The paper investigates whether complex property market forecasting techniques are better at forecasting than simple specifications. As the research and initial modelling results suggest, simple models outperform the more complex structures. It therefore calls analysts to make forecasts more user-friendly, and for researchers to pay greater attention to the development and improvement of simpler forecasting techniques or simplification of more complex structures. Further planned research presents an alternative simple modelling approach, which was successfully employed by economists and business researchers, helping to achieve greater predictive outcomes.

Keywords: Modelling, Property, Simple, UK.

PROPERTY MARKET MODELLING AND FORECASTING

Property market modelling and forecasting is an indispensable activity in property investment (Mitchell and McNamara, 1997). The issue has been the subject of a number of studies. As a result, numerous models have been developed to forecast property markets (Brooks and Tsolacos, 2010). According to Harris and Cundell (1995, p.76), “the market crash which traumatized the property industry between 1991 and 1994 has led the institutions in particular to seek greater predictive input to their portfolio management and investment decisions”. As McDonald (2002) pointed out, after the 1980s property boom property researchers responded to the crisis situation, and as a result substantial progress has been made in property market research and forecasting. Subsequently, property researchers, including McGough and Tsolacos (1995), Wheaton et.al. (1997), and Barras (2009), suggest that the commercial property market is forecastable. Although, Tonelli et.al. (2004, p.1) argued that “numerous econometric models have been proposed for forecasting property market performance, but limited success has been achieved in finding a reliable and consistent model to predict property market movements”.

SIMPLE AND COMPLEX MODELS

According to Caminiti (2004, p.992), models are “invaluable tools”. Models help their users to develop a better understanding of complex systems, allow testing for possible scenarios, predicting outcomes, as well as they can assist in the setting of priorities. Byrne et.al. (2010) added that models have been produced for a range of different reasons, i.e. to improve one’s understanding on the subject and its processes, to predict, forecast or explore possible scenarios, or to provide a basis for decision-making. Despite all the benefits of various models, many concerns have been expressed regarding their application. According to STOWA/RIZA (1999), extensive use of

1Corresponding author: a.jadevicius@napier.ac.uk
models increases the risk of inexpert use which as a result can lead to unreliable modelling outcomes. Similarly, Middlemis et.al. (2000) observed that if the model is poorly designed, or it does not represent the system being modelled properly, all the efforts to create the model are virtually in vain, or it will then generate inaccurate forecasts. Subsequently, the problems associated with models were articulated by Jakeman et.al. (2006, p.603), who identified that the use of models can bring unwanted outcomes due to "limitations, uncertainties, omissions and subjective choices in models".

The issue was also addressed by various researchers, including Box and Draper (1987, p.424), who earlier presented a more rigorous critique, that “essentially all models are wrong, but some are useful”, Mellor et.al. (2003, p.16) who indicated that “models offer more hindrance than help”, and Sterman (2002, p.525) who stated that ”all decisions are based on models, and all models are wrong”. As Sterman observed, models are wrong because they are only a simplification, an abstraction of the system with no solid foundation, and, what is more, they are based on human perception and knowledge which are fundamentally limited. However, despite all of this criticism, researchers including Parker et.al. (2002) and Caminiti (2004) consider models to be assets rather than liabilities, and essential elements in understanding and forecasting complex systems.

What is the difference between simple and complex models?

Looking at the various fields of science, there have been a very few direct comparisons between simple and complex models. What is more, there are only a few descriptions of what constitutes simple and complex models and what differences between the two structures are (Buede, 2009). Researchers, including Armstrong et.al. (1984), Armstrong (1986), Wilkinson (1999), and Sterman (2002) referred to simple and complex models in their publications. However, none of these researchers defined precisely what they meant by each structure.

In their systemic study on modelling complexity, Batty and Torrens (2001) proposed a working definition of a complex system. According to the authors, a complex system is “an entity which is coherent in some recognisable way but whose elements, interactions, and dynamics generate structures admitting surprise and novelty which cannot be defined a priori” (ibid., p.2). According to Holland (1995), a “complex” or “adaptive” model is one that maintains its composition and exhibits coherence through time. As Allen and Strathern (2005) suggested, in the case of the socio-economic environment, a simple model is a structure of fixed, predictable behaviour, while a complex one is a system in which a range of possible structural changes can appear. According to Buede (2009, p.11), complex, or as the author indicated, “more realistic [science-based] models”, usually require greater amounts of data, as well as a more complex set of relationships.

