



Delft University of Technology

The accuracy of consensus real estate forecasts revisited

McAllister, Pat; Nase, Ilir

DOI

[10.1080/09599916.2020.1720784](https://doi.org/10.1080/09599916.2020.1720784)

Publication date

2020

Document Version

Final published version

Published in

Journal of Property Research

Citation (APA)

McAllister, P., & Nase, I. (2020). The accuracy of consensus real estate forecasts revisited. *Journal of Property Research*, 37(2), 147-170. <https://doi.org/10.1080/09599916.2020.1720784>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

The accuracy of consensus real estate forecasts revisited

Pat McAllister & Ilir Nase

To cite this article: Pat McAllister & Ilir Nase (2020) The accuracy of consensus real estate forecasts revisited, Journal of Property Research, 37:2, 147-170, DOI: [10.1080/09599916.2020.1720784](https://doi.org/10.1080/09599916.2020.1720784)

To link to this article: <https://doi.org/10.1080/09599916.2020.1720784>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 05 Feb 2020.



[Submit your article to this journal](#)



Article views: 378



[View related articles](#)



[View Crossmark data](#)



The accuracy of consensus real estate forecasts revisited

Pat McAllister^a and Ilir Nase^b

^aReal Estate and Planning, University of Reading, Reading, UK; ^bDepartment of Management in the Built Environment, Faculty of Architecture and the Built Environment, TU Delft, Delft, Netherlands

ABSTRACT

This study updates and expands upon the existing work on the accuracy of the IPF's Consensus Forecasts. The paper evaluates the extent to which the consensus forecasts were able to predict the relative performance. It also assesses the accuracy of implied yield forecasts and concludes that failure in yield forecasting is the main source of failure in forecasts of capital growth and total returns. A high level of agreement between the actual and forecasted sector rankings was found. Evidence of a pessimism bias was identified. Yield forecasts are consistently found to perform worst using a range of forecast performance metrics.

ARTICLE HISTORY

Received 8 August 2019
Accepted 20 January 2020

KEYWORDS

Forecasting accuracy;
consensus forecasts;
commercial real estate

I Introduction

This paper investigates the accuracy of the Investment Property Forum (IPF)¹ Consensus Forecasts in the UK commercial real estate market. Given the global significance of real estate as an asset class and a common perception that the market is characterised by inefficiency, the body of knowledge on the performance of real estate forecasts is surprisingly thin. For institutional real estate investors, forecasts of future rents, vacancy rates, yields² etc. and returns for assets across classes are key inputs into decisions about stock selection and asset allocation. Consequently, the ability of real estate forecasters to provide investment organisations with forecasts that add value should be a concern to both users and producers of real estate forecasts. There is a substantial real estate forecasting 'industry' with most institutional investors, their advisers and specialists producing regular forecasts. Beginning in 1999, the IPF Consensus Forecasts have provided a rich source of data on market expectations of real estate forecasters in the UK.

Building on previous research that has evaluated the accuracy of the IPF Consensus Forecasts, this study contributes to the body of knowledge on the topic in three ways. Firstly, it updates and expands existing research on the performance of the IPF's Consensus Forecasts of rental growth, capital growth and total return applying a range of error metrics and other accuracy measures commonly used in forecast evaluation. Previous studies focussed only at the All Property level. This study now includes sector level forecasts and expands the sample period. Secondly, the availability of sector level forecasts presented an opportunity to assess the extent to which the consensus forecasts

CONTACT Ilir Nase  I.Nase@tudelft.nl

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

were able to predict the *relative* performance. Finally, since it has been argued that failures in forecasting real estate returns are caused largely by limitations in yield forecasting, a further contribution of the research is that it evaluates the performance of yield forecasts implied in the relationship between the explicit forecasts of rental and capital growth.

The remainder of the paper is organised as follows. In the next section, focussing mainly on the most closely related previous research, a brief discussion of the literature on the topic is provided. This is followed by a description and short evaluation of the metrics used to measure the accuracy of forecasts. The third section discusses the data. The fourth section describes and discusses the results of the forecasting performance tests. Finally, conclusions and implications for practice are drawn.

II Previous research

The scale of the empirical research on forecast accuracy in numerous fields precludes a comprehensive review of the literature in this context. However, it is important to acknowledge that it is widely accepted that uncertainty and disagreement are intrinsic to economic forecasting. One source of forecast uncertainty can lie in the forecast modelling process. The innate modelling problem in forecasting is largely summed up by George Box's widely cited dictum that '[T]he most that can be expected from any model is that it can supply a useful approximation to reality: All models are wrong; some models are useful' (Box, Hunter, & Hunter, 1978, 440). In the context of economic forecasts, Hendry and Clements (2003, 303) echo George Box when they argue that 'all econometric models are mis-specified, and all economies have been subject to unanticipated shifts'.

In economic forecasting, Hendry and Clements (2003) argue that it is seldom forecasting models that are the main source of forecast failure. Although failure may be partially caused by weak theory and incorrect observations, they argue that it mainly arises from structural instability in the underlying causal relationships. Unanticipated shocks, in particular, can lead to systematic forecast failure. Structural breaks – sudden large changes, invariably unanticipated – have been a major source of forecast failure. It is also worth pointing out the inherent smoothing that is a feature of econometric models. Hendry (2003) highlights the intrinsic averaging that occurs in econometric modelling which averages over likely future shocks and produces signals of the 'average future'.

The prevalence of unanticipated shocks can mean that, model specification can be irrelevant to forecast performance with well-specified models outperformed by poorly specified models in forecasting competitions. In line with findings outside the real estate sector, real estate researchers have found that in many instances, simple forecasts (e.g. via naïve predictors) to be more accurate than using complex econometric models (Chaplin, 1998, 2000; Wilson et al, 2000). Further, in macro-economic forecasts, non-causal models often tend to dominate causal models (Clements & Hendry, 1999).

The optimal response to the challenges of incorporating structural breaks in forecasting models has been judgemental overlay so that economic and financial forecasts are often the product of a blend of quantitative modelling and qualitative judgements. Focussing on the UK commercial real estate forecasting sector, Watkins, White, and Keskin (2012) state that most forecasters cite mood and sentiment as the main reasons for making such overlay adjustments. There has been growing research on the

information content of sentiment in the forecasting process (see Marcato and Nanda, 2016) Whilst poor data, model mis-specification etc. can be categorised as honest mistakes, there are a number of long-recognised behavioural biases and agency problems in forecasting.

Batchelor (2007) provided an overview of how different informational environments and incentive structures can bias individual financial and economic forecasts. Gallimore and McAllister (2005) found that real estate forecasters in the UK often self-censored or were censored when forecasting models generate contentious or conspicuous forecasts. In the US commercial real estate market, Ling (2005) found evidence of adaptive rather than rational expectations in RERC's survey of expected sector performance. It was found that consensus predictions were correlated with market returns in the two years *prior* to the survey. No correlation between the prediction of the RERC survey respondents and actual return performance was identified. It was concluded that investment strategies based on the specific measure of consensus expectation would not have improved total return performance – the criterion of prime importance to most institutional investors.

There have been a number of previous studies that have evaluated the accuracy of the IPF Consensus Forecasts of rental growth, capital growth and total returns. Two studies focused on individual forecaster disagreement rather than the performance of the consensus forecasts (see McAllister, Newell, & Matysiak, 2006; Papastamos, Stevenson and Matysiak 2015). Focussing on a comparison of the performance of various econometric models with the IPF Consensus Forecasts of rental growth and total return for the UK market for period 1999–2004, Tsolacos (2006) found no consistent performance differences in forecasting ability between the consensus and the output of relatively simple regression models. Again, over a relatively short sample period (1999–2004) McAllister, Newell, and Matysiak (2008) found that the consensus forecasts were more successful in terms of forecasting rental compared to capital returns. They concluded that the fact that capital growth consensus forecasts tended to track rental growth forecasts probably reflected the challenges of forecasting yields.

Drawing upon a longer sample period (1999–2011) using similar data and methods to McAllister et al. (2008), Papastamos et al. (2015) also concluded that, compared to capital growth and total return forecasts, rental growth forecasts performed much better. However, both studies seemed to have used a flawed approach to comparing the forecasting performance of the consensus compared to naïve forecasts. Their focus was on the performance of individual forecasters compared to naïve forecasts for individual years rather than comparing the forecasts of the consensus and naïve forecasts for the relevant sample periods. This suggests that the performance of the consensus relative to simple naïve forecasts has not been effectively evaluated in the existing research. Nevertheless, both also identified large forecasting errors in capital returns which they also attribute to the difficulties of forecasting yields.

Bond and Mitchell (2011) compared the accuracy of implied capital returns from real estate derivative prices to the IPF Consensus Forecasts for the period 2006 to 2009. Acknowledging the short sampling period, consequent small number of observations and looking only at the mean absolute error as a measure of forecast uncertainty, they found evidence that the (implied) forecasts in the swaps market for one year ahead were marginally more accurate than consensus forecasts with the latter performing better over two and three year horizons. It is also worth noting that there has been

a longstanding body of work comparing the accuracy of different real estate forecasting models assuming either perfect foresight for independent variables or common sources of forecasts (see Tsolacos, 2006; Stevenson & McGarth, 2003; Chaplin, 1998).

