Journal of Property Research

ISSN: 0959-9916 (Print) 1466-4453 (Online) Journal homepage: www.tandfonline.com/journals/rjpr20

A comparison of alternative rental forecasting models: empirical tests on the London office market

Simon Stevenson & Oliver McGarth

Journal of Property Research, September 2003, **20**(3) 235–260 Spon Press Taylor & Francis Group

A comparison of alternative rental forecasting models: empirical tests on the London office market

SIMON STEVENSON[1]* and OLIVER MCGARTH[2]

[1]*Centre for Real Estate Research, Department of Banking & Finance, Graduate School of Business, University College Dublin, Carysfort Avenue, Blackrock, Dublin, Ireland. E-mail: simon.stevenson@ucd.ie*

[2]*Trinty Real Estate, Duke Street, Dublin 2*

Received 6 November 2000
Revised 16 September 2002; 12 August 2003
Accepted 20 August 2003

Summary

The study examines four alternative rental forecasting models in the context of the London office market. The forecasting ability of an ARIMA model, a Bayesian Vector Autoregression approach, an OLS based single equation model and a simultaneous equation model are compared and contrasted. The models are estimated using the CB Hillier Parker London Office index over the period 1977–1996, with out-of-sample testing undertaken on the following three years of data. Diagnostic testing is also conducted on the alternative models. The findings reveal that the Bayesian VAR model produces the best forecasts, while the ARIMA model fails to pick up on the large uptake in rental values during the testing period.

Keywords: Rental forecasting, Vector Autoregression, ARIMA modelling, London office market

1. Introduction

The issue of rental forecasting has been the subject of an extensive amount of empirical study and investigative research. Additionally, it is also an area that has seen academic research being quickly adopted by institutions as an aid in asset allocations decisions.[1] Research in the forecasting field has largely centred on three alternative approaches. The first has largely been based around vacancy rates (Wheaton, 1987; Wheaton and Torto, 1988), while the second has tended to use demand and supply variables in reduced form models. In addition, studies such as McGough and Tsolacos (1994, 1995b) and Brooks and Tsolacos (2000), have examined the use of ARIMA and Vector Autoregressive (VAR) models in the context of rental forecasting.

However, despite the large literature in the area, few studies have explicitly compared the forecasting ability of alternative models. Brooks and Tscolacos (2000) is one of the few papers to have compared alternative approaches. This paper compares the forecasting

*Author to whom correspondence should be addressed.

[1] See Mitchell and MacNamara (1997).

Journal of Property Research
ISSN 0959–9916 print/ISSN 1466–4453 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0959991032000162338

performance of vector-autoregressive and ARIMA based models with respect to the UK retail sector. The paper finds that the relative accuracy of the different approaches alters with regard to the forecasting horizon adopted. Hendershott *et al.* (2002) highlight the divergence in approach between the UK and US literature and the preference for reduced form and equilibrium rental adjustment models respectively. The papers shows that both approaches can be effectively viewed as special cases of a more general model that the authors illustrate within an error correction framework. The error correction model proposed has the advantage over the formal rental adjustment model in that it incorporates the flexibility of reduced form models, this being particularly important in the context of markets where supply-side data is limited. Most previous studies have tended to provide evidence as to the forecasting ability of the final model selected. Chaplin (1998, 1999), however, provides evidence that in the majority of cases, the best fitting model actually fails to provide the best forecast of rental movements. It is this issue that the current study aims to examine in the context of the London office market.

This paper compares four alternative methods of forecasting, using the CB-Hillier Parker Rental Index for Central London offices.[2] The index is available on a semi-annual basis from May 1977, with the period up until May 1996 used for model construction and the following three years of data used to assess out-of-sample performance. The four models examined are ARIMA modelling, a Bayesian Vector Autoregression approach, the single equation model and the simultaneous model proposed by Wheaton *et al.* (1997). All of the analysis is conducted in real terms. The use of a three-year forecasting period is specifically used due to the importance of long-term forecasts within the development process.[3] The use of alternative models in the forecasting processes should allow an examination of whether the use of alternative techniques results obtains stronger results over different horizons. In particular, previous studies of ARIMA based models have often found it strongest at short-horizons, which given the nature of the model and the resulting forecasts in not surprising. One key hypothesis to be examined in this study is whether those models relying more on fundamental demand and supply side variables provide more accurate forecasts over the three-year horizons assumed and tested in this paper.

The CB-Hillier Parker index holds an advantage over other comparable publicly available indices on two counts. First, it allows an examination of a specific market, both in geographic and sector terms and secondly its use of hypothetical properties, rather than an actual portfolio, allows more rapid incorporation of rental changes into the index. This second point should ensure that less smoothing is evident in the index figures in comparison to a portfolio-based index. In addition, the index should be better than valuation based measures in reflecting changes in effective rents. This is of particular relevance in the context of the current study due to its analysis of the London market and the extensive use of incentives in the downturn of the early 1990s. While the index construction methods do not guarantee the elimination of the impact of incentives, the CB-Hillier Parker index can account for them far more affectively than other indices due to its construction method and the use of hypothetical properties.

The remainder of the paper is laid out as follows. Initially, each of the four models are discussed and modelled. Each section discusses the rationale behind the respective

[2] Details of the other data used in this study are given in appropriate section of the paper.
[3] See Tsolacos *et al.* (1998) for an examination and modelling of the British office market.

model, the data required and the preliminary diagnostic statistics of them. The following section then discusses, and compares, the out-of-sample performance of each of the four models. The final section provides concluding comments.

2. ARIMA model

An ARIMA model is a univariate model that seeks to depict a single variable as an Autoregressive Integrated Moving Average process. Herein, the series is fully described by p, the order of the AR component, q, the order of the MA component and d, the order of integration. The AR component is built upon the assumption that future realizations can be approximated and predicted by the behaviour of current and past values. The MA component, on the other hand, seeks to depict the processes where the effects of past environmental innovations continue to reverberate for a number of periods. If is an ARIMA p,d,q process, then the series evolves according to the following specification:

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \quad (1)$$

Where θ_0 is a constant, ε is the error term, q is the number of lagged terms of ε and p is the number of lagged terms of y_t. The ARIMA model can be described as atheoretical, as it ignores all potential underlying theories, except those that hypothesize repeating patterns in the variable under study. Hence, its forecasting performance vis-à-vis the structural models will bear testimony to the influence that structural forces have on the market. There has been some debate as to the minimum number of observations required to generate an ARIMA model, with both McGough and Tsolacos (1995b) and Tse (1997) recommending at least 50 observations, while papers such as Weiss and Andersen (1984) argue that 30 observations can be sufficient. A total of 38 observations are used in the model construction phase of this paper, therefore, the relative limited number should be borne in mind both in terms of the estimation of the model selected and the forecasting ability of that model.

This paper adopts a three-phase process in the construction of the model, namely, identification, estimation and application. The first phase is concerned with one of the underlying principles governing the generation of ARMA models, in that the series analysed is stationary, i.e. that its stochastic properties are invariant to time transition. Failure to observe this condition will introduce multicollinearity problems. To assess this issue we use the Dickey–Fuller unit root test is used, which can be represented as follows:

$$\Delta y_i = a_0 + p Y_{t-1} + \varepsilon_t \quad (2)$$

In addition, the Ljung–Box modified Q statistic is used, which can be formulated as:

$$Q = T(T+2) \sum_{k=1}^{s} r_k^2 / (T-k) \quad (3)$$

The unit root results are reported in Table 1, while the modified Q statistics are contained in Table 2. It can be seen that with both the series levels and when first differenced,

Table 1. Unit root tests for stationarity

Null hypothesis	No differencing	First differences	Second differences
A(1) = 0	−2.8572	−1.6506	−3.2387*
A(0) = A(1) = 0	4.0876	1.3625	5.2710*

Note: The appropriate asymptotic critical values at a 5% level for the unit root tests are –2.86 for the t-test and 4.59.