In property, as in other fields of research, there have been only minor references as to what constitute simple and complex models. According to Chaplin (1999), econometric models, which are considered as complex structures, differ from simple competitors as they include more variables and contain a greater number of estimations. Brooks and Tsolacos (2000) suggested that simple autoregressive structures (Long-Term-Mean and Random-Walk models) constitute only a part of more complex Vector Autoregressive (VAR) system. According to Stevenson and McGarth (2003), more complex forecasting techniques, such as Bayesian Vector Autoregression approach “improve economic modelling and forecasting” over the simple forecasting methods by offering a more flexible forecasting process (ibid., 246). What is more, Brooks and Tsolacos
(2010) commented that time-series extrapolative models such as Exponential Smoothing or ARIMA are atheoretical, which means that they are not based on any underlying economic theory. These models produce forecasts capturing only empirically relevant properties of selected time-series. In contrast, econometric or more complex structures are based on the economic theory relevant to the subject. These models also contain a system of equations representing the economic theory derived for the subject being modelled.

As the discussion above suggests, the more complex analytical and econometric modelling structures are “newer”, include more variables, contain a greater number of estimations, and accounts for attributes of the external environment. In contrast, simple models constitute an uncomplicated combination of rules, a limited number of variables, in most cases they are of a fixed structure and usually extrapolate into the future from the past values of the time-series itself. As such, following classification of the real estate forecasting methods presented by Lizieri (2009), it can be suggested that Exponential Smoothing, Simple Regression, Multiple Regression and ARIMA models are all simple forecasting techniques. The Econometric and VAR specifications constitute more complex modelling structures.

**Which models are best?**

The question that needs to be asked however is whether complex forecasting structures are any better at forecasting than simple ones. This issue has been raised within various scientific areas including environment, economics, and physiology.

Outside the real estate discipline, the findings of Dorn (1950) and Hajnal (1955), who investigated demographics forecasting, suggested that complex population forecasting models, which typically incorporate large amounts of inputs, become overly complicated, and thus exhibit poorer accuracy.

In their analytical survey, Armstrong et.al. (1984) assessed the relative accuracy of both complex and simple extrapolative methods. The commentators identified that simple methods (e.g. exponential smoothing) exhibit a comparable degree of accuracy to more complex ones (e.g. Box-Jenkins approach). In the subsequent paper, Armstrong (1986) made a qualitative review of the forecasting methods of the period from 1960 to 1984. The author arrived at the same conclusion that forecasters should be in favour of simple forecasting techniques over more complex econometric structures.

In the discussion on the issue within econometrics, Kennedy (2002, p.575) subsequently called for a “keep it sensibly simple” modelling approach. Kennedy adopted this terminology from Zellner (1991, p.6), who recommended “KISS” or “Keep It Sophistically Simple” modelling style. As both authors indicated, simple specifications are just as good as complex empirical techniques. Kennedy also substantiated his approach by referencing to Wilkinson et.al. (1999, p.601) who recommended “Choosing a Minimally Sufficient Analysis”. According to Wilkinson, the researcher should not “choose an analytic method to impress [his] readers or to deflect criticism”. The commentator suggested that ”if the assumptions and strength of a simpler method are reasonable for your data and research problem, use it” (ibid., p.601).