Whilst there is a longstanding body of work modelling the determinants of real estate market yields, the literature on yield forecasting is much thinner. At the individual asset level, a forecast of the exit yield is a significant variable in estimating future returns. Projections of total returns at the individual property, portfolio or index level require estimates of yield changes. Watkins et al. (2012) identified a range of potential yield forecasting techniques. These included autoregressive integrated moving average models (ARIMA), a vector error correction mechanism and vector autoregressive models relating yields to lagged yields, rental growth, bond yields and stock market variables. More recent studies also apply an error correction framework.

Drawing upon semi-structured interviews with forecasters, Gallimore and McAllister (2003) found that they generally regarded yield forecasting as inherently more difficult than rent forecasting. It was found that real estate market forecasters had little confidence in their ability to provide reliable yield forecasts using econometric techniques and, given this lack of confidence, that judgement seemed to be a much more important element of the yield forecasting process. There seems to be only one study that has evaluated yield forecasts. For the period 2001–2011 in Australia, Perera, Higgins, and Wong (2018) included ‘six months out’ real estate yield forecasts as well as for a range of economic and other real estate market variables in their forecast accuracy study. Whilst there is limited detail and discussion of the findings on yield forecasts in their paper, their results suggest extremely poor performance of yield forecasts compared to forecasts of other economic and real estate market variables.

III Measuring forecast performance

The terminology regarding the measurement of forecast uncertainty has echoed much of the debate on appraisal uncertainty and smoothing. Similar distinctions are drawn between random variations in the actual and predicted outcomes (error) and systematic tendencies towards optimism or pessimism (bias). ‘Forecast smoothing’ is a well-established stylised fact. In the 1980s, Nordhaus (1987) speculated that the commonly observed lower volatility in forecasts, relative to actual outcomes, is due to behavioural and incentives structures factors such as the need to reach a consensus and to maintain forecast credibility by avoiding major ‘jumps’. The previous research by McAllister et al. (2008) and Papastamos et al. (2015) found that forecasted time series of rental and capital growth tended to be much less volatile than the actual time series.

A common problem in comparing the performance of forecasts for different variables or different markets is differences in scale and other attributes of the variables forecasted. For instance, for MSCI IPD, the average annual growth in market rental levels for the retail sector between 1981 and 2017 has been 4.1% per annum with a standard deviation of 5.2%. Given that a high auto-correlation coefficient implies predictability, this coefficient is very high at 0.83. Capital growth on average has been 3.8% per annum with a standard deviation of 8.5%. The auto-correlation coefficient is much lower at 0.42 which is consistent with a more volatile and less smooth time series. Despite their similar average rates of growth, the different levels of variation in the

series and different levels of 'predictability' (as proxied by the level of auto-correlation) suggest that comparing 'raw' forecasts of capital growth with forecasts of rental growth may not be an appropriate like for like comparison. In turn, different real estate sectors or sub-markets can have different long-term rates of growth, variability and auto-correlation. The yield time series tend to be at the other end of the scale relative to capital growth. The average equivalent yield for the retail sector between 1981 and 2017 was 7.1% with a standard deviation of 1.1% and an auto-correlation coefficient of 0.86. Hence, it is worth bearing in mind the different attributes of the different variables to be forecasted when evaluating forecast performance and evaluating *measures of* forecast performance.

Typically, forecast accuracy measures are classified into two types. The first category is independent measures, i.e. forecast performance measures are solely a function of the forecast under evaluation. The second category is relative measures that compare forecasts to a benchmark forecast, e.g. a forecast of no change. The most commonly applied independent error metric measures are Mean Error, Mean Absolute Error, square root of the Mean Squared Error. Theil's U2 statistic is a common measure of relative performance which can be decomposed into three parts namely bias, variance and covariance proportions. These measures have been covered in detail in previous papers. Further information on their calculation is provided in [Appendix 1](#).

Since the contribution of forecasts to organisational decision-making will depend on the specific application of forecasts in the organisation, forecast evaluation is to some extent an inherently subjective process. In terms of forecast evaluation, Granger & Pesaran, (2000, 538) have advocated a decision theoretical approach where there is a 'consideration of the linkage between the modeler who produces forecasts and the decision-maker who consumes them'. The most appropriate measure of accuracy is therefore dependent upon the utility of the forecast to the forecast user. Whilst it is possible to focus on loss functions, these also are typically user specific. In practice, Gallimore and McAllister (2003) found that UK real estate forecasters generally perceived forecast success and failure in relative rather than absolute terms. Forecasters placed importance on being correct about the performance rankings rather than getting the absolute numbers right. There is a need for some caution in generalising about whether all or most real estate forecasters now regard relative rather than absolute performance as more important. The Gallimore and McAllister (2003) study was qualitative and essentially exploratory. The composition of the real estate investment sector has changed significantly over the last two decades. Some investors or funds can have specialised sector and/or geographical foci or, alternatively, may be diversified internationally. As discussed further below, there are incentives for forecasting organisations to strategically bias outputs. Nevertheless, the data do provide an opportunity to evaluate the relative performance of the consensus forecasts.

In order to evaluate the extent to which forecasters agree in terms of ranking the various sectors, we use Kendall's coefficient of concordance (also known as Kendall's W). This is commonly used for assessing agreement among raters. Kendall's W ranges from 0 (no agreement) in ranking to 1 (complete agreement). If W is 0, then rankings may be regarded as being random with intermediate values suggesting a higher of lower level of agreement among the various sets of rankings. In this specific case for rankings of the sectors (Office, Retail, Shopping Centres etc.), Kendall's W is used to

measure the extent to which the predicted sector rankings of the consensus ‘agree with’ the actual observed rankings. Kendall’s W is calculated as

$$W = \frac{12S}{m^2(n^3 - n)} \quad (1)$$

Where m is the number of rankers and n is the number of categories being ranked. In this study we investigate two cases where for every year there are two rankers; the actual outcomes and either the consensus forecasts or the ARMA forecasts ranking eight sectors (categories) for four variables. The relative performance of consensus with regard to ARMA forecasts is assessed by comparing each year’s Kendall’s W values across the four variables. S represents the sum of the squared differences between each rank R_i where $i = 1:n$ and the mean rank for each category.

IV Data

The core data on the consensus forecasts has been provided by the IPF. Whilst there are numerous potential permutations of forecasts that could be evaluated, we focus here only on the one year ahead consensus forecasts from 2006 until 2017. For nearly two decades, the IPF have been surveying forecasters employed by a broad range of institutional investors and their advisors classified into property advisors (including research consultancies) and fund/investment managers.³ Surveyed four times per annum (in February, May, August and November), respondents are asked to provide forecasts of the annual rate of rental growth, capital value growth and total returns until the end of current calendar year, until the end of the subsequent calendar year and until the end of the fifth calendar year. For instance, in the August 2018 survey, respondents are asked to estimate the average rate of rental growth per annum for the period from January 2018 until December 2018, December 2019 and December 2022. We focus only on the February forecasts which approximates to a one year horizon producing one annual observation for each sector for each of the 13 years. In addition to UK All Property, forecasts are requested for seven sectors – Standard Retail, Shopping Centres, Retail Warehouses, Offices, City Offices, West End Offices and Industrial. It is also worth noting that there is overlap between some of the sectors. Most obviously, the All Property indices are composed of assets in the other seven sectors. In addition, the assets in the City and West End sectors comprise a substantial proportion of the Office sector.

Due to mergers and acquisitions and new entrants, re-entrants etc. into the real estate advisory services and fund management sectors, the number of contributing firms has changed from forecast period to forecast period. As financialisation and globalisation of commercial real estate markets has deepened over the last two decades, a broader user base of real estate forecasts has emerged. Baym and Hartzell (2012, 9) referred to the creation of ‘a more complex industry structure and a confusion of ownership and management’. This increased complexity creates challenges in generalising about what constitutes ‘practice’ in terms of the production and use of forecasts. Recent empirical evidence on investment and forecasting practice for commercial real estate markets is thin. However, there can be little doubt that investing institutions now have access to a range of paid for and ‘free’ forecasts for the various MSCI market sectors. Many large fund management organisations will internally produce their own ‘house’ forecasts and/or subscribe to external forecast

providers. It can be more nuanced. For instance, a fund management organisation may subscribe to an external forecasting service but ask the real estate forecasting organisation to customise forecasts to their specific expectations about, for example, interest rate movements or GDP growth or a subscriber may simply subjectively adjust externally provided forecasts. Whilst the main applications of forecasts still seem to be asset allocation and stock selection, there is anecdotal evidence that the nascent real estate derivatives market has increased market sensitivity about the production and release of the consensus forecasts. In context of this research, the IPF Consensus Forecasts at sector level are likely to be applied for strategic and tactical asset allocation. More localised forecasts or significant adjustment to wider sector forecasts are usually needed for specific submarkets and assets.