Table 2. Modified Box–Pierce statistics[a]

No differencing		First differences		Second differences	
Lag	*Q*	Lag	*Q*	Lag	*Q*
1	36.45***	1	28.50***	1	1.10
2	64.20***	2	45.07***	2	1.54
3	81.24***	3	51.39***	3	2.06
4	88.69***	4	51.93***	4	2.37
5	90.26***	5	52.75***	5	2.59
6	90.30***	6	58.75***	6	8.94
7	92.34***	7	68.83***	7	12.05*
8	97.91***	8	80.18***	8	17.29**
9	106.78***	9	87.93***	9	17.33**
10	117.37***	10	93.03***	10	17.42*
11	127.78***	11	95.71***	11	17.54*
12	136.53***	12	96.56***	12	18.56*
13	143.07***	13	96.87***	13	18.99
14	147.57***	14	96.87***	14	21.85*
15	150.50***	15	96.88***	15	22.14
16	152.32***	16	96.95***	16	22.36
17	153.45***	17	97.25***	17	23.52
18	154.32***	18	97.58***	18	23.54
19	155.17***	19	98.00***	19	23.62
20	156.09***	20	98.25***	20	23.65
21	157.11***	21	98.50***	21	24.73
22	158.07***	22	98.55***	22	24.75
23	158.88***	23	98.55***	23	25.04
24	159.45***	24	98.76***	24	25.06

Notes: [a]Modified Box–Pierce statistics for the rental series under different levels of differencing over the estimation period. * indicates significance at a 10% level, ** at a 5% level and *** at a 1% level.

there is strong evidence of nonstationarity. Only, when the second differenced series is examined, is there consistent evidence of stationarity.[4]

The second phase of the ARIMA modelling is concerned with the estimation of *p* and *q*, which are estimated using both the autocorrelation and partial autocorrelation

[4] These results are also confirmed by the autocorrelation functions. Stationarity necessitates that the autocorrelation function should decline geometrically to zero. This does not occur in either of the first two cases, however for the second differenced series does tend to converge to zero, albeit in an oscillatory fashion.

Table 3. Autocorrelation and partial autocorrelation functions of stationary series[a]

Number of lags	Autocorrelation function	Partial autocorrelation function
1	0.17	0.17
2	0.11	0.08
3	0.11	0.09
4	−0.09	−0.13
5	−0.07	−0.06
6	−0.38	−0.38
7	−0.26	−0.15
8	−0.33	−0.30
9	0.03	0.25
10	−0.04	−0.10
11	−0.05	0.02
12	0.14	−0.15

Notes: [a] Autocorrelation and partial autocorrelation functions over 12 lags for the rental series over the estimation period.

functions. The Autocorrelation Function (ACF) plays a major role on modelling the dependencies among the observations as it characterizes, together with the mean and variance, the stationary stochastic process describing the evolution of the series. Thereby, indicating the length and strength of the process, by measuring the extent to which one value of the process is correlated with previous values. While the MA order can often be directly observed from this correlogram, problems of ambiguity necessitates that the Partial Autocorrelation Function (PACF) be used to infer the order of the AR process. The order of the AR process is the value of p beyond which the partial autocorrelations are insignificantly different from zero. The two functions, up to lags of 12, are reported in Table 3. The objective criteria are imprecise, an imprecision which is magnified by the fact that the ACF and PACF are only sample representations of the population values. Whilst both functions exhibit *damped sine wave behaviour*, their combined patterns do not accord with any strict model order.

Because of the inconclusive results obtained, a general class of models is proceeded with and goodness-of-fit and diagnostic tests executed to isolate the 'correct' model. Eight tentative models are thus identified, the choice of which was guided by the Box–Jenkins principles of parsimony, with the highest order model comprising only two p and q elements. As Kennedy (1998) illustrates, a combination of small p and q values can capture an amazingly wide variety of time series structures. To supplement the visual model selection rules the Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC) are employed to ensure that the most accurate model is selected from the class of possible models. The two selection rules are calculated as:

$$AIC = T\ln(RSS) + 2n \tag{4}$$

$$SBC = T\ln(RSS) + n\ln(T) \tag{5}$$

Table 4. AIC and SBC statistics for alternative ARIMA models autocorrelation and partial autocorrelation functions of stationary series[a]

Number of lags	Akaike Information Criterion	Schwartz Bayesian Criterion
AR (1)	3.8874	3.9754
MA (1)	3.9471	4.0376
AR (2)	4.0291	4.1638
MA (2)	4.0335	4.1682
ARMA (1,1)	4.0293	4.1640
ARMA (2,1)	4.1194	4.2989
ARMA (1,2)	4.1175	4.2971
ARMA (2,2)	3.9622	4.1867

Notes: [a]AIC and SBC results for the alternative ARIMA based models over the estimation period. The lower the statistic the better fitting the model is.

Where T is the sample size, n is the number of regressors and RSS is the residual sum of squares. The most accurate model is that which emerges with the smallest figure on both criterion. Although, McGough and Tsolacos (1995b) and Tse (1997) both rely solely on the AIC, this paper shifts the emphasis more onto the SBC. As Mills (1990) notes, the SBC is strongly consistent, determining the true model asymptotically, whereas the AIC will always select an overparameterized model, regardless of the length of the available realization. The results reported in Table 4, show that according to both criteria, the AR(1) model emerges with smallest criteria. It is therefore, selected as the final model.[5]

Diagnostic tests are conducted to assess the adequacy of the AR(1) model. Adequacy here is concerned with ascertaining whether the residuals of the series evolve as a structureless white noise process. If the errors of the model are white noise, the modified Box–Pierce statistic is distributed approximately as a χ^2 distribution, with h–m degrees of freedom, where h is the number of lags used in the statistic and m is the number of fitted parameters. Table 5 illustrates that the error sequence does accord to a χ^2 distribution at a 5% level, with all calculated Q values failing to exceed their respective critical values.

3. Single equation model

The single equation model seeks to isolate the most fundamental forces driving the supply and demand for offices. Herein, rents are assumed to be a linear function of a series of independent variables, each proxying these forces. Providing certain assumptions are met a standard OLS approach can be used for this purpose. In this situation, the most important of these assumptions are that the error term exhibits uniformity of variance (homoscedasticity), insignificant autocorrrelation and that the equation is

[5] A viable alternative further step would be to evaluate the forecasting performance of all the possible models against a holdout set, and then use this as a further selection criteria. However, the small sample size would force the use of a post May 1996 holdout period

Table 5. Modified Box–Pierce statistics to assess AR(1) model residuals

Lag	Q
2	0.31
3	0.63
4	0.98
5	0.98
6	6.64
7	7.64
8	12.44
9	12.98
10	13.09
11	13.30
12	14.78
13	16.08
14	20.13
15	20.86
16	21.04
17	22.09
18	22.13
19	22.50
20	22.93

Notes: [a]A test of the residuals of the AR(1) model using the Modified Box–Pierce statistic.

correctly specified. Appropriate diagnostic tests will illustrate whether or not compliance with these assumptions is achieved.

The initial set of independent variables is selected according to the existing literature, with a total of ten series analysed. Wherever the series is provided in monetary terms, it is deflated by the Retail Prices Index (RPI). The ten series used are Gross Domestic Product, GDP of the Banking, Finance, Insurance and Leasing Industry, the value of construction contracts for new office space within London, interest rates, employment in Banking, Finance and Insurance (SIC 8) in South East England, changes in building costs, the number of property transactions, UK company gross trading profits and finally the Central Statistics Office (CSO) Longer and Shorter Leading indicators. Each of the variables is discussed in more detail below, including the rationale as to their inclusion.

