with them. Revealing the duration and magnitude of cycles allow for better understanding of the course of house prices, which, in turn, helps government policy makers take the best stance in reaction to the property price changes. Knowing this can allow the government to monitor the property price behavior and control the speculation activities that cause dramatic changes. For the investors' side, future prices might increase the possibility of diligent investor in the real estate market to reap abnormal profits by using a trading rule mainly based on the observed behavior of property prices.

Note

1. Schwartz Info Criterion have been used in ADF unit root tests. Kernal sum of covariances estimator with Barlett weights and Newey West automotic bandwith selection methods have been used in PP and KPSS unit root tests. The KPSS output only provides the asymptotic critical values.

References

Belsky, E. and Prakken, J. (2004), "Housing wealth effects: housing's impact on wealth accumulation, wealth distribution and consumer spending", Joint Center for Housing Studies, No. W04-13, Harvard University, Boston, MA.

Bönner, A. (2009), "Forecasting models for the German office market", dissertation thesis, Graduate School of Business Administration, Economics, Law and Social Science, The University of St Gallen, St Gallen.

Box, G.P.E. and Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control*, 2nd ed., Holden-Day, San Francisco, CA.

Brooks, C. and Tsolacos, S. (2010), *Real Estate Modeling and Forecasting*, 1st ed., Cambridge University Press, New York, NY.

Case, K.E. (1990), "Forecasting prices and excess returns in the housing market", *AREUEA Journal*, Vol. 18 No. 3, pp. 253-73.

Case, B. and Wachter, S. (2005), "Residential real estate price indices as financial soundness indicators: methodological issues", BIS Papers, No. 21, Bank for International Settlements, Basle, pp. 197-211.

Crawford, G.W. and Fratantoni, M.C. (2003), "Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices", *Journal of Real Estate Economics*, Vol. 31 No. 2, pp. 223-43.

Dubai Statistics Center (2010), *Gross Domestic Product at Basic Current Prices – Emirate of Dubai*, available at: www.dsc.gov.ae/Reports/GDP_CurrentPrices_08_Ar.pdf (accessed 1 July 2010).

Finocchiaro, D. and Queijo, H.V. (2009), "Do central banks react to house prices?", Sveriges Riksbank Working Paper Series, No. 217, Sveriges Riksbank, Stockholm.

Giussani, B. and Tsolacos, S. (1993), "The office market in the UK: modeling the determinants of rental values", paper presented at the International Real Estate Research Session of the 1993 ASSA/AREUEA Conference, Anaheim.

Grebler, L. (1979), *The Inflation of Housing Price: Its Extent, Cause and Consequences*, Lexington Books, Lexington, MA.

Hekman, J.S. (1985), "Rental price adjustment and investment in the office market", *AREUEA Journal*, Vol. 13 No. 1, pp. 31-47.

Johnston, J. and Dinardo, J. (1997), *Econometric Methods*, 4th ed., McGraw-Hill, New York, NY.

McGough, T. and Tsolacos, S. (1995), "Forecasting commercial rental values using ARIMA models", *Journal of Property Valuation & Investment*, Vol. 13, pp. 6-22.

Makridakis, S. and Hibon, M. (1997), "ARIMA models and the Box-Jenkins methodology", *Journal of Forecasting*, Vol. 16 No. 3, pp. 147-63.

Reidin.com (2010), Real Estate Investment & Development Information Network, available at: www.reidin.com

Rosen, K.T. (1984), "Toward a model of the office building sector", *AREUEA Journal*, Vol. 12 No. 3, pp. 261-9.

Seko, M. (2003), "Housing prices and economic cycles", paper presented at The International Conference on Housing Market and the Macro Economy, Hong Kong.

Stevenson, S. (2007), "A comparison of the forecasting ability of ARIMA models", *Journal of Property Investment & Finance*, Vol. 25 No. 3, pp. 223-40.

Tse, R.Y.C. (1997), "An application of the ARIMA model to real-estate prices in Hong Kong", *Journal of Property Finance*, Vol. 8 No. 2, pp. 152-3.

Tsolacos, S., Keogh, G. and McGough, T. (1998), "Modeling use, investment and development in the British office market", *Environment and Planning*, Vol. 30, pp. 1409-27.

Wilson, P.J., Okuneu, J., Ellis, C. and Higgins, D. (2000), "Comparing univariate forecasting techniques in property markets", *Journal of Real Estate Portfolio Management*, Vol. 6 No. 3, pp. 283-306.

Wold, H. (1938), *A Study in the Analysis of Stationary Time Series*, Almqvist and Wiksell, Stockholm.

Further reading

Agung, I.G.N. (2009), *Time Series Data Analysis Using EViews*, 1st ed., Wiley, Singapore.

Edigar, V.Ş. and Akar, S. (2007), "ARIMA forecasting of primary energy demand by fuel in Turkey", *Energy Policy*, Vol. 35 No. 3, pp. 1071-708.

Corresponding author

Ali Hepşen can be contacted at: alihepsen@yahoo.com