In addition, there are other *producers* of commercial real estate forecasts external to fund managers. Watkins et al. (2012) stated that commercial real estate forecasts are produced by combining econometric predictions with a more subjective market overlay process. There may be different incentives to herd and/or to strategically bias their forecasts between the different forecasters and organisations. Real estate advisors and research consultancies produce forecasts that are sold externally. Advisors may also publish some of their forecasts essentially as part of marketing their services. There is anecdotal evidence that institutional investors believe that the forecasts of the global real estate advisory firms are optimistically biased for fear that 'negative' forecasts could jeopardise transactions.⁴ Watkins et al. (2012) noted that organisations that publish forecasts for external use tend to employ models where the assumptions and methods are explicit since, even if it is acknowledged that the forecast is expected to be incorrect, the outputs can be more readily justified.

To assess forecast performance, the actual outcome is the growth rate or return for the relevant annual index produced by MSCI (formerly produced by Investment Property Databank (IPD)). In terms of evaluating forecasting performance, the relevant MSCI market or sector index is taken as the actual outcome. The performance data is available for the period 1981–2017. This paper investigates the forecasting performance (forecast error measures are outlined in [Appendix 1](#)) as the difference between the consensus forecasts and the actual outcomes. Data on the IPF's Consensus Forecasts are taken from 2006 because the sector composition of the forecast survey has been stable since then providing a consistent time series.

In addition, we use the actual series to estimate one year ahead forecasts with simple trend models. In previous research on the IPF Consensus Forecasts, the benchmark naïve forecast against which the performance of forecasts was compared was typically a 'no change' or 'same as last period' forecast. Given that this is a rather undemanding hurdle to jump over, some basic trend-based forecasts have been made as a forecasting performance benchmark against which to compare the IPF Consensus Forecasts. We undertook preliminary analysis of the MSCI data series for three different periods namely 1981–2005; 1981–2010 and 1981–2017. These comprised of Augmented Dicky-Fuller (ADF) tests and Autocorrelation and Partial Autocorrelation Function (ACF – PACF) graph analysis. They indicated that, for the majority of the series, the predominant structure was a second order autoregressive one (AR2). This was the predominant specification used to produce the alternative forecasts against which to compare the consensus forecasts.

The MSCI data analysed here were the changes for all the variables under investigation – effectively first differences of the log transformed original quantities of each series. Hence, the integration term D in the standard annotation of ARIMA (p, D, q) models was not applicable in any of the models estimated. We performed Ljung-Box and Engel heteroscedasticity tests with the 32 actual series for the periods 1981–2005 and 1981–2017 and found that both autocorrelation and heteroscedasticity were present. We repeated the tests after fitting the best performing models finding no significant correlation or conditional heteroscedasticity in the (standardised) residuals with the exception of two cases in the 1981–2017 series. By and large, the best fitted models were second order autoregressive AR(2) with the exception of rental growth in Retail Warehouses, AR(1). Sector series for equivalent yield were generally dominated by AR(2) models with the exception of years 2007 and 2008 where AR(3) specifications provided the best fit alternative (see [Appendix 2](#) and [Appendix 3](#) for further details).

As noted earlier, in this context there is a specific focus on implied yield forecasts and on the ability of real estate forecasters to predict relative sector performance. Whilst there are no publicly available forecasts of yields, with an element of approximation implied forecasts of yield shifts⁵ can be ‘backed-out’ of the relationship between capital and rental growth forecasts. The underlying rationale is that, since capital values are fundamentally a function of rents and yields, *assuming stable vacancy rates and reversionary periods*, at the aggregate index level it is only changes in rental growth and yields that will produce changes to the capital value of the index.

The intuition here is that, if a forecaster is forecasting rents to grow by 10% in a year and is also forecasting capital value growth to be 10% in that year, it is expected that the increase in Market Rent by itself will produce all the capital value growth and it can be inferred that no change in yield is being forecasted. However, if the forecast was for capital value growth to be 15% and rental growth to be 10%, then there is an implied forecast of a (downward) yield shift that is producing the added 5% capital growth. Hence, when rental and capital growth forecasts are known, implicit forecasts for future yields can be estimated. The implied forecast of future yield shifts can be calculated more formally as:

$$\Delta Y_{t+1} = \left[\frac{(1 + \Delta \text{rent})}{(1 + \Delta \text{capital})} - 1 \right] \quad (2)$$

Where

ΔY_{t+1} is the one period ahead implied yield shift in percentage terms

Δrent is the one period ahead forecasted rate of rental growth

$\Delta \text{capital}$ is the one period ahead forecasted rate of capital growth

For example, if the current yield is 5%, the one year ahead rental growth forecast is 20% and the capital growth forecast is 9%, the implied forecast of the relative yield *shift* in a year’s time is $(1.20/1.09) - 1 = 0.1009$ (10.09%) and the yield is forecasted to increase to 5.5% ($0.1009 \times 0.05 + 0.05$).

Given that at the index level negligible variation in average reversionary period would be expected, expected or actual differences in vacancy levels could also produce changes in expected or actual capital growth. In the past, changes in vacancy rates seemed to have had limited impact on capital growth shifts. Ranging from an upward shift of 210 basis

points in 2008 to a downward shift of 75 basis points in 2010, on average the actual MSCI All Property Equivalent Yield series has by shifted 53 basis points every year since 2005. An implied series of equivalent yield shifts estimated from actual changes in rental growth and capital growth alone are closely aligned to the actual equivalent yield shifts. There is an average difference between the actual and implied yield shift of seven basis points per annum. This suggests that forecasts of capital growth would largely be a product of rental and yield forecasts. However, where capital growth forecasts are low, it is possible that they may reflect expectations of fluctuations in vacancy rates instead of or as well as expected yield shifts and rental growth. In addition, it should be noted that the MSCI Equivalent Yield time series is based upon a changing sample of properties. However, its Equivalent Yield Shift time series is based upon a 'held' sample. Put simply, a small part of the yield shift will also be due to changing sample composition rather than market change. As such, implied forecasts of small yield shifts based only on the relationship between rental and capital growth forecasts should be treated with caution.

V Results and findings

Table 1 displays the summary statistics for the (consensus) forecasted and actual outcomes for the one year ahead forecasts. The summary data provides preliminary evidence of a pessimism bias in the forecasts. With few exceptions, the mean forecast *over the whole period* for all sectors and variables forecasted is lower than the actual outcome. This pessimism bias is less clear for the implied equivalent yield shift forecasts. The broad retail sector (incorporating standard shops, shopping centres and retail warehouses) has been the worst performing in terms of capital growth, rental growth and total returns with City and West End Offices performing best. This pattern is also found in the forecasts.

The evidence on forecast smoothing is more mixed. For rental growth, the standard deviation of the forecasted and actual growth rates are similar with the forecasted series displaying slightly higher volatility on average. This is not in line with expectations. Whilst the standard deviations were similar, for all sectors the consensus forecasts systematically underestimated the peak of the market whilst overestimating its lowest point over the sample period. The 'cancelling out' effect caused by the under-estimation of the upside variance combined with the over-estimation of the downside variance resulted in similar levels of volatility in the times series for the actual rental growth rates and forecasted rental growth rates. For capital growth, there is some evidence of forecast smoothing. For all sectors, the standard deviation of the forecasted capital returns is lower than the actual capital returns. This is reflected in the forecasts of the market peaks and troughs. For all sectors, the consensus forecasts systematically underestimated the peak and the low point of the market cycle. This pattern is largely replicated for the total return forecasts.

Bearing in mind that positive yield shifts reflect decreases in capital values, the interpretation of positive and negative figures is the opposite as for the other variables. The forecasted yield shifts are typically more optimistic than the actual. For All Property, the mean forecasted Equivalent Yield Shift is 1.2 whilst the actual is 1.6 indicating that the consensus expectation was for lower increases in yields. The standard deviation of the



Table 1. Summary statistics: Actual outcomes and consensus forecasts for MSCI IPD UK property indices 2006–2017.