As demand for office space is a function of the demand for goods and services produced by occupants of this space, a proxy is required to represent this demand. GDP representing the total value of economic activity has been consistently and successfully used for this purpose. In this study changes in GDP are used. In addition, a more specific measure of output is also assessed. Overall GDP captures the effect of economic growth in the whole economy on the office sector, including economic sectors such as agriculture. Therefore, the more specific measure, purely examining the services sector is also used. The importance of such focused variables was confirmed in McGough and

Tsolacos (1994), who found that in their VAR model, the output of business services and the output of financial and business services proved more significant than overall GDP. Additionally, Wheaton *et al.* (1997) report that 75% of the space in larger office buildings in the USA is occupied by firms whose SIC code is Finance, Insurance, Real Estate, Business or Professional Services. The comparable industry code in the UK encompasses similar sectors. Statistics were however only available on an annual basis, therefore to obtain semi-annual measures the series was interpolated.

The third series to be used measures construction contracts for London office space. In order to obtain a measure proxying the supply of office space, this series is deflated by the Commercial Building Cost Index to provide a representation of the amount of new office space entering the market. Although this proxy will not capture the supply of new space from existing stock and the possibility of incomplete contracts, it should provide a relatively accurate measure of overall market movements. The fourth measure is interest rates, which play an integral part in the evolution of the economic and office cycle. The proxy of interest rates used is the LIBOR rate, deflated by the RPI. The next variable is employment in SIC8 industries and is used to capture economic activity in the sectors most pertinent to the office rental cycle. While the majority of European studies have reported that such measures fail to exert a significant influence, Wheaton *et al.* (1997) argue that it is the major driving force behind office demand. The strong theoretical appeal therefore leads to its inclusion.

The next three variables are changes in building costs, the number of property transactions and UK Company gross trading profits. The inclusion of changes in building costs is based on the hypothesis that higher building costs should translate into higher rents, a phenomenon that would be expected to occur as a development boom nears its peak, placing greater emphasis on limited resources. Furthermore, this variable should assist in the forecasting of new office construction, and considering the significant results obtained in the structural model, its inclusion should aid in providing more accurate rental forecasts. The property transactions variable measures the number of property transactions and its inclusion is designed to capture some measure of sentiment in the investment market. It is a national sector-wide index, and thus some allowance should be made in terms of potential contamination that may occur. Specifically, it would be expected that the number of transactions would decrease as market sentiment turns negative. The gross trading profits of UK companies is included in order to capture any demand side effects that the GDP, employment or indicator variables may have missed. It is a more focused measure, concentrating solely on profits of firms and not output. As with a number of variables used in the model, it is deflated by the RPI.

The final two series that form the initial set of independent variables are the CSO Longer and Shorter Leading indicators. These measures aim to capture expected business and economic conditions. The conventional measures of economic activity are purely contemporaneous, however, firms will also incorporate expectations into their decisions regarding real estate requirements. This factors potential importance is magnified in markets such as the UK; due to the long lease structure commonly in place. It would be expected that a positive contemporaneous relationship would emerge. Any lagged values that provide significant coefficients would merely reflect the current health of the economy and not capture the intended influence. McGough and Tsolacos (1994) used both of these variables in their use of the VAR model.

The initial model can be formulated as follows:

$$
\begin{aligned}
r_t = \alpha &+ \sum \beta gdp_{t-i} + \sum \gamma gdpbfi_{t-i} + \sum \lambda nofc_{t-i} + \sum \delta libor_{t-i} \\
&+ \sum \xi emp_{t-i} + \sum \kappa bc_{t-i} + \sum \psi pt_{t-i} + \sum \nu cp_{t-1} + \sum \varpi lli_{t-i} + \sum \kappa sli_{t-i} + \varepsilon
\end{aligned} \tag{6}
$$

where r is the rental series, *gdp* is changes in overall real GDP, *gdpbfi* is the change in real GDP for the service sector, *nofc* represents the construction series, *libor* is the real interest rate series, *emp* is service employment in the South East, *bc* is changes in building costs, *pt* is the number of property transactions, *cp* is gross company trading profits, *lli* is the CSO Longer Leading Indicator and *sli* is the CSO Shorter Leading Indicator. The real interest rate series was adjusted using actual inflation. While it may be argued that the use of expected inflation would be more appropriate due to need to account for inflationary expectations in financing decisions, there is a degree of subjectivity when estimating expected inflations. In a real estate context studies of inflation hedging have utilized a wide variety of proxies for expected inflation. These include ARIMA based models, autoregressive models and published inflation forecasts. Studies such as Stevenson (2000) find that results can differ substantially depending on the proxy used. In addition, the expected inflation figures obtained using the different proxies also can vary substantially. In order to avoid the variable being dependent on the proxy chosen, actual inflation was used to adjust the interest rate. In terms of the number of lags initially analysed, the maximum length used was dictated by previous literature. For the two GDP series', the employment variable and company profits a maximum allowable lag length of four was used, while for the remaining variables, six lags were initially examined for the construction, property transaction and building cost series', two for real interest rates and one for the CSO leading indicators.

To avoid multicollinearity some method is required to pare down this initial equation and to isolate the most influential variables. D'Arcy *et al*. (1997a) used the Wald variable deletion test, while D'Arcy *et al*. (1999) employed t-statistics. In this study a simple stepwise regression procedure is used. This is an iterative procedure, with each stage involving a search for the variable with the highest explanatory power. F-tests for joint significance are then calculated with variables exceeding this test at a 5% level of significance being entered into the final equation. Recognizing that the addition of a new variable could impact upon the explanatory power of all variables in the presence of the addition, variables failing this F-test at a 10% significance level are removed form the final model.

The stepwise approach effectively provides that model with the greatest explanatory power in-sample. It can be criticized on the basis that the approach may not address multicollinearity adequately, while Davidson and MacKinnon (1993) also criticize its use in the stead of specifications grounded in economic theory. However, in situations such as commercial rent modelling, its application can be justified due to the fact that empirical evidence regarding the theoretical frameworks so far established remains in its infancy. Theory dictates that rents are a function of the economy's demand for space and current supply of space to the market. How to best proxy these forces remains open to a large degree of debate. This is particularly so in the case of a large number of variables and alternative lag lengths. However, in order to ensure that diagnostic and specification problems are minimized, a variety of diagnostic tests are undertaken on the final specification obtained in order to ensure that no severe problems result from the use of the stepwise approach.

Table 6. Final structured equation specification and diagnostics[a]

Panel A: Specification of structured equation

Variable	Coefficient	*t*-statistic
Intercept	−42.883	2.032
EMP	0.12437	4.358
EMP (lag 1)	0.12097	4.139
EMP (lag 3)	0.11747	4.443
NOFC (Lag 5)	−11.514	2.926
LLI	0.64197	3.553

Panel B: Ramsey RESET specification tests

Proxy	RESET value	Degrees of freedom	Critical value 5%
Y^2	0.21433	1,31	4.16
Y^3	0.49784	2,30	3.32
Y^4	1.1352	3,29	2.94

Panel C: Tests for heteroscedasticity

Test	Test statistic	Degrees of freedom	Critical value 5%
Breusch-Pagan-Godfrey	4.049	5	11.07
Glejser	6.599	5	11.07
ARCH	0.028	1	3.84

Notes: [a]Final specification of the structured model. The model is estimated using OLS and a stepwise procedure to pare down the initial set of variables. The table also presents test statistics in relation to specification and heteroscedasticity.

The final specification of the model is reported in Table 6, and as can be seen, it uses both demand and supply variables. On the demand side, employment in the SIC8 sectors influence both contemporaneously and at lags one and three. Their influence is remarkably strong, subsuming both of the GDP variables. This result is theoretically pleasing due to the variable's status as the only one to specifically measure economic activity within the Greater London region, and its concentration on sectors most pertinent to the office sector. Furthermore, as all three sectors are involved in the provision of services their financial health depends on the fortunes of their customers and clients. Therefore, the variable may be acting as a barometer for the health of the overall economy, thereby offering further reasons as to the non-inclusion of the GDP variables.