In their research Beven and Freer (2000) found that simple and complex models, which are used to reproduce the observed behaviour of the system, may exhibit equifinality. As Byrne et.al. (2010, p.7) explained, the term equifinality implies the situation where “different parameter sets may yield equivalent model outputs”, i.e. similar outcomes can be achieved from different models and in different ways.
Clements and Hendry’s (2003) investigation into economic forecasting was also not in a favour of complex forecasting models. As the authors indicated, “although which model does best in a forecasting competition depends on how the forecasts are evaluated and what horizons and samples are selected, “simple” extrapolative methods tend to outperform econometric systems” (ibid., p.304). More recent evidence from Buede (2009) suggests that although simple models contain a large variance in their predictions, complex models still have a large probability of producing wrong results. Orrell and McSharry (2009) also observed that, as models become more complex and parameterised, the number of elements they contain increases significantly. As a result, even small changes in these parameters can have significant consequences to modelling outcomes. Certainly, more parameterised models may fit historic data better, their structure can also be more flexible. However, as the authors observed, such models are less helpful at predicting the future. Accordingly, Orrell and McSharry (ibid.) referred to Occam’s razor principle (cited in Standish, 2004, pp.256), which states that “entities should not be multiplied unnecessarily”. In other words, models should be as simple as possible with the minimum number of parameters. Subsequently, Orrell and McSharry suggested that instead of developing a “model of everything”, one should aim in developing a set of models which could be adapted to a particular situation (ibid., p.741).

A more recent argument for simplicity is so called "frugal innovation" approach. Pioneered by Prahalad (2006) and now becoming popular within the business and management community, this approach suggests that one should keep things simple and look for what people actually need (Immelt et.al., 2009; Sehgal et.al., 2010; Radjou et.al., 2012). The main principle of this approach is to produce ideas that are affordable and flexible. Radjou et.al. (2012 p.4) call it “Jugaad Innovation”. This Hindi term translates as "an innovative fix; an improvised solution born from ingenuity and cleverness". According to the authors, Jugaad is about doing more with less.

The opposite findings, however, were presented by Armstrong (1975), Pandy (2003) and Li et.al. (2005). In his research, Armstrong (ibid.) identified that naïve (simple) forecasting structures are less accurate than causal (more complex) methods. Pandy’s (ibid.) investigation into muscle function in walking and comparison of two different models revealed that although a simple model can help to explain basic characteristics of movements, it tend to exhibit misleading results. According to Li et.al. (2005), the Basic Structural Model (BSM) specification performed better than less complex forecasting technique in the context of the international tourism demand modelling.

In the property forecasting literature, as the evidence suggests, simple models such as Exponential Smoothing, Simple Regression, or ARIMA specifications outperform the more complex forecasting techniques, including VAR and Econometric models, or at least generate highly comparable outcomes (Chaplin, 1999; Newell et.al., 2002; Stevenson and McGarsh, 2003). This therefore led Newell et.al. (2002) to suggest that despite the increased complexity in property market modelling methodologies, simple methods are often found to be as good as complex econometric structures.

It all therefore suggests that, regardless of the comments that complex modelling structures contain more variables and equations, consider elements of the external environment, and seem to fit historic data with greater accuracy, simple forecasting methods are in many instances more accurate, or at least as accurate as complex structures. According to Makridakis (1988, p.475), “forecasting errors can result from many sources and cannot be eliminated by more complex models or more gifted forecasters”. Pant and Starbuck (1990, p.442) also suggested that “more complex, subtle, or elegant techniques give no greater accuracy than simple, crude or naïve ones.
More complex methods might promise to extract more information from data, but such methods also tend to mistake noise for information. As a result, more complex methods make more serious errors, and they rarely yield the gains they promised. Therefore, if to follow Mahmoud (1984), the implications of these findings are to make forecasts more user-friendly, and for researchers to pay greater attention to the development and improvement of simpler forecasting techniques or simplification of more complex structures. This proposition can be well generalised by Thoreau (1897, pp.144) who more than one hundred years ago wrote: “simplicity, simplicity, simplicity! I say let your affairs be as two or three and not a hundred or thousand <...> simplify, simplify”.