	Actual				Consensus Forecasts					
	Mean	Standard Deviation	Max.	Min.	Serial correlation	Mean	Standard Deviation	Max.	Min.	Serial correlation
<i>Rental value growth (%)</i>										
OFFICES	2.7	6.5	10.8	-13.5	0.3	1.2	6.0	6.8	-14.7	0.3
INDUSTRIAL	1.0	2.9	5.3	-4.6	0.7	0.0	3.0	3.5	-7.5	0.5
STANDARD SHOPS	1.0	3.1	4.1	-7.2	0.5	-0.3	2.9	2.5	-7.3	0.5
SHOPPING CENTRES	0.2	2.4	4.8	-4.6	0.5	-0.4	2.7	2.7	-6.5	0.4
RETAIL WAREHOUSE	0.2	2.2	3.0	-6.0	0.3	0.1	3.0	3.7	-7.7	0.4
ALL PROPERTY	1.2	3.4	5.6	-7.6	0.4	0.2	4.0	3.8	-10.3	0.3
City Offices	4.2	9.1	13.6	-17.7	0.3	1.6	8.9	11.6	-21.8	0.2
West End Offices	4.8	10.3	17.8	-19.9	0.3	2.5	8.2	10.4	-19.7	0.2
<i>Capital value growth (%)</i>										
OFFICES	2.7	12.1	17.3	-26.4	0.3	0.9	8.6	9.6	-20.4	0.2
INDUSTRIAL	2.0	11.7	16.4	-26.0	0.4	-0.2	6.7	6.6	-15.9	0.2
STANDARD SHOPS	1.5	9.8	12.1	-23.1	0.4	-0.7	6.3	5.9	-15.3	0.1
SHOPPING CENTRES	-1.3	10.5	10.6	-26.6	0.3	-1.0	6.5	5.9	-16.0	0.1
RETAIL WAREHOUSE	-1.3	11.3	10.6	-29.7	0.1	-0.4	7.1	6.9	-17.2	0.1
ALL PROPERTY	1.1	10.5	12.6	-25.9	0.3	-0.1	7.3	7.0	-17.9	0.1
City Offices	4.3	13.7	19.8	-27.9	0.3	1.5	10.3	12.2	-23.7	0.2
West End Offices	6.3	13.4	25.0	-25.8	0.3	2.4	10.1	11.5	-23.3	0.2
<i>Total return (%)</i>										
OFFICES	7.9	12.5	23.1	-22.4	0.3	6.6	8.5	14.6	-14.0	0.1
INDUSTRIAL	8.3	12.1	23.4	-21.2	0.3	6.4	6.5	13.5	-8.4	0.2
STANDARD SHOPS	6.6	10.1	17.5	-18.8	0.3	4.7	6.1	12.6	-8.4	0.1
SHOPPING CENTRES	4.1	11.0	17.7	-22.6	0.3	4.9	6.2	12.6	-8.4	0.1
RETAIL WAREHOUSE	4.4	12.1	16.5	-25.6	0.1	5.3	7.0	13.5	-10.5	0.1
ALL PROPERTY	6.6	11.0	18.1	-21.7	0.2	5.7	7.2	13.4	-11.3	0.1
City Offices	9.1	14.2	25.1	-24.2	0.3	6.7	10.3	17.2	-17.6	0.1
West End Offices	10.8	13.8	30.6	-22.3	0.2	6.9	10.1	16.0	-17.9	0.1
<i>Equivalent Yield Shift</i>										
OFFICES	1.2	12.0	33.4	-10.2	0.3	1.0	14.1	29.0	-18.5	0.0
INDUSTRIAL	0.7	12.7	35.8	-12.9	0.3	0.4	15.2	37.0	-17.3	0.1
STANDARD SHOPS	0.7	11.4	31.7	-8.8	0.3	0.1	14.4	32.0	-22.1	0.0
SHOPPING CENTRES	2.8	12.4	37.0	-11.9	0.3	1.3	14.4	39.1	-11.8	0.2
RETAIL WAREHOUSE	3.1	14.0	40.9	-10.0	0.2	3.4	17.7	43.2	-26.3	0.0
ALL PROPERTY	1.6	12.1	35.1	-9.3	0.3	1.2	14.6	34.1	-19.6	0.0
City Offices	1.1	12.2	31.3	-11.4	0.4	1.0	13.6	31.8	-16.3	0.0
West End Offices	-0.3	11.7	28.5	-13.4	0.2	-0.2	14.5	29.3	-25.9	-0.1

forecasted and actual equivalent yields are broadly similar. However, the forecasted series displays consistently higher volatility than the actual series on average. Finally, whilst the sample time frame provides a fairly small number of annual observations, the results for serial correlation provide little evidence to indicate that forecasters tend to be more ‘sticky’ in their forecasts compared to the index. For all sectors and all of the variables, the level of serial correlation in the forecasted time series is generally lower for the forecasted series than for the actual series.

Error metric measures of forecast accuracy

Table 2 presents the results for the conventional error metric tests of forecasting accuracy for consensus and the ARMA models. The Mean Error is essentially a measure of forecasting bias. For rental growth, the results are consistent with the analysis above that identified a systematic tendency for the consensus forecast to be more pessimistic than the actual outcome for all sectors. However, it should be noted that there are specific periods, such as 2008, when the consensus forecasts were optimistic and underestimated

Table 2. Forecasting accuracy: Error metrics.

	Consensus Forecast			ARMA		
	ME	MAE	RMSE	ME	MAE	RMSE
Rental value growth (% p.a.)						
OFFICES	-1.5	3.0	3.5	0.5	4.2	6.2
INDUSTRIAL	-1.0	1.3	1.7	1.9	3.0	3.5
STANDARD SHOPS	-1.4	1.4	1.6	3.5	3.5	4.7
SHOPPING CENTRES	-0.6	0.9	1.2	4.6	4.6	5.1
RETAIL WAREHOUSE	-0.1	1.0	1.2	4.9	4.9	5.4
ALL PROPERTY	-1.0	1.7	1.9	2.6	2.9	4.2
City Offices	-2.6	4.8	5.5	-1.2	6.7	8.8
West End Offices	-2.3	4.5	6.0	-0.2	7.1	9.8
Capital value growth (% p.a.)						
OFFICES	-1.7	8.1	9.5	-0.1	8.9	11.7
INDUSTRIAL	-2.1	7.5	9.1	0.9	8.5	11.4
STANDARD SHOPS	-2.2	6.5	7.8	2.8	6.8	10.0
SHOPPING CENTRES	0.2	5.3	7.1	4.8	8.2	11.3
RETAIL WAREHOUSE	0.9	7.3	10.4	7.0	8.7	13.2
ALL PROPERTY	-1.2	6.8	8.5	2.3	7.4	10.5
City Offices	-2.8	10.2	11.4	-1.6	10.6	13.2
West End Offices	-4.0	9.9	11.2	-1.5	9.9	12.9
Total return (% p.a.)						
OFFICES	-1.3	8.0	9.8	1.4	9.2	12.2
INDUSTRIAL	-1.9	7.6	9.5	3.0	8.9	12.3
STANDARD SHOPS	-1.9	6.5	7.9	3.6	7.2	10.6
SHOPPING CENTRES	0.8	5.1	7.3	5.7	8.8	12.2
RETAIL WAREHOUSE	0.9	7.5	10.7	8.5	9.5	14.7
ALL PROPERTY	-0.9	6.8	8.7	3.5	7.6	11.3
City Offices	-2.4	10.2	11.6	-0.1	10.8	13.6
West End Offices	-3.9	10.0	11.4	0.4	10.0	13.2
Equivalent Yield Shift (% p.a.)						
OFFICES	-0.2	9.9	13.5	-0.6	8.5	12.1
INDUSTRIAL	-0.3	10.4	14.4	-0.3	8.6	12.9
STANDARD SHOPS	-0.7	8.8	13.3	-0.6	7.9	11.6
SHOPPING CENTRES	-1.5	7.9	12.4	-2.4	7.8	12.8
RETAIL WAREHOUSE	0.3	11.9	18.4	-3.3	9.4	14.7
ALL PROPERTY	-0.4	9.4	13.7	-1.3	8.1	12.3
City Offices	-0.1	9.1	12.7	-0.5	9.2	12.3
West End Offices	0.1	9.2	14.1	-0.2	8.8	11.8

the extent of the market fall. Given the low rates of nominal rental growth in the sample period, the Mean Absolute Error in the rental forecast is typically higher than the actual rate of rental growth. The potential scaling problem can be seen in that the Mean and Mean Absolute Errors tend to be higher when the actual rate of growth is higher. Hence, the West End Office and City Office consensus forecasts had the highest levels of inaccuracy. In contrast, the ARMA forecasts were generally optimistic. For instance, in the consensus forecasts the ME for All Property was -1.0% . For the ARMA forecasts, the equivalent figure was 2.6% . A lot of the error in the ARMA models was in the retail sectors where the ARMA models performed particularly poorly.

Turning to capital growth forecasts, the level of Mean Error of the consensus forecast is broadly similar to rental growth. The most notable result is the size of the Mean Absolute Error. For instance, All Property had an average rate of capital growth of 2% per annum in the sample period. Forecasters were systematically pessimistic producing a Mean Error of -1.2% with the actual being higher than the forecasted. However, the average difference between the achieved capital growth and the forecasted capital growth in any given year was 6.8% . For the whole sample period, this average figure was highest for City Offices at 10.3% . It peaked in 2008 with the average difference between the actual and forecasted rate of capital growth at 20% . Again, the ARMA models were generally optimistic. For instance, in the consensus forecasts the ME for All Property Capital Growth was -1.2% . For the ARMA

forecasts, the equivalent figure was 2.3% . Whilst usually higher in terms of MAE, the ARMA forecasts had similar levels of forecasting performance as the consensus forecasts.