Previous studies have long proclaimed the explanatory power of service sector employment on office rent determination, albeit with little empirical support. The results contained in this study add further weight to recent studies to have found a significant influence for such a variable. In the current paper the final specification included the first, second and third lag of BFI employment. Similarly, the emergence of a significant negative influence from lagged values of the office construction variable is consistent with *a priori* expectations. Although theory has long obviated the importance of supply-side variables, data availability has limited the degree of empirical evidence. The significance of the fifth lag may appear somewhat distant, but it is not without its theoretical and empirical support. The variable tested is based on the value of new contracts awarded for new construction, thus the actual development process' have not yet commenced. At the very minimum, 18 months would usually be required to fully complete any reasonably large office development. The inclusion of the CSO Longer Leading Indicator in the initial set of variables was motivated by a desire to examine the potential impact of economic expectations on rental values. The results also confirm the findings of McGough and Tsolacos (1994) in their VAR model. Of the variables that failed to enter the final specification, the ones of most interest are the national GDP measures and interest rates. The two GDP variables may have suffered to some degree due to their national focus, especially when a more localized measure was available in the form of service sector employment in the South East. In addition, the Banking, Insurance and Real Estate GDP measure may have been affected by the need to interpolate the series in order to obtain a semi-annual variable. In relation to the lack of significance observed with regard to interest rates it is possible that this is due to the choice of the LIBOR as the proxy.

An adjusted R^2 of 82.9 was obtained for the final model, demonstrating the powerful explanatory ability of the model, especially considering the fact that changes in the index were being modelled. The model accurately tracks the evolution of the rental index, with the model capturing both the speed and direction of all major movements very accurately. The one exception is that the increase in rental values in the late 1980s is slightly underestimated by the model, which, due to the relatively unique nature of that boom, is not entirely unexpected.

To test the validity of the model and the results, diagnostic tests were executed. To examine potential functional misspecification the Ramsay RESET test is used. This tests for omitted variables and the existence of non-linearities, both of which would be contrary to OLS' BLUE qualities, rendering estimates unreliable and inferences infeasible. The RESET test revolves around using the second, third and fourth powers of the predicted dependent variable as a proxy for any omitted variables and adding this proxy to the original set of explanatory variables. An F-test is then used to test the proxy's set of coefficients against the zero vector. The results are reported in Table 6 and show that all of the RESET values fail to exceed their respective critical values. A common problem of time series models is autocorrelation, referring to the presence of non-zero off-diagonal elements in the variance–covariance matrix of the disturbance terms. One of the most frequently observed reasons for this is the prolonged influence of shocks, continuing to reverberate for a number of periods. If the residuals are not independent, therefore displaying autocorrelation, the employment of F and t tests and confidence intervals are not strictly valid and estimates of the coefficients may be subject to instability. To investigate the presence of autocorrelated disturbances the Durbin–Watson test is used, which can be displayed as follows:

$$DW = \frac{\sum_{t=2}^{N} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{N} \varepsilon_t^2} \tag{7}$$

The test statistic falls between two and four, and in this case, with six explanatory variables including the constant, and 38 observations, an upper test statistic d_U of 1.85, and a lower test statistic d_L of 1.16 are estimated. The reported value is 1.74, therefore falling into the region of indeterminacy, where the presence of autocorrelation cannot be confirmed or denied. Pindyck and Rubinfeld (1998) state that this occurrence can be attributable to autocorrelation in the independent variable and not in the error term.

A second condition of the OLS assumption of nonspherical disturbances is that the diagonal elements in the variance–covariance matrix of the error term are all equal, i.e. homoscedasticity. Violation of this assumption will again obscure the validity of the statistical tests and confidence intervals. A range of tests are available to investigate for the presence of heteroscedasticity, with a total of three such tests employed in this paper. The first such test is the relatively broad Breusch–Pagan–Godfrey test, which tests the hypothesis that the variance is some function of a linear combination of known variables against a null of homoscedasticity. The second test is the Glejser tests in which the absolute values of the residuals are regressed on the independent variables. A significant non-zero coefficient would confirm the presence of heteroscedasticity. The final form is an ARCH, Autoregressive Conditional Heteroscedasticity model. The results for the three diagnostic tests are displayed in Table 6 and it can be seen that in each case the test statistics fail to exceed their critical values, thereby failing to reject the null hypothesis of homoscedasticity.

4. Bayesian VAR model

The Vector Autoregression (VAR) model is built upon the postulation that all variables in a system of equations are endogenous and that each can be depicted as a linear function of its own lagged values and the lagged values of all other variables in the system. The basic form of this model can be formulated as follows:

$$\mathbf{Y}_t = \alpha_0 + \alpha_1 \mathbf{Y}_{t-1} + \alpha_2 \mathbf{Y}_{t-2} + \ldots + \alpha_n \mathbf{Y}_{t-n} + \varepsilon_t \tag{8}$$

Where $\mathbf{Y}$ is the vector of all variables in the equations and n is the desired lag length. Thus a simultaneous equation system is constructed, enabling continuous one-step ahead forecasts for all variables within the model. Unrestricted VAR models often suffer from overfitting, a problem that emerges from the fact that the number of observations typically available is inadequate for the number of coefficients to be estimated. If degree of freedom exhaustion has not already created problems with the model, the statistical procedure will not only choose coefficients to explain the salient, enduring features, but will also fit some of the too numerous coefficients to random, non-recurring, relationships. This over-fitting thereby, often leads to large out-of-sample errors and renders the VAR overly sensitive to changes in the variables. The alternative method, and the one used in this study, is to adopt a Bayesian approach. Bayesian VARs (BVARs) seek to improve economic modelling and forecasting by offering a more flexible, yet objective, method to tackle the specification problems of the arbitrary procedures adopted hereto. BVARs utilize prior statistical and economic information to construct restrictions on the coefficients that can then be overridden by the data. All of the data is initially modelled in levels. The use of levels, rather than differenced series, is due to the loss of information that accompanies differences, for example, cointegrating relationships. In addition, evidence would suggest that differencing data provides no improvement in asymptotic efficiency.

The methodology used in this study broadly follows the Minnesota prior. The first step in such an approach is to quantify prior beliefs regarding the value for each of the coefficients. These are expressed in the form of probabilities regarding which value will generate the best forecasts. In this study a relatively loose prior variance of 0.25 was imposed thus permitting large deviations between the best coefficient and the coefficient guesses. This appears reasonable, given that many of the series, including rent, should be somewhat forecastable and secondly, that lags of many of the variables would be expected to exert some influence. Thus it would not be justifiable to place too much confidence in the best estimates. Generally as lag length increases, the less importance the role the lagged variable plays in forecasting. Thus, to ensure less noise pollution by these variables, the Minnesota prior advocates that prior variances should become tighter as the lag increases. This would ensure that lagged variables remained closer to their best guess coefficients of zero. However, as already witnessed in the structural model, the fifth lag of new office construction plays an influential role in explaining rental movements. Therefore, it was decided to reduce the standard decay factor in the prior variance from $1/(k+1))$, where k is the length of the lag, to 0.05 per extra lag.

The second stage of the modelling of the BVAR is to specify the cross weights, the influence of variable j, and its lags, on variable i. Two alternatives are available. First, a symmetric prior, placing equal weight on the interaction between each and every variable, could be constructed, thus placing no demands for information regarding the influence, for example, that building cost movements would have on the long run indicator. On the downside, however, it would mean assuming that, for example, office rents were just as significant a factor in explaining GDP as GDP was in explaining rents. Obviously, this is unreasonable. Symmetric priors like this can be of use in explaining market microstructures, where the variables could have a reciprocal relationship with each other, but their use is limited here, where the prediction of a single variable is the sole focus, and reciprocal relationships are not expected. Hence, a general prior must be constructed, specifying the relationship between each and every variable. Exact specification of all of the relationships is beyond the bounds of this study, and hence approximations based on economic theory are employed.