**Initial Modelling Results**

The property market modelling was performed using Exponential Smoothing (Simple Exponential Smoothing, Holt’s Linear Trend and Brown’s Linear Trend), Simple Regression, Multiple Regression, Vector Autoregression, ARIMA and ARIMAX techniques. The IPD and Scott’s (1996) combined All Property Rental Value Growth Index for the UK was the dependent variable. Bank Rate, Construction Costs, Construction Orders, Construction Output, Construction Starts, Employment, and GDP were all explanatory variables selected for the research. All time-series were for 1963-2010 period in annual numbers. Time-series were tested for stationarity following Koop (2006). All models were parameterised and tested on the initialisation period from 1964 to 2000. Forecasts were made on the holdout set from 2001 to 2010. The in-sample accuracy was assessed by computing R-square, Mean Error (ME), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) accuracy measures. The out-of-sample accuracy was examined from Theil’s second inequality coefficient “U”. In addition to that, Akaike Information Criterion (AIC) was used as a measure of the best parameterised specification. What is more, two supplementary tests were computed. One was Durbin-Watson (DW) test for autocorrelation, which assessed whether autocorrelated disturbances are present within the model. The other test was White’s test (WT) which assessed the presence of heteroscedasticity (Chaplin, 1999; Makridakis et.al., 1998; Stevenson and McGarth, 2003; Brooks and Tsolacos, 2010). Following on from this, 140 model specifications were computed, including 3 Exponential Smoothing, 7 Simple Regression, 1 Multiple Regression, 1 VAR, 16 ARIMA, ranging from ARIMA (1,0,1) to ARIMA (4,0,4), and 112 ARIMAX specifications, ranging from ARIMA (1,0,1) to ARIMA (4,0,4) with seven explanatory variables.

The statistical results indicate the VAR (1) specification to be the best fitting model. Its R-squared is the greatest of all sample models. The AIC also indicates it to be the best parameterised specification of all the candidate models (Table 1). However, these results do not come as a surprise. The VAR model comprises lagged values of all explanatory variables, as well as past values of the dependent variable itself. It all therefore explains its goodness of fit to the historic data.

However, when it comes to the out-of-sample forecasting performance, VAR’s accuracy is not so impressive. It’s Theil’s U value is poorer than that of some less complex ARIMAX and Simple Regression models (Table 1). All that adds further to the suggestion that goodness of fit does not imply good forecasting performance, and that increased model complexity does not necessarily yield greater forecasting accuracy (Chaplin, 1999; Newell et.al., 2002; Stevenson and McGarth, 2003). Subsequently, it connotes to the above noted suggestion that forecasts should be made more user-friendly, and that researchers should pay greater attention to the development and improvement of simpler forecasting techniques or simplification of more complex structures.
Table 1: Summary model fit statistics

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>R-squared</th>
<th>MAE</th>
<th>MAPE</th>
<th>AIC</th>
<th>Theil’s U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Exponential Smoothing</td>
<td>-0.03</td>
<td>4.40</td>
<td>109.26</td>
<td>130.87</td>
<td>0.94</td>
</tr>
<tr>
<td>Holt’s Linear Trend</td>
<td>-0.03</td>
<td>4.39</td>
<td>110.32</td>
<td>131.15</td>
<td>0.93</td>
</tr>
<tr>
<td>Brown’s Linear Trend</td>
<td>-0.00</td>
<td>4.36</td>
<td>100.24</td>
<td>131.23</td>
<td>1.00</td>
</tr>
<tr>
<td>Simple Regression (Bank Rate)</td>
<td>0.00</td>
<td>4.33</td>
<td>98.26</td>
<td>130.47</td>
<td>0.95</td>
</tr>
<tr>
<td>Simple Regression (Construction Costs)</td>
<td>0.00</td>
<td>4.36</td>
<td>98.86</td>
<td>130.49</td>
<td>0.97</td>
</tr>
<tr>
<td>Simple Regression (Construction Orders)</td>
<td>0.34</td>
<td>3.74</td>
<td>142.38</td>
<td>115.26</td>
<td>0.41</td>
</tr>
<tr>
<td>Simple Regression (Construction Output)</td>
<td>0.02</td>
<td>4.53</td>
<td>108.08</td>
<td>129.94</td>
<td>0.88</td>
</tr>
<tr>
<td>Simple Regression (Construction Starts)</td>
<td>0.00</td>
<td>4.35</td>
<td>97.93</td>
<td>130.46</td>
<td>0.93</td>
</tr>
<tr>
<td>Simple Regression (Employment)</td>
<td>0.03</td>
<td>4.09</td>
<td>87.64</td>
<td>129.36</td>
<td>0.82</td>
</tr>
<tr>
<td>Simple Regression (GDP)</td>
<td>0.32</td>
<td>3.63</td>
<td>120.19</td>
<td>118.50</td>
<td>0.47</td>
</tr>
<tr>
<td>Multiple Regression</td>
<td>0.55</td>
<td>3.06</td>
<td>141.65</td>
<td>109.35</td>
<td>0.46</td>
</tr>
<tr>
<td>Vector Autoregression</td>
<td>0.79</td>
<td>2.47</td>
<td>91.85</td>
<td>85.180</td>
<td>0.48</td>
</tr>
<tr>
<td>ARIMA (1,0,2)</td>
<td>0.52</td>
<td>2.60</td>
<td>66.04</td>
<td>109.85</td>
<td>0.85</td>
</tr>
<tr>
<td>ARIMAX (1,0,2) (Bank Rate)</td>
<td>0.52</td>
<td>2.58</td>
<td>66.40</td>
<td>112.32</td>
<td>0.82</td>
</tr>
<tr>
<td>ARIMAX (1,0,2) (Construction Costs)</td>
<td>0.52</td>
<td>2.61</td>
<td>66.69</td>
<td>112.46</td>
<td>0.74</td>
</tr>
<tr>
<td>ARIMAX (1,0,2) (Construction Orders)</td>
<td>0.60</td>
<td>2.61</td>
<td>80.83</td>
<td>103.24</td>
<td>0.33</td>
</tr>
<tr>
<td>ARIMAX (1,0,2) (Construction Output)</td>
<td>0.52</td>
<td>2.60</td>
<td>65.49</td>
<td>112.66</td>
<td>0.84</td>
</tr>
<tr>
<td>ARIMAX (1,0,2) (Construction Starts)</td>
<td>0.52</td>
<td>2.60</td>
<td>67.45</td>
<td>112.27</td>
<td>0.83</td>
</tr>
<tr>
<td>ARIMAX (1,0,2) (Employment)</td>
<td>0.52</td>
<td>2.60</td>
<td>66.65</td>
<td>112.90</td>
<td>0.84</td>
</tr>
<tr>
<td>ARIMAX (4,0,0) (GDP)</td>
<td>0.69</td>
<td>2.26</td>
<td>69.83</td>
<td>99.09</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Figure 1. VAR (1) (model fit and forecasting accuracy) (1st.dif.)
FURTHER PLANNED RESEARCH