For the equivalent yield shift forecasts, the level of Mean Error looks small compared to rental and capital growth, except for shopping centres. The results do not indicate significant bias in the *implied* consensus forecasts of equivalent yield shift. Whilst the yield shift forecasts for most sectors were more optimistic than the actual, the large yield forecasting errors in the 2007–2009 period produced this result. The ARMA conditional mean forecasts had much higher levels of Mean Error compared to the consensus. Again, the most notable result in the consensus forecasts is the size of the Mean Absolute Error. The average difference between the actual equivalent yield and the forecasted equivalent yield in any given year was approximately 0.5% . The effect of such an error on capital growth depends on the equivalent yield at the start of the forecasting period. For instance, a 0.5% downward equivalent yield shift from 4.5% to 4% would, all else equal, produce an increase in capital value of approximately 12.5% . The comparative increase if the starting equivalent yield were 6% is 8.3% . The ARMA models performed much more poorly in terms of Mean Error. For instance, in the consensus forecasts the ME for All Property Equivalent Yield was -0.4 . For the ARMA forecasts, the corresponding figure was -1.3% . In terms of MAE, the ARMA forecasts had similar but consistently lower levels of error relative to the consensus forecasts.

The discussion above has hinted at the fact that forecasting uncertainty may be concentrated in certain periods and that there may be variations in the reliability of forecasts in different market conditions. Figures 1–3 display the actual and forecasted rental growth, capital growth and equivalent yield shift figures for the All Property Index. It is clear that the largest forecasting errors occurred during the Global Financial Crisis. As the downturn started in 2007, the one year ahead forecast implied minimal change in the equivalent yield. The All Property equivalent yield rose from 5.4% to 6.1% . As the

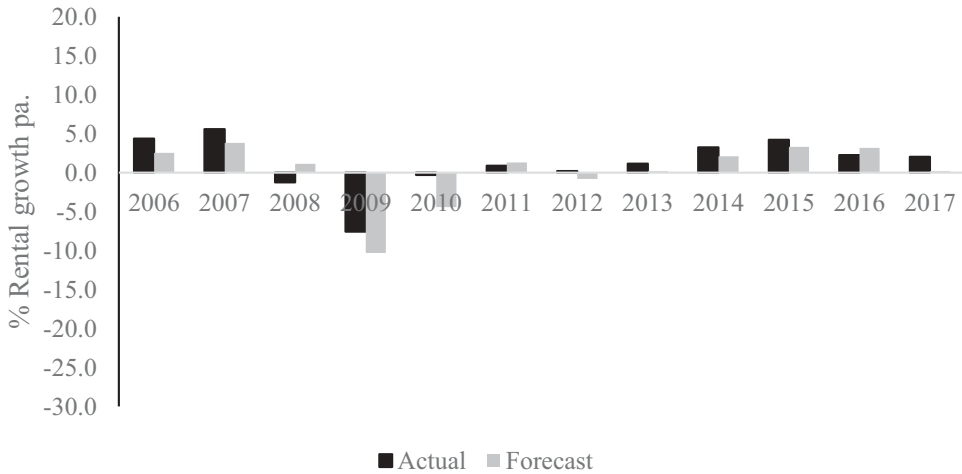


Figure 1. All Property – Rental growth.

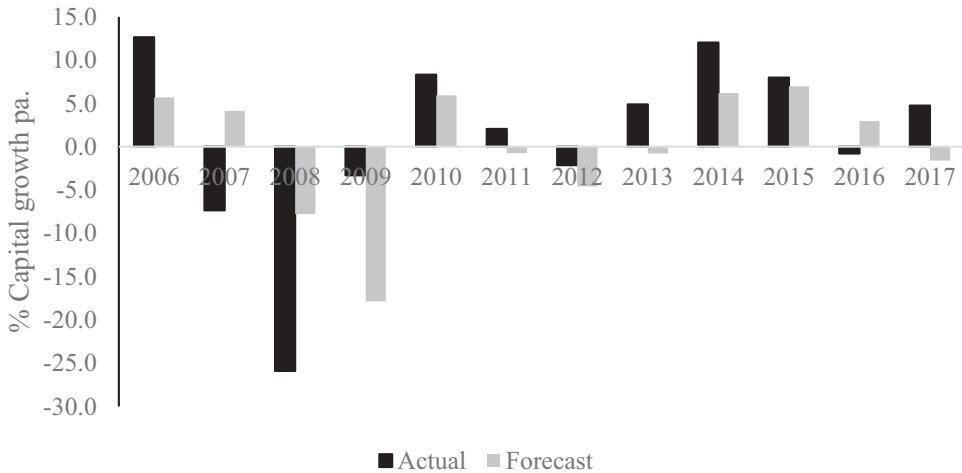


Figure 2. All Property – Capital growth.

downturn intensified in 2008, the consensus forecast was for an increase in the equivalent yield from 6.1% to 6.7%. All else equal, this yield shift would result in a capital growth of approximately -9%. This proved to be a considerable underestimation of the actual yield shift with the All Property Equivalent Yield rising to 8.2%. In reality, the yield shift itself produced a fall in capital values of around -25%. The consensus forecast anticipated that the market would continue to deteriorate in 2009 with the All Property Equivalent Yield rising to 8.9%. In fact, there was a degree of market stabilisation with the All Property equivalent yield falling from 8.2% to 8.0%. In terms of the total Mean Absolute Error in the equivalent yield forecasts over the whole 13 year period, the years 2007, 2008 and 2009 account for approximately 72% of the forecasting error. The comparable figure for rental growth is 42%. If errors were spread equally, it would be expected that this period would account for around 23% of the total forecast error.

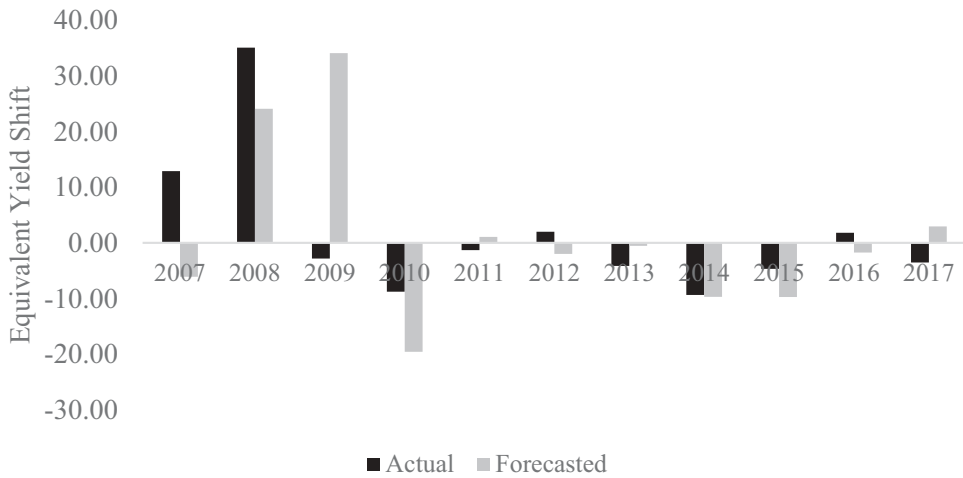


Figure 3. All property equivalent yield shift.

Forecasts of relative performance

We now turn to the relative measures of forecasting performance of the various sectors. Figures 4 and 5 display the Kendall's W Coefficient of Concordance (equation 1) for the various performance metrics over the sample period. Figure 4 shows results for the comparison between actual and consensus forecasts and Figure 5 shows the results for the comparison between actual and ARMA model forecasts. Broadly, in contrast to Ling's (2005) findings in the US for the RERC survey, the results suggest that the consensus forecasts display a high degree of reliability in forecasting the relative performance of the

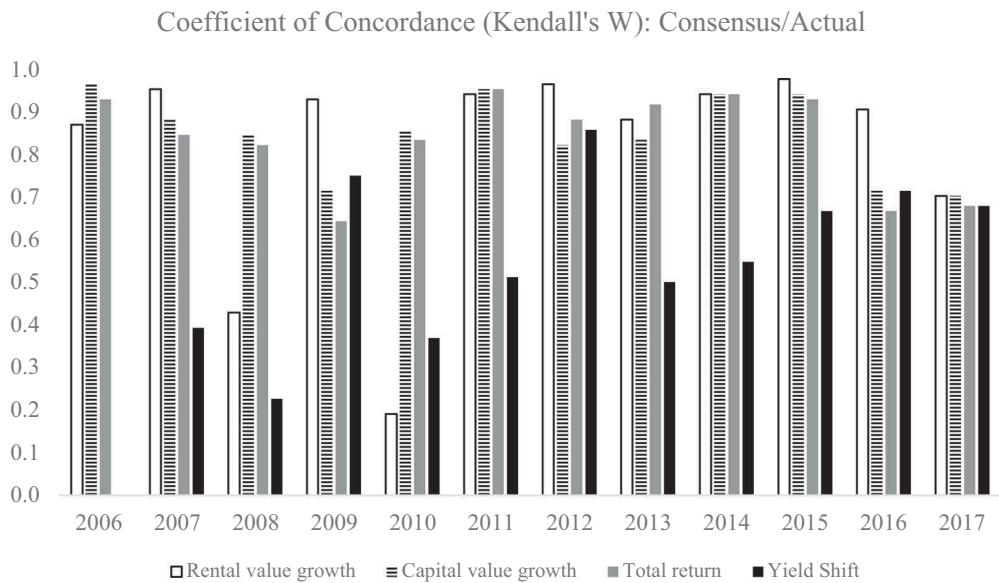


Figure 4. Coefficient of concordance in sector ranking for consensus forecasts.