The initial priors are displayed in Table 7, with the weights assigned aimed at depicting how likely the coefficients of the respective variable are to deviate from their prior mean of zero. In order to reduce the discretionary nature of the assignment of these weights, the Theil U statistic[6] is used. This method incorporates running three systems of equations: a system of univariate OLS models, a system of Bayeisan autoregressions and a standard symmetric BVAR. Using the first 22 observations a forecast for the 23rd observation is calculated using each of the three systems. Using the actual observation, Theil U statistics are calculated for each method. A variable is added each time for each of the remaining observations. The Theil U statistic evaluates how each method performs against a naive method, namely one that assumes no change. The forecast errors are squared, giving greater

[6] Alternative methods include the use of Kalman filers to generate one-step ahead forecast errors and using the log of the determinant of the sample covariance matrix of these errors as a measure of fit or alternatively using the likelihood function of the data conditional on the hyperparameters.

Table 7. Initial prior[a]

	RENT	NOFC	GDP	GDP BFI	EMP	LLI	LIBOR	PROFITS	TRANS	BC	SLI
RENT	1	0.8	0.4	0.4	0.8	0.8	0.2	0.2	0.2	0.2	0.2
NOFC	0.5	1	0.5	0.5	0.5	0.2	0.2	0.2	0.2	0.35	0.2
GDP	0.1	0.1	1	0.3	0.3	0.3	0.35	0.35	0.1	0.1	0.3
GDP BFI	0.1	0.1	0.3	1	0.35	0.2	0.35	0.35	0.2	0.1	0.2
EMP	0.1	0.1	0.3	0.35	1	0.2	0.2	0.2	0.2	0.1	0.2
LLI	0.1	0.1	0.35	0.35	0.2	1	0.35	0.35	0.1	0.1	0.35
LIBOR	0.1	0.1	0.2	0.1	0.1	0.2	1	0.1	0.1	0.1	0.1
PROFITS	0.1	0.1	0.35	0.35	0.1	0.2	0.2	1	0.1	0.1	0.35
TRANS	0.25	0.35	0.2	0.2	0.1	0.2	0.2	0.1	1	0.2	0.1
BC	0.1	0.5	0.2	0.2	0.2	0.1	0.1	0.1	0.1	1	0.1
SLI	0.1	0.1	0.35	0.35	0.2	0.5	0.35	0.35	0.1	0.1	1

Notes: [a]Initial priors used in the estimation of the Bayesian Vector Autoregression model.

weight to the disproportionate cost of large errors. The statistic can be formulated as:

$$U = \left\{ \frac{\sum_{t=1}^{n-1} (FPE_{t+1} - APE_{t+1})^2}{\sum_{t=1}^{n-1} (APE_{t+1})^2} \right\}^{1/2} \tag{9}$$

where FPE represents the forecast relative change and APE the actual relative change. Thus, when the statistic is less than unity, the forecasting method is superior to the naive approach while if it is greater than one the forecasting method is inferior. The procedure centres around two rules. First, that if the OLS statistics are superior than in the univariate BVAR, then the unitary value applied to the own lags is too tight and needs to be relaxed. Secondly, that if the U statistics in the univariate BVAR are superior to the symmetric BVAR, then less emphasis should be applied to the cross lags, necessitating a reduction in the off-diagonal elements. The results, displayed in Table 8, show that the U statistics for the univariate BVAR are superior to the OLS, but are in turn inferior to the symmetric BVAR. Hence, the unitary value for own lags is maintained, while the weights applied to other variables are reduced. In the application of the first rule, the unitary value is increased to 1.5, and in the application of rule 2, the weights are deflated by a factor of 0.25. The modified priors are shown in Table 9. As with the ARIMA model, out-of-sample forecasting performance would be a desirable further step in this process. This would have permitted the estimation of a vast quantity of priors and the fine-tuning of individual weights to minimize out-of-sample forecast errors. However, as with the ARIMA model, the limited number of observations, combined with the possibility of introducing pre-competition bias, rendered this step unfeasible. To obtain the forecasts the system of equations is estimated using Theil's mixed estimation technique.

Table 8. Theil U statistics[a]

Variable	OLS	UBVAR	SBVAR
RENT	1.852	0.9877	0.8927
NOFC	3.395	1.0737	1.503
GDP	0.5253	0.5833	0.7626
GDP BFI	0.45408	0.5832	1.2816
EMP	0.6845	0.7415	0.919
LLI	1.2755	1.092	0.6526
LIBOR	0.56627	0.5913	0.7803
PROFITS	0.7363	0.7756	0.7964
TRANS	1.949	0.84977	1.2387
BC	0.5115	0.48492	0.6259
SLI	0.60934	0.8546	0.801

Notes: [a]Theil U test statistics for the system of OLS equations, the Bayesian autoregressions and the standard symmetric VAR model. The statistic is used to assesses priors used in the BVAR model.

Table 9. Adjusted prior[a]

	RENT	NOFC	GDP	GDP BFI	EMP	LLI	LIBOR	PROFITS	TRANS	BC	SLI
RENT	1	0.6	0.3	0.3	0.6	0.2	0.15	0.15	0.15	0.15	0.15
NOFC	0.376	1	0.375	0.375	0.375	0.15	0.15	0.15	0.15	0.2625	0.15
GDP	0.075	0.075	1	0.225	0.225	0.225	0.2625	0.2625	0.075	0.075	0.225
GDP BFI	0.075	0.075	0.225	1	0.2625	0.15	0.2625	0.2625	0.15	0.075	0.15
EMP	0.075	0.075	0.225	0.2625	1	0.15	0.15	0.15	0.15	0.075	0.15
LLI	0.075	0.075	0.2625	0.2625	0.15	1	0.2625	0.2625	0.075	0.075	0.2625
LIBOR	0.1	0.1	0.2	0.1	0.1	0.2	1	0.1	0.1	0.1	0.1
PROFITS	0.075	0.075	0.2625	0.2625	0.075	0.15	0.075	1	0.05	0.075	0.2625
TRANS	0.182	0.2625	0.15	0.15	0.075	0.15	0.15	0.075	1	0.15	0.075
BC	0.075	0.375	0.15	0.15	0.15	0.075	0.075	0.75	0.075	1	0.075
SLI	0.1	0.1	0.35	0.35	0.2	0.5	0.35	0.35	0.1	0.1	1

Notes: [a]Adjusted priors used in the estimation of the Bayesian Vector Autoregression model.

5. Simultaneous equation model

The final model to be analysed is the simultaneous model proposed by Wheaton *et al.* (1997), which is an extension of that developed by Wheaton (1987). The advantage to such an approach is that while the structured equation focuses merely on one aspect of the larger system of real estate investment, a simultaneous approach extends this focus to incorporate all of the interrelationships and interactions within the system. However, such theoretically-based models can limit the flexibility in modelling, especially in a forecasting sense. In addition, as with both the single equation model and the BVAR, the spatial scales of the variables used is not necessarily in line with the rents used. This fact should be considered, not only in the simultaneous model but also in the previous two alternatives.[7] The model purports that the demand for occupied space is a function of service sector employment and past returns. Service sector employment is the major demand-side force in this model, which is consistent with the results already obtained in the structural model. However, the fixed lease environment that was prevalent for much of the period under study would limit firms in achieving their desired level of office space. This desired level of occupancy would eventually emerge if employment and rents are held constant, and thus predicted values from the following model can be viewed as one-step ahead forecasts.

$$OS_t = a_0 + \mathrm{E}_t(a_1 + \mathrm{a}_2 R_{t-1}) \tag{10}$$

where OS is occupied office stock, E is service sector employment and R is rents. Estimation of the model yields the following results.

$$Os_t = 39335000 + \mathrm{E}_t(174.57 - 0.18685\, R_{t-1})$$

Both coefficients take their expected signs, with increases in employment exerting a positive influence on the desired level of occupied stock and with increases in rent having a negative impact. In each period some fraction of firms, F, will adjust their space consumption to long run demand, as defined by the predicted values in the above model. This can be formulated as follows:

$$OS_t - OS_{t-1} = AB_t = F_1[OS_t^* - OS_{t-1}\} \tag{11}$$

where OS^* is the desired, or predicted, level of occupancy and AB is the absorption rate. Therefore, F can be calculated as the ratio of actual absorption to desired absorption, or the percentage of the market that moves towards the equilibrium level in any one period. Combining the above equations generates a linear equation in which absorption and the occupied stock progressively move towards a target defined by office employment and rents.

$$AB_t = F_1 + (a_0 + E_t[a_1 + a_2 R_{t-1}]) - F_1\, OS_{t-1} \tag{12}$$

[7] Hendershott *et al.* (1999) provide an alternative simultaneous model specifically for the City of London office market. The primary difference between the two models is the direct link between with the capital markets through a time-varying equilibrium rent. Hendershott *et al.* (1999) also use effective rather than headline rents in their model. Ball *et al.* (1998) explicitly compare the two alternative models.