The further planned research presents the principle of combination forecasting as an alternative simple modelling approach helping to achieve greater predictive outcomes. This modelling principle was successfully employed by economists and business researchers, including Makridakis (1989), De Gooijer and Hyndman (2006), Goodwin (2009), Pesaran and Pick (2011) and Wallis (2011). The researchers were motivated that by combining forecasts from different methods and sources greater predictive results can be achieved. What is more, their theoretical and empirical findings suggested usefulness of this procedure. Although, Bates and Granger (1969) and more recently Kapetanios et.al. (2008) observed that the combination forecasting does not necessarily lead to a better forecasting performance, the empirical results presented in Pagourtzi et.al. (2005) and Gupta et.al. (2011) indicated a benefit of this procedure and suggest further research in this area.

CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

The current study questioned whether complex forecasting techniques are better at forecasting than simple ones. The issue has been raised within various scientific areas. As the research and the initial modelling results revealed, simple and complex models exhibit equifinality. The initial modelling results indicated that a simple ARIMA model with Construction Orders as an explanatory variable (ARIMAXCoR) was more accurate in producing the out-of-sample forecasts than the more complex VAR model. It therefore calls for the adoption of a “keep it sensibly simple” modelling approach. The recommendation for analysts is to make forecasts more user-friendly, and for researchers to pay greater attention to the development and improvement of simpler forecasting techniques or simplification of more complex structures.

Further planned research assesses the main reasons behind forecasting inaccuracy. It then presents an alternative simple modelling approach, i.e. combination forecasting which was successfully employed by economists and business researchers, helping to achieve greater predictive outcomes.
REFERENCES


Lizieri, C.M. (2009) "Forecasting and Modelling Real Estate". School of Real Estate & Planning, Henley Business School, University of Reading [Internet]. Available at: <http://www.henley.reading.ac.uk/web/FILES/REP/Forecasting_Version_2.pdf>, Accessed [02 November, 2011].


Thoreau, H.D. (1897) "Walden (Volume 1)". Houghton, Mifflin and Company, Boston, pp.263.