Coefficient of Concordance (Kendall's W): ARMA/Actual

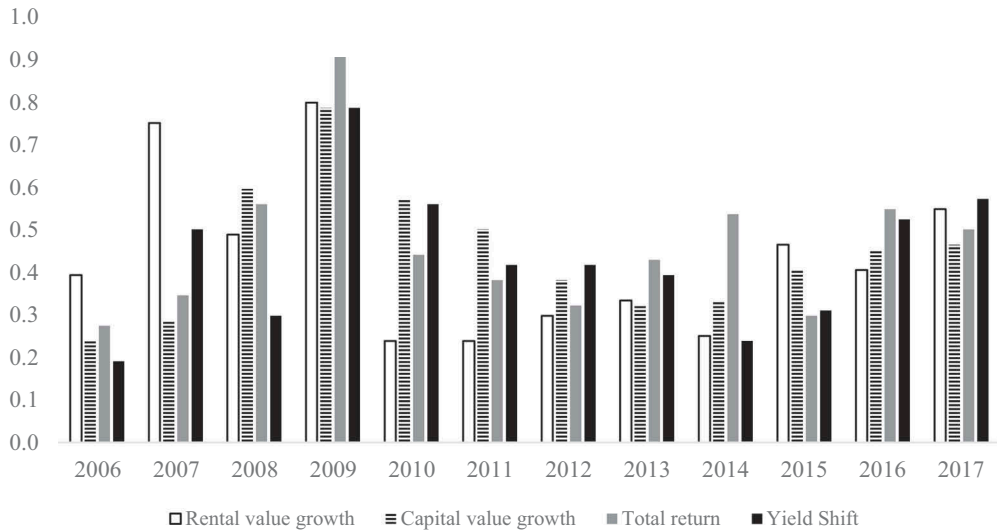


Figure 5. Coefficient of concordance in sector ranking for ARMA Forecasts.

various sectors. In the vast majority of cases, there is little ‘disagreement’ between the actual sector ranking and the forecasted sector ranking. For rental growth, capital growth and total return, in 60%-70% of observations, the forecasted ranking was the same as or different by one rank from the actual rank.

It is difficult to reconcile the stark contrast between the findings of Ling’s study in the US published in 2005 and these current findings for the UK. A key difference between the two studies is the composition of the survey samples. The US study is based upon a survey of institutional investors and managers whilst the IPF Consensus Forecasts are based on a survey of specialist professional forecasters. Investors may tend to anchor on past real estate market performance whilst forecasters may be more inclined to anchor on forecasted variables e.g. GDP or inflation forecasts. However, the investor expectations survey by RERC may also have been influenced by real estate forecasts and that forecasters as well as investors display adaptive expectations. Finally, it is also possible that the difference in findings may be due to cultural differences between UK and US markets.

Again the consensus forecasts for equivalent yield shifts were less reliable. The comparable figure for the equivalent yield forecasts is approximately 40% of the forecasted rankings being the same as or different by one rank from the actual rank. Once more, there are variations in the level of forecast reliability over time. Consistent with the findings above, the forecasts of equivalent yield displayed their lowest correlation with the actual rankings during the Global Financial Crisis. There was actually a negative (Spearman’s Rank) correlation coefficient between the sector rankings in terms of equivalent yields and the actual sector ranking of equivalent yields in six of the 13 years. However, it is notable that this did not typically result in poor relative forecast performance for total returns or capital growth. In periods when there was poor performance in sector ranking of equivalent yields (e.g. in 2007, 2008, 2010 and 2103),

there was not necessarily poor forecasting performance in ranking sectors in terms of total return and/or capital growth. A clear-cut finding is that the conditional mean ARMA forecasts tend to perform much more poorly in terms of predicting the relative performance of the various sectors.

The final evaluation of the consensus forecasts assesses their performance relative to simple (so-called naïve) forecasts and ARMA models via Theil's inequality coefficient (U2). Additionally, for both consensus and ARMA forecasts we investigate the source of the forecast error by decomposing the U2 coefficient into bias, variance and covariance proportions. The results are displayed in Table 3. Consistent with the analysis above, consensus rental growth forecasts tend to display the best performance. The RMSE of the consensus forecasts of rental growth is typically just over half the naïve forecasts. In contrast, the conditional mean ARMA forecasts tend to be beaten by naïve predictions. The decomposition of the error source indicates that most problems lie with systematic forecasting error in the three retail sectors. On the other hand, the error source in the consensus forecasts is dominated by unsystematic forecasting error – the 'least bad' source of forecasting error.

Table 3. Forecasting accuracy: Theil's U2 statistic.

	Consensus Forecasts				ARMA Forecasts			
	U2 stat	Error source			U2 stat	Error source		
		bias	variance	covariance		bias	variance	covariance
Rental value growth (%)								
OFFICES	0.51	19%	2%	79%	0.91	1%	97%	2%
INDUSTRIAL	0.56	37%	1%	62%	1.19	30%	51%	19%
STANDARD SHOPS	0.52	68%	2%	30%	1.47	55%	29%	16%
SHOPPING CENTRES	0.50	23%	5%	72%	2.21	82%	10%	8%
RETAIL WAREHOUSE	0.59	0%	45%	55%	2.54	83%	8%	9%
ALL PROPERTY	0.55	29%	8%	63%	1.19	38%	49%	13%
City Offices	0.57	23%	0%	77%	0.90	2%	97%	1%
West End Offices	0.55	14%	13%	73%	0.89	0%	100%	0%
Capital value growth (%)								
OFFICES	0.80	3%	14%	83%	0.98	0%	100%	0%
INDUSTRIAL	0.81	5%	30%	65%	1.00	1%	94%	5%
STANDARD SHOPS	0.82	8%	20%	72%	1.05	8%	84%	8%
SHOPPING CENTRES	0.70	0%	32%	68%	1.12	18%	74%	8%
RETAIL WAREHOUSE	0.95	1%	16%	83%	1.21	28%	62%	10%
ALL PROPERTY	0.84	2%	14%	84%	1.04	5%	91%	4%
City Offices	0.83	6%	9%	85%	0.96	1%	99%	0%
West End Offices	0.78	12%	9%	79%	0.90	1%	99%	0%
Total return (%)								
OFFICES	0.68	2%	17%	81%	0.85	1%	98%	1%
INDUSTRIAL	0.67	4%	34%	62%	0.86	6%	84%	10%
STANDARD SHOPS	0.67	6%	26%	68%	0.90	12%	80%	8%
SHOPPING CENTRES	0.65	1%	43%	56%	1.08	22%	69%	9%
RETAIL WAREHOUSE	0.87	1%	22%	77%	1.19	34%	55%	11%
ALL PROPERTY	0.70	1%	19%	80%	0.91	10%	84%	6%
City Offices	0.71	4%	11%	85%	0.83	0%	100%	0%
West End Offices	0.67	12%	10%	78%	0.77	0%	100%	0%
Equivalent Yield Shift								
OFFICES	0.96	0%	3%	97%	1.01	0%	100%	0%
INDUSTRIAL	0.95	0%	3%	97%	1.02	0%	99%	1%
STANDARD SHOPS	0.93	0%	6%	94%	1.01	0%	99%	1%
SHOPPING CENTRES	0.85	1%	3%	96%	1.01	4%	94%	2%
RETAIL WAREHOUSE	1.02	0%	4%	96%	1.02	5%	93%	2%
ALL PROPERTY	0.94	0%	4%	96%	1.01	1%	99%	0%
City Offices	0.93	0%	1%	99%	1.00	0%	100%	0%
West End Offices	0.97	0%	4%	96%	1.01	0%	100%	0%

Turning to capital growth, whilst all sectors beat the naïve prediction, the consensus forecasts display weaker performance compared to consensus rental growth forecasts. The ARMA capital growth forecasts improve compared to their rental growth counterparts. They broadly match the performance of the naïve predictions. Even more so than for rental growth, the error source in the consensus forecasts is dominated by unsystematic forecasting error. In contrast, the major source of error in the ARMA forecasts is an inability to capture the volatility of the times series. To a large extent, this is intrinsic to the ARMA forecasting model which does not condition upon variance.

Throughout the paper, the forecasting performance for total returns has not been discussed in detail because it tends to be approximately between the forecasting performance of capital and rental growth. This pattern holds in the U2 statistics for consensus forecasts of total returns but not for the ARMA forecasts. In terms of error source in the consensus forecasts, it is unsystematic forecast error that is the predominant cause of forecast failure. For the ARMA forecasts, similar to rental and capital growth it is the intrinsic inability to reflect the volatility of the time series that is the key determinant of forecast error. For equivalent yield shift forecasts, there is little to distinguish the performance of the naïve and consensus forecasts. A similar pattern can be seen for the ARMA forecasts. However, in terms of error source, for consensus forecasts practically all forecast error can be attributed to unsystematic factors. Whereas the dominant source of forecast error in the ARMA is once again the inability to model the variance of the time series.