Table 10. Specification and diagnostics for absorption model[a]

Panel A: Specification

Variable	Coefficient	t-statistic
Intercept	1 820 200	4.59
	1.048	3.592
	166.71	4.086
	0.0278	3.500
	−0.31954	3.957

Panel B: Ramsey RESET specification tests

Proxy	RESET value	Degrees of freedom	Critical value 5%
	4.889	1,31	4.16
	10.332	2,30	3.32
	10.442	3,29	2.94

Panel C: Tests for heteroscedasticity

Test	Test statistic	Degrees of freedom	Critical value 5%
Breusch–Pagan–Godfrey	17.933	4	9.49
Glejser	16.844	4	9.49
ARCH	0.029	1	3.84

Notes: [a]Estimated coefficients from the absorption equation in the Wheaton *et al*. (1997) model. Details are also provided with regard to specifification and heteroscedasticity.

The results from the estimation of this equation are reported in Table 10, and it can be seen that all coefficients emerge with the expected sign and are statistically significant at conventional levels. A constant was however required to stabilize the equation, while the R^2 of 41% suggest that the specification failed to fully capture the absorption series. More serious signs appear in the form of the specification and diagnostic tests. As with the single equation model the Ramsey RESET test, the Durbin–Watson test for autocorrelation and the three alternative tests for heteroscedasticity are used. Table 10 also reports the specification and heteroscedasticity tests. It is clear that in each case the RESET test statistic exceeds the appropriate critical value, while the DW statistic of 1.1508, when compared to a lower limit of 1.26 provides compelling evidence of autocorrelation. In addition, the test statistics for both the Breusch–Pagan–Godfrey and Glejser tests exceed their critical values. Conventional techniques such as conversion to logs and first differences failed to optimally address these violations, therefore to ensure correct parameter estimation and the validity of the inference techniques, the equation was re-estimated using the Newey–West autocorrelation and heteroscedasticity consistent estimator.

The model stresses the importance of vacancy rates in the process of rent determination. This approach is not without its critics, for example D'Arcy *et al*. (1999), who highlight that the endogenous nature of vacancy rates in a rent equation could lead to spurious results. The equilibrium rent is modelled as:

$$R^* = a_0 - a_1 V_{t-1} + a_2(AB_{t-1}/OS_{t-1}) \tag{13}$$

where V represents the vacancy rate. If vacancy and absorption were to remain fixed, this equilibrium level of rent would eventually emerge in the market, implying that the predicted values of rent in this equation can be interpreted as the long-run equilibrium, which will hold until vacancy and absorption change. The estimation of this equation yields:

$$R^* = 139.23 - 171.86V_{t-1} - 88.362(AB_{t-1}/OS_{t-1})$$

Similar to the absorption specification, a certain percentage of market rents, K, will move towards this rent in the following period. This can be deduced from the following.

$$R_t - R_{t-1} = K[R^* - R_{t-1}] \tag{14}$$

Combining these two models leads to a specification for the change in rental values:

$$R_t - R_{t-1} = K[a_0 - a_1 V_{t-1} + a_2(AB_{t-1}/OS_{t-1})] - K\Delta R_{t-1} \tag{15}$$

The term in parentheses represents the equilibrium rent, which is a function of the preceding vacancy rate and absorption as a percentage of occupied stock. This absorption rate feeds in from the previous equation. Hence, an increase in office employment will boost absorption, and through this will lower the vacancy rate. Providing that new office supply does not respond accordingly, this will lead to an increase in the equilibrium rental level. Table 11 provides details of the estimated coefficients of this model. While an R^2 of 71% demonstrates the good fit of the model, both the vacancy and absorption variables fail to achieve statistical significance. In addition, the sign of the vacancy rate is not in line with expectations, while the diagnostic tests, again reveal the presence of misspecification and the presence of autocorrelation in the disturbance terms, with the DW statistic outside of the lower boundary at 1.0038. The Newey–West estimator was used again to address the problem of autocorrelation.

The final equation in the system depicts the new office development process. This portrays new construction as a linear function of rents, vacancy rates, interest rates and building costs, and can be formulated as follows:

$$C_t = b_0 + b_1 R_t - b_2 V_t - b_3 I_t - b_4 RC_{ti} \tag{16}$$

where C represents new construction, I is interest rates and RC is real construction costs. Therefore, while rents and vacancy rates capture demand-side forces the supply-side is represented by building costs and interest rates. The estimation of the equation yielded the results reported in Table 12. An R^2 of 91% illustrates the explanatory power of the model, while all of the coefficients take the expected sign, with the exception of building costs, whose coefficient also failed to achieve statistical significance. Furthermore, the diagnostic tests fail to reveal any major problems with the equation. The Durbin–Watson statistic obtains a value of 1.4334, falling into the indeterminacy area.

Table 11. Specification and diagnostics for simultaneous equation rent model[a]

Panel A: Specification			
Variable		Coefficient	t-statistic
Intercept		2.0039	1.137
		−8.5453	3.592
		−61.945	0.3363
		−35.2364	0.0463
		−0.77628	7.296

Panel B: Ramsey RESET specification tests			
Proxy	RESET value	Degrees of freedom	Critical value 5%
	4.8928	1,31	4.16
	4.2002	2,30	3.32
	3.5722	3,29	2.94

Panel C: Tests for heteroscedasticity			
Test	Test statistic	Degrees of freedom	Critical value 5%
Breusch–Pagan–Godfrey	1.636	4	9.49
Glejser	3.749	4	9.49
ARCH	1.937	1	3.84

Notes: [a]Estimated coefficients from the equilibrium rent equation in the Wheaton *et al.* (1997) model. Details are also provided with regard to specification and heteroscedasticity.

6. Forecasting performance

This section of the paper initially details the forecasting ability of each of the four models over the three-year period to May 1999. The ARIMA and Bayesian VAR models are relatively easy to forecast. For the ARIMA model point estimates are made for six periods, while for the BVAR, the system of equations is estimated using Theil's (1971) mixed estimation technique. This method incorporates the extraneous information of the prior into classical GLS estimation. The forecasts for the rental series are then generated from this process. The single equation and simultaneous models are however, more complicated to use in a forecasting capacity due to the need to obtain forecasts for the explanatory variables. This study proceeds on the basis of using professional forecasts where available. When such data are not available, private forecasts are calculated. Makridakis *et al.* (1998) compare a number of alternative forecasting techniques, finding that the deseasonalized exponential smoothing method generally provides the most accurate results. D'Arcy *et al.* (1999) use a similar method, bequeathed of the seasonal component, known as the double exponential smoothing technique. Because of the strong evidence as to its forecasting performance it is the model used throughout this study.