VI Conclusion

Evaluating forecast accuracy is an inherently subjective process. The extent to which a failure to correctly predict the future is costly to the forecast user depends on a wide range of factors. As a result, in the broader forecasting literature the identification of appropriate measures of forecast accuracy in order to compare forecasters against each other and in terms of their ability to improve decision-making has generated a large body of analysis and discussion. However, in economic, business and financial market forecasting, it is generally accepted that forecasts will be prone to non-trivial errors, can produce 'forecasting debacles' following major unanticipated shocks and can be prone to systematic biases. The fact that previous research has found evidence of these phenomena in real estate forecasts should be expected.

Many of the findings from this research are consistent with the results of previous research. There was evidence of a systematic tendency towards pessimism. In line with previous research, rental growth forecasts tend to be more reliable than capital growth forecasts. Compared to capital growth, they have lower (mean absolute) error and negligible difference in the volatility of forecasted and actual rental growth rates. It is the typically large errors in capital growth forecasts that are then the source of most of the error in total return forecasts. There is evidence of forecast smoothing here with the forecasted capital growth time series exhibiting lower volatility than the actual capital growth rates. However, it is also notable that the forecasts tend to display less serial correlation than the actual series. Given that the main sources of error in capital growth forecasts will be due to errors in the forecasts of rental growth and equivalent yield, the relatively low (mean absolute) error in rental growth forecasts indicates that it must be the error in the yield forecasts that is the guilty party.

Whilst the IPF do not include forecasts of yields in the survey, it is failure in yield forecasting that is the main source of failure in forecasts of capital growth and also in forecasts of total return. With an average variation from the actual yield of 0.5%, the consequence tends to be large errors in the estimation of capital growth. However, the picture is more nuanced. The implicit yield forecasts inferred from rental and capital growth forecasts display little evidence of forecast smoothing. The actual and forecasted yield time series have similar levels of volatility. Much of the yield forecast error was concentrated in the Global Financial Crisis. The period from 2007 to 2009 accounted for approximately 60% of the total error in the forecasts of equivalent yield shift. In terms of measuring relative sector performance, the equivalent yield forecasts also experienced their worst performance in 2007 and 2008.

Whilst there has been little previous research on the issue, the findings indicate that the consensus forecasts generally provide reliable indicators of the relative performance of the sectors included in the survey. Although there are some years of comparatively poor prediction of relative performance, for rental growth, capital growth and total return, the forecasted sector rankings at the beginning of the year tend to be similar to the outcome at the end of the year. Again, forecasted sector rankings for the equivalent yield tended to have the least reliability. However, the focus here has been on one-year ahead forecasts. Given the time scales involved in acquisition and disposal, the two and five year horizons may be more appropriate for investors. The findings of the error decomposition analysis provide some reassurance to forecasting practitioners. The vast majority of the error source in consensus forecasts is in the unsystematic component caused by random shocks. However, there is also evidence, particularly for rental growth, that there is a conservative bias. On average, the consensus forecasts tended to be lower than the actual outcome for all sectors across the study period.

Finally, it is with some caution that we turn to the practical implications of these findings. To paraphrase Niels Bohr, it is difficult to predict the commercial real estate market, especially market yields. This weakness in yield forecasting is transmitted directly to capital growth forecasts. Forecast users can place much more confidence in rental growth predictions but it is also worth bearing in mind that the consensus forecasts have been systematically pessimistic about this variable. The main challenge for producers of commercial real estate forecasts is yield shifts which are more closely related to the more volatile capital markets (compared to the relatively 'sticky' real economy). Predictability is likely to stay lower for yields. The most recent research incorporating a range of sentiment indicators into yield forecasting models has found little improvement in forecasting performance compared to more established and even naïve models (see Heinig, Nanda, & Tsolacos, [forthcoming](#)). The performance of real estate forecasts during the GFC particularly highlighted the need for users of forecasts to place less weight on them in periods of market volatility.

Notes

1. The Investment Property Forum (IPF) is a London based individual members organisation with a 2000+ diverse membership including investment agents, fund managers, bankers, lawyers, researchers, academics, actuaries and other related professionals operating in the UK property investment market.

2. The terms 'yield', 'equivalent yield' and 'capitalisation rate' should be regarded as effectively interchangeable in this context. 'Equivalent yield', which is mainly used in the UK and Ireland, incorporates particular attributes of the typical lease structures. Throughout the paper, the term 'equivalent yield' is used.
3. Whilst investment banks, advisors and institutional investors have been the three main categories, not all contributors fit into these categories. In particular, a small number of specialist forecasting organisations have participated e.g. Capital Economics, Real Estate Forecasting Limited. In addition, over the last decade the participation of investment banks has decreased. In 2019, the asset management arm of UBS was the only investment bank participating.
4. For instance, several years ago, one former forecaster for a global real estate advisory firm informed one of the authors that they had been successfully pressured by one of the firm's leading investment agents to change an initial forecast for City of London office rents because it might result in a buyer withdrawing from a large deal.
5. It should be noted the Equivalent Yield Shift is measured in relative terms by MSCI. For instance, if the yield at $t = 0$ is 5% and at $t = 1$ it increases to 5.5%, the equivalent yield shift is calculated as $(0.055 - 0.05) / 0.05 * 100 = 10$. This effectively represents a 10% increase in the yield relative to the reference year. In absolute terms, the equivalent yield shift is 0.5% since the equivalent yield increases from 5% to 5.5%.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Pat McAllister is a professor of Real Estate at the Department of Real Estate and Planning, University of Reading. His research interests are focused on international real estate investment, valuation and real estate development.

Ilijir Nase is an assistant professor (Real Estate Management) at the Department of Management in the Built Environment, Delft University of Technology. His main areas of research activity include real estate and urban economics, the interface of urban planning/design and real estate development and corporate real estate management.

References

- Batchelor, R. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting*, 23(2), 189–203.
- Baym, A., & Hartzell, D. (2012). *Global property investment: Strategies, structures, decisions*. Chichester: Wiley.
- Bond, S., & Mitchell, P. (2011). The information content of real estate derivative prices. *Journal of Portfolio Management*, 37(5), 170–181. (Special Real Estate Issue).
- Box, G. E., Hunter, W. G., & Hunter, J. S. (1978). *Statistics for experimenters*. New York, NY: John Wiley & Sons.
- Chaplin, R. (1998). An ex-post comparative evaluation of office rent prediction models. *Journal of Property Valuation & Investment*, 16, 21–37.
- Chaplin, R. (2000). Predicting real estate rents: Walking backwards into the future. *Journal of Property Investment and Finance*, 18, 352–370.
- Clements, M., & Hendry, D. (1999). *Forecasting Non-stationary Economic Time Series*. Boston, MA: MIT Press.

- Gallimore, P., & McAllister, P. (2003). *Judgment and quantitative forecasts in commercial property investment decisions*. London, UK: Report for RICS Education Trust.
- Gallimore, P., & McAllister, P. (2005). Expert judgement in the processes of commercial property market forecasting. *Journal of Property Research*, 21, 337–360.
- Granger, C., & Pesaran, M. (2000). Economic and statistical measures of forecast accuracy. *Journal of Forecasting*, 19, 537–560.
- Heinig, S., Nanda, A., & Tsolacos, S. (forthcoming). Which sentiment indicators matter? Evidence from the European commercial real estate market. *Journal of Real Estate Research*.
- Hendry, D. F. (2003). Forecasting pitfalls. *Bulletin EU and US inflation and macroeconomic analysis* (BIMA), 100, 65–82.
- Hendry, D. F., & Clements, M. P. (2003). Economic forecasting: Some lessons from recent research. *Economic Modelling*, 20(2), 301–329.
- Ling, D. (2005). A random walk down Main Street: Can experts predict returns on commercial real estate? *Journal of Real Estate Research*, 27, 137–154.
- Marcato, G., & Nanda, A. (2016). Information content and forecasting ability of sentiment indicators: Case of real estate market. *Journal of Real Estate Research*, 38(2), 165–203.
- McAllister, P., Newell, G., & Matysiak, G. (2006). *Disagreement and uncertainty in UK commercial property market forecast: Research findings*. London, UK: Report for Investment Property Forum.
- McAllister, P., Newell, G., & Matysiak, G. (2008). Agreement and accuracy in consensus forecasts of the UK commercial property market. *Journal of Property Research*, 25, 1–22.
- Nordhaus, W. (1987). Forecasting efficiency: Concepts and applications. *Review of Economics and Statistics*, 69, 667–674.
- Papastamos, D., Matysiak, G., & Stevenson, S. (2015). Assessing the accuracy and dispersion of real estate investment forecasts. *International Review of Financial Analysis*, 42, 141–152.
- Perera, T., Higgins, D., & Wong, W. W. (2018). The evaluation of the Australian office market forecast accuracy. *Journal of Property Investment & Finance*, 36(3), 259–272.
- Stevenson, S., & McGarth, O. (2003). A comparison of alternative rental forecasting models: Empirical tests on the London office market. *Journal of Property Research*, 20(3), 235–260.
- Theil, H. (1966). *Applied economic forecasting*. Amsterdam: North-Holland Publishing Company.
- Tsolacos, S. (2006). An assessment of property performance forecasts: Consensus versus econometric. *Journal of Property Investment & Finance*, 24(5), 386–399.
- Watkins, C., White, M., & Keskin, B. (2012). *The future of property forecasting*. London, UK: Investment Property Forum.
- Wilson, P., Okunev, J., Ellis, C., & Higgins, D. (2000). Comparing Univariate Forecasting Techniques in Property Markets. *Journal of Real Estate Portfolio Management*, 6(3), 283–306.