Table 12. Specification and diagnostics for construction model[a]

Panel A: Specification

Variable	Coefficient	t-statistic
	61.458	1.151
	3.8297	15.73
	−1330.5	5.504
	−31.456	7.121
	0.16242	0.4737

Panel B: Ramsey RESET specification tests

Proxy	RESET value	Degrees of freedom	Critical value 5%
	5.4761	1,31	4.16
	3.0331	2,30	3.32
	1.9585	3,29	2.94

Panel C: Tests for heteroscedasticity

Test	Test statistic	Degrees of freedom	Critical value 5%
Breusch–Pagan–Godfrey	8.441	4	9.49
Glejser	7.337	4	9.49
ARCH	2.034	1	3.84

Notes: [a]Estimated coefficients from the construction equation in the Wheaton *et al.* (1997) model. Details are also provided with regard to specifification and heteroscedasticity.

With regard to the single equation model, as new construction was only significant at the fifth lag, actual values can be used. Therefore, forecasts are only required for the employment and leading indicator variables. The National Institute of Economic & Social Research provide business service employment forecasts, and these were available up until the fourth quarter of 1998. Although not strictly equal to the BFI series, it was felt that owing to the use of changes in employment, there was enough comparability between the two for it to be used. The forecasts accuracy is impressive; with an average one year forecast error of only 2.1%. Forecasting ahead from the first quarter in 1996, the NIESR predicted business service employment would increase by 1.438% to January 1997, a further 1.418% in the following six months and by 1.379% to the end of 1997. For the remaining three periods of the forecast, the double exponential smoothing method was used. This method revolves around the calculation of two smoothing variables (α and β) and three equations:

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + b_{t-1}) \tag{17}$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta) \tag{18}$$

$$F_{t+1} = L_t + b_t \tag{19}$$

where L denotes an estimate of the level of the series at time t, and b denotes an estimate of the slope of the series at time t. The first equation adjusts L directly for the trend of the previous period by adding to it the last smoothed value. This helps to estimate the lag and brings L to the approximate level of the current data value. The second equation then updates the trend, which is defined as the difference between the last two smoothed values. This is appropriate because if there is a trend in the data, new values should be higher or lower. Since there may be some randomness remaining, the trend is modified by smoothing with β the trend in the last period and adding it to the previous estimate of the trend multiplied by $(1-\beta)$. The final equation simply produces the one-step ahead forecast. To calculate the two coefficients a dataset of the 20 observations directly preceding June 1996 and a range of coefficient values with the goal of minimizing the mean squared error between the forecasts and actual values are used. A value for α of 0.90 was obtained, and 0.50 for β. For the Longer Leading Indicator, the same method is used for the whole the three-year period, yielding estimates of 0.70 for α and 0.85 for β. For the simultaneous model, forecasts are required for BFI employment, interest rates and building costs. Forecasts from the NIESR are again used where available, in this case for interest rates, while the employment variable is the same as used in the structural model. For building costs the double exponential smoothing model is used. In order to forecast the simultaneous model, the absorption model initially is calculated, the output of which enters the second equation for rent, via the absorption and vacancy variables.

Table 13 and Fig. 1 provide details of the forecasts estimated for each of the four models. All four alternatives have tended to capture the directional movement in the index, however, all of the models, with the exception of the Bayesian VAR model, fail to capture the large upward movement in rental values. This is particularly the case with the ARIMA model. In part the poor performance of the ARIMA may be due to the fact that the final model selected was a simple AR1, therefore not capable of picking up the large upswing in rental values that occurred during the out-of-sample period. The simultaneous model moved closely alongside the actual values during the first half of the test period, but fail to pick up the upswing in rents from mid-1997 onwards. The single equation model also performs well, although it displays a tendency to lag the movements in the index.

To more formally measure forecasting ability a variety of measures are used. Table 14 reports a variety of measures, while Table 15 presents the Theil's U statistics for each

Table 13. Comparison of forecasting performance[a]

Period	Actual	ARIMA	Structured	BVAR	Simultaneous
May 1996	100.0000	100.0000	100.0000	100.0000	100.0000
Nov 1996	104.0081	97.8323	106.7175	103.9190	103.4794
May 1997	107.2817	100.0240	109.9589	109.4223	106.9542
Nov 1997	118.8388	102.2710	113.6607	115.8853	110.8004
May 1998	128.3513	104.5750	117.4542	123.1507	114.6326
Nov 1998	131.7227	106.9330	122.5925	129.7427	118.3386
May 1999	135.3178	109.3470	128.1718	134.8162	122.7449

Notes: [a]Report of the actual index values over the forecast period May 1996–May 1999. The table also reports the forecasted index values using the four alternative models.

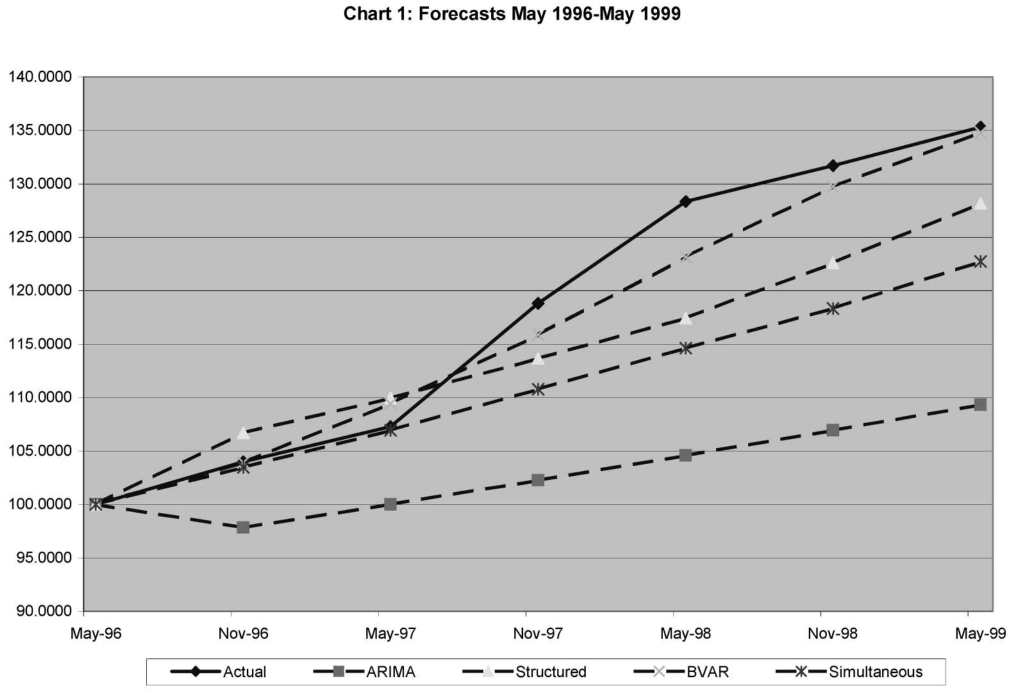

Fig. 1. Forecasts May 1996–May 1999: the comparative index figures for the actual movements in the CB Hiller Parker index together the forecasted figures from the four alternative models.

Table 14. Forecasting accuracy measures[a]

	ARIMA	Structured	BVAR	Simultaneous
Mean error	−17.4230	−4.4941	−1.4307	−8.0951
Mean absolute error	17.4230	6.2897	2.1442	8.0951
Mean squared error	369.9385	49.0823	7.4219	98.4029
Mean squared percent error	2.2351%	1.6128%	0.1589%	0.3830%
Error variance	66.3770	28.8850	5.3750	32.8731

Notes: [a]Details of five alternative measures of forecast accuracy for the four different models in relation to actual observed index figures.

alternative model.[8] It can be seen quite clearly how the ARIMA was the worst performing of the five models. Of the remaining models, the structured and simultaneous models obtained broadly similar results, however, the Bayesian-VAR was by far the most accurate of the alternatives. This result confirms the results of McGough and Tsolacos (1994) and provides strong encouragement for a model that has been relatively

[8] More formal statistical comparisons of the alternatives in limited due to the small number of forecasted observations.

Table 15. Theil's U statistics for forecasting models[a]

Period	ARIMA	Structured	BVAR	Simultaneous
Nov 1996	1.45	0.63	0.02	0.12
May 1997	0.96	0.35	0.28	0.04
Nov 1997	0.87	0.27	0.15	0.42
May 1998	0.83	0.38	0.18	0.48
Nov 1998	0.78	0.29	0.06	0.42
May 1999	0.73	0.20	0.01	0.35
Mean	0.935	0.354	0.119	0.306

Notes: [a]Theil's U statistic for each of the four forecast models relative to the actual index levels reported in the out-of-sample period.

under-utilized in the literature. The second most accurate model was the relatively simple structured model and despite the specification problems encountered in its estimation, the Wheaton *et al.* (1997) model did perform reasonably well. The results for the Theil U statistics are contained in Table 15, with the results further confirming the initial findings. The Bayesian VAR was by far the most accurate model, while with the exception of the ARIMA model in the first period; all of the models outperformed the naive strategy of no change in rental value.

The poor performance of the ARIMA has to be viewed in the light of the low number of observations available for its calculation, with only 38 observations available in this study. This shortfall in turn led to conventional model selection techniques, the ACF and PACFs, becoming redundant and forcing the emphasis to shift to using the AIC and SBC criteria and the somewhat arbitrary selection of eight models. Of considerable note is the ARIMA's poor short-term performance. Such models have long been heralded as capable one- and two-point estimators, a capability not proved in this study. Indeed, the models worst performing periods were the first two. These results do however confirm the findings of studies such as Makridakis *et al.* (1998). In addition, while the poor short-term accuracy of the ARIMA model is not in line with much of the previous empirical evidence in real estate, its poor long-term forecasting ability would be consistent with expectations due to the nature of the model. ARIMA models rely totally on past observations in the series examined and are therefore particularly vulnerable at market turning points. The upswing in the index in the forecasting period used in this paper is probably the primary reason as to why the ARIMA method provides the least accurate forecast over the three-year period.

Despite the specification problems in the estimation of the simultaneous model, the forecasting ability of the model was relatively strong. The accuracy of its one- and two-steps forecasts was extremely strong, however, it failed to pick up the strong upward movement in rents in the second half of the testing period. A closer examination of the forecasts reveal the similarity between each step forecast, all of them measuring between 3.6 and 4.2. This pattern of stable growth can be traced back to the employment forecasts, the major demand side force in the model. The violations in the estimation of the model do however bring into question the replicability and transferability of the model. However, the forecasting results should be viewed in light of the models nature as an equilibrium model. Therefore, the strong rental growth in the late 1990s and the nature of the model may account for the forecasting results. The strong performance of the structured model confirms the existing literature with regard to the ability of this relatively simple model. In part the accuracy of the model in this study may be due to the use of long lags for both office construction and employment, thereby allowing the use of actual figures rather than having to estimate the independent variables as well as the overall model.

7. Concluding comments

This study has compared the forecasting ability of four alternative models designed to forecast office rental in the City of London. The paper compares ARIMA, Bayesian VAR, structural and simultaneous equation models. The forecasting performance of the specified models is then compared in a three-year sample period. Of the individual models tested the BVAR provides the most accurate forecasts, while the ARIMA consistently produces the worst forecast. It is hypothesized that this is due to the use of a simple AR1 model. The simultaneous equation and structural models perform relatively well; however they do not manage to pick up the large upward movement in rental values that occurred from mid-1997 onwards.

The results are naturally unique to both the in-sample period in terms of the specifications adopted and the out-of-sample period. The findings do however highlight some of the relative advantages and disadvantages of the alternative approaches available in forecasting. In particular, the ARIMA model approach does suffer over a long-term forecasting horizon and especially at points in the market cycle where turning points or major movements occur. The OLS model highlights previous findings with regard to its suitability and particularly the work of Chaplin (1998, 1999) in that what may be the best model in sample may not necessarily provide the most accurate forecast. In contrast, while the simultaneous equation suffers from a number of diagnostic problems in its estimation, the strong theoretical grounding of the model results in a relatively accurate forecast. The flexibility present in the VAR approach confirms previous findings as to its suitability, the main limitation in its implementation being the requirement of a relatively long time-series.

Acknowledgements

The authors would like to thank the anonymous referees, participants at the European Real Estate Society conference and seminar participants at California State University-Fullerton for comments on earlier versions of the paper.

References

Ball, M., Lizieri, C. and MacGregor, B.D. (1998) *The Economics of Commercial Property Markets*, Routledge, London.

Brooks, C. and Tsolacos, S. (2000) Forecasting models of retail rents, *Environment and Planning A* **32**, 1825–39.

Chaplin, R. (1998) An ex-post comparative evaluation of office rent prediction models, *Journal of Property Valuation & Investment* **16**, 21–37.

Chaplin, R. (1999) The predictability of real office rents, *Journal of Property Research* **16**, 21–50.

D'Arcy, E., McGough, T. and Tsolacos, S. (1997a) An empirical investigation of retail rents in five European cities, *Journal of Property Valuation & Investment* **15**, 308–22.

D'Arcy, E., McGough, T. and Tsolacos, S. (1997b) National economic trends, market size and city growth effects on European office rents, *Journal of Property Research* **14**, 297–308.

D'Arcy, E., McGough, T. and Tsolacos, S. (1999) An econometric analysis and forecasts of the office rental cycle in the Dublin area, *Journal of Property Research* **16,** 309–22.

Davidson, R. and MacKinnon, J.G. (1993) *Estimation and Inference in Econometrics,* Oxford University Press, Oxford.

Hendershott, P.H., Lizieri, C.M. and Matysiak, G.M. (1999) The workings of the London office market, *Real Estate Economics* **27**, 365–87.

Hendershott, P.H., MacGregor, B.D. and Tse, R.Y.C. (2002) Estimation of the rental adjustment process, *Real Estate Economics* **30**, 165–84.

Kennedy, P. (1998) *A Guide to Econometrics*, fourth edition, Blackwell, Oxford.

McGough, T. and Tsolacos, S. (1994) Forecasting office rental values using vector autoregressive models, *Cutting Edge 1994* 303–20.

McGough, T. and Tsolacos, S. (1995a) Property cycles in the UK: an empirical investigation of the stylised facts, *Journal of Property Finance* **6**, 45–62.

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

Makridakis, S., Wheelwright, S. and Hyndman, R. (1998) *Forecasting Methods and Applications,* John Wiley & Sons, New York.

Mills, T.C. (1990) *Time Series Techniques for Economists*, Cambridge University Press, Cambridge.

Mitchell, P. and MacNamara, P. (1997) Issues in the development and application of property market forecasting: the investors perspective, *Journal of Property Finance* **8**, 363–76.

Pindyck, R.S. and Rubinfeld, D.L. (1998) *Econometric Models and Economic Forecasts*, fourth edition, McGraw-Hill, Boston, MA.

Theil, H. (1971) *Principles of Econometrics,* North Holland, Amsterdam.

Tse, R.Y.C. (1997) An application of the ARIMA model to real estate prices in Hong Kong, *Journal of Property Finance* **8**, 152–63.

Tsolacos, S., Keogh, G. and McGough, T. (1998) Modelling, use, investment and development in the British office market, *Environment and Planning A* **30**, 1409–27.

Weiss, A.A. and Andersen, A.P. (1984) Estimating time series models using the relevant forecast evaluation criterion, *Journal of the Royal Statistical Society. Series A (General)*, **147**, 484–487.

Wheaton, W. (1987) The cyclical behavior of the national office market, *Journal of American Real Estate and Urban Economics Association* **15**, 281–99.

Wheaton, W. and Torto, R. (1988) Vacancy rates and the future of office rents, *Journal of American Real Estate and Urban Economics Association* **16**, 430–6.

Wheaton, W., Torto, R. and Evans, P. (1997) The cyclical behavior of the Greater London office market, *Journal of Real Estate Finance and Economics* **15**, 77–92.