Appendix 1. Independent and Relative Measures of Forecast Performance

Commonly, forecast error is measured as the difference between actual outcome and forecasted outcome. In more formal notation, this can be expressed as

$$e_t = F_t - A_t \quad (3)$$

Where e_t represents the forecast error, A_t is the actual outcome and F_t is the forecasted outcome. For a series of forecasts, the Mean Error (ME) measures the average difference between the actual and forecasted outcomes. It is a simple and frequently used measure of forecast accuracy. As negative and positive errors cancel each other out, it is a measure of bias – whether there is a systematic tendency for forecasts to be over or under-estimated. For a set of N forecasts, it can be expressed more formally as

$$ME = \frac{1}{N} \sum_{t=1}^N e_t \quad (4)$$

Given that positive and negative errors can cancel each other out, the ME metric can mask substantial forecasting inaccuracy. It provides no indication of the variance in the errors with the possibility that lots of the individual errors could be large. The Mean Absolute Error (MAE) overcomes this limitation by providing the average error regardless of whether it is positive or negative.

$$MAE = \frac{1}{N} \sum_{t=1}^N |e_t| \quad (5)$$

A commonly used alternative to the MAE is the Mean Squared Error (MSE). Whilst similar to the MAE, it puts a larger penalty on large forecast errors than the MAE.

$$MSE = \frac{1}{N} \sum_{t=1}^N e_t^2 \quad (6)$$

The square root of the MSE (typically labelled RMSE – root mean squared error) is also a commonly used forecast accuracy metric. The RMSE is sometimes preferred to the MSE as it is on the same scale as the data. It is the square root of the average of squared errors. The effect of each error on RMSE is proportional to the size of the squared error; thus larger errors have a disproportionately large effect on RMSE with the result that the RMSE is sensitive to outliers. This is a weakness and a strength in that large errors can be the most harmful from the perspective of forecast users. However, at the same time, the effects of unpredictable major shocks may be given undue weight.

Turning to relative measures of forecast performance, as stated above, a weakness of error metrics that use the same scale as the forecasted variables measures is that it is difficult to compare them across variables. A range of methods have been suggested to surmount this problem that involve normalising the values. A common approach to dealing with the scaling issue is to divide the forecast error by the forecast error obtained using another method of forecasting. Partially to overcome scaling issues, we evaluate the forecasts using Theil's U2 statistic. This statistic is commonly a measure of the naïve time series model (e.g. same as last year). The main rationale is based upon an expectation that forecasters should be able to outperform simple models.

$$U2 = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (F_t - A_t)^2}}{\sqrt{\frac{1}{N} \sum_{t=1}^N (A_t)^2}} \quad (7)$$

Where, following Theil's (1966) notation, F_t and A_t are pairs of predicted and realised *changes* of any given time series. In particular, in interpreting Theil's U2-statistic:

- $U = 1$ indicates the naïve forecasting method is as good as the forecasting technique being evaluated
- $U < 1$ indicates the forecasting technique being evaluated is better than the naïve forecasting method
- $U > 1$ indicates the forecasting technique being evaluated is worse than the naïve forecasting method.

Typically the naïve forecasting method used in calculating Theil's U_2 -statistic is the 'same as last period' forecasting strategy, in which the previous actual annual return is used as the real estate forecast for the subsequent annual period. Theil's U_2 statistic can be decomposed into three parts – bias, variance and covariance proportions (see Theil, 1966). The bias proportion measures how far the mean of the forecast is from the mean of the actual series. Hence, it measures the percentage of the forecast error that is caused by systematic forecasting error. The variance proportion represents the ability of the forecasts to reproduce the degree of variability in the actual outcomes. It measures the percentage of the forecast error that is caused by failure to capture the volatility of the actual. Finally, the covariance proportion measures the remaining unsystematic forecasting errors. The quality of the forecast is judged according to the relative sizes of the bias and variance proportions (the smaller the better) so that most of the forecast error should be in the covariance proportion.

Appendix 3. ARMA forecasts

ARMA forecasts	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Rental value growth (%)												
OFFICES	3.58	3.88	3.92	2.81	2.63	3.56	2.94	2.77	2.93	3.11	3.16	2.98
INDUSTRIAL	3.61	3.51	3.48	3.18	2.81	3.06	2.76	2.63	2.60	2.67	2.67	2.58
STANDARD SHOPS	5.46	5.42	5.30	4.85	3.94	4.72	4.20	3.93	3.89	4.08	3.93	3.85
SHOPPING CENTRES	6.06	5.80	6.04	4.83	4.56	4.92	4.51	4.27	4.24	4.14	4.17	3.85
RETAIL WAREHOUSE	6.27	6.11	5.93	5.62	4.90	4.97	4.90	4.71	4.52	4.44	4.37	4.26
ALL PROPERTY	4.56	4.57	4.58	3.85	3.58	4.18	3.64	3.43	3.44	3.48	3.44	3.27
City Offices	2.91	3.76	3.75	2.44	2.13	3.37	2.71	2.76	2.87	3.26	3.43	3.17
West End Offices	4.84	5.38	5.43	4.09	3.96	5.17	4.34	4.14	4.39	4.60	4.68	4.50
Capital value growth (%)												
OFFICES	2.99	3.42	2.70	2.16	2.64	2.34	1.95	1.98	2.33	2.63	2.72	2.97
INDUSTRIAL	3.73	4.01	3.46	2.67	2.90	2.47	2.27	2.10	2.33	2.67	2.74	3.26
STANDARD SHOPS	5.39	5.51	4.84	3.99	4.42	4.11	3.80	3.69	3.86	4.04	4.14	4.18
SHOPPING CENTRES	4.73	4.99	4.35	3.43	3.64	3.59	2.87	2.86	2.91	3.06	2.96	2.85
RETAIL WAREHOUSE	7.28	7.42	6.96	5.64	5.97	5.56	5.37	5.10	5.08	5.12	4.92	4.67
ALL PROPERTY	4.20	4.52	3.89	3.24	3.62	3.21	2.88	2.83	3.01	3.18	3.17	3.31
City Offices	2.90	3.57	2.41	1.83	2.43	2.47	2.01	2.19	2.61	2.97	3.20	3.40
West End Offices	4.86	5.59	4.70	4.11	4.86	4.76	4.21	4.47	4.80	5.17	5.25	5.40
Total return (%)												
OFFICES	9.96	10.34	9.63	9.05	9.48	9.08	8.68	8.69	8.99	9.24	9.30	9.14
INDUSTRIAL	12.57	12.74	12.09	11.14	11.40	10.86	10.62	10.39	10.60	10.92	10.91	10.86
STANDARD SHOPS	11.39	11.51	10.86	10.00	10.38	10.02	9.72	9.61	9.76	9.91	9.98	9.84
SHOPPING CENTRES	11.15	11.36	10.73	9.75	9.96	9.85	9.12	9.11	9.14	9.27	9.13	8.98
RETAIL WAREHOUSE	14.87	14.93	14.18	12.79	13.22	12.76	12.53	12.22	12.18	12.20	11.95	11.66
ALL PROPERTY	11.07	11.33	10.70	9.98	10.34	9.85	9.52	9.44	9.60	9.74	9.70	9.58
City Offices	9.48	10.08	8.94	8.32	8.86	8.81	8.34	8.50	8.88	9.19	9.37	9.07
West End Offices	11.56	12.24	11.34	10.72	11.42	11.20	10.63	10.85	11.11	11.41	11.43	11.21
Equivalent Yield												
OFFICES	8.16	8.01	8.12	8.26	7.99	7.94	8.00	7.91	7.82	7.70	7.75	7.61
INDUSTRIAL	8.96	7.49	8.89	9.61	9.48	9.40	9.36	9.35	9.22	8.98	8.82	8.69
STANDARD SHOPS	6.90	6.69	6.96	7.06	6.93	6.89	6.88	6.87	6.81	6.74	6.65	6.57
SHOPPING CENTRES	7.89	7.60	7.92	8.36	8.20	7.95	7.98	7.91	7.85	7.67	7.61	7.54
RETAIL WAREHOUSE	7.31	6.83	7.53	8.10	7.71	7.57	7.55	7.58	7.49	7.33	7.29	7.37
ALL PROPERTY	7.94	7.72	7.94	8.16	7.93	7.86	7.82	7.82	7.73	7.61	7.53	7.49
City Offices	7.49	7.43	7.50	7.56	7.43	7.39	7.38	7.38	7.30	7.26	7.18	7.16
West End Offices	7.88	7.64	7.80	7.91	7.63	7.50	7.39	7.29	7.16	6.89	6.78	6.86