David S. Bywaters*
Gareth Thomas**

The Rôle of Forwarding-Looking Price Expectations Data published by the Royal Institution of Chartered Surveyors (RICS) in explaining and forecasting U.K. house prices

Keywords: House prices, Expectations, Logistic, Forecasting

Abstract The objective of this paper is to make use of the forwarding-looking price expectations data published by the RICS to explain and forecast U.K. house prices, three-months ahead. The Nationwide index is used to test the performance of the RICS Survey based on the process of bounded rationality. An S-shaped logistic effect is shown to fit the data, assumed to be on account of a diffusion path of expectations from surveyors to other agents.

INTRODUCTION

Most empirical articles about the housing market, if they include expectations, do so only as a backward-looking conceptualisation of the general price level and not as the expected future values of the housing price (Meen, 2000). There is, however, a strong link between forward-looking expectations within the housing market and the state of the economy (Muellbauer and Murphy, 2008). The boom and bust cycles of rising and falling house prices have occurred on several occasions in the past (Garino and Sarno, 2004). One of the main factors which drove the latest recession was the expectations of falling house prices.

This paper outlines the theory on the formation and diffusion of house price expectations, building on the work of Mitchell *et al*. (2002 and 2005). The theory is then applied to the *RICS* Survey data on expectations of house prices over the next three months in conjunction with the actual house price index provided by the *Nationwide Building Society,* explained in the next section. The article ends on the empirical findings and the implications of the econometric study with the forecasting model. The paper is not concerned with any fundamental determination of house prices, but with expectations of future house prices.

THE DATA

The forward-looking price expectations are provided by the *RICS*, which is an independent, representative professional body, regulating property professionals and surveyors in the U.K. and other countries. These agents have expert knowledge of the housing market because they offer advice on mortgage valuations, and provide various surveying and auctioneering services to buyers and sellers. They are at the forefront of the market in providing survey information on a number of leading variables about housing such as price, sales and stock that can be formulated by the Institution into meaningful data. The surveys provide, in particular, details on the expected change in price, either in the form of 'ups' or 'downs' or 'the sames', over the next three months, which may indicate where the actual observations are heading in the future.

There are several sources of data on actual house prices in the U.K., in particular mortgage suppliers such as the *Halifax* and the *Nationwide* as well as the *Land Registry.* The *Financial Times House Price Index* is based on the information recorded at the *Registry;* it is smoothed and seasonally adjusted.

* University of Hertfordshire, Accounting, Finance and Economics, Business School, de Havilland Campus, Hatfield, Hertfordshire, AL10 9 AB, U.K
** University of Hertfordshire, Accounting, Finance and Economics, Business School, de Havilland Campus, Hatfield, Hertfordshire, AL10 9 AB, U.K.

These sources give rise to four sets of indices giving estimates of actual movements in residential prices. The *Nationwide* and the *Halifax* house prices are derived from standardised indexes based on a basket (or mix adjusted); thus they follow representative house prices, which are weighted over time, using data extracted from monthly mortgage information (or lending), meaning that the sample size varies from month to month. The *Land Registry* derives a price index based on information provided from completed sales, and therefore lags behind the mortgage lenders' observations. Furthermore, the method of calculation adopted means that the source can be influenced by a change in the mix, namely the proportion of different types of housing as well as the locations. The actual data set on monthly house prices from the *Nationwide*, not seasonally adjusted, is adopted in this empirical investigation, but it would be relatively easy to use the *Halifax index* instead.

THE THEORETICAL MODEL

The formation and spread of forward-looking price expectations within the housing market can be modelled on the basis of bounded rationality and the inter-dependence of agents. It can be described as a process of diffusion, implying slow adjustment because the whole process of buying and selling of houses takes considerable time, and for most market participants, happens infrequently. The majority of agents may seek to minimise costs of obtaining expectations by relying on forecasts formed by others (Carroll, 2003).

Those agents with the resources to form expectations from the available, imperfect information are likely to be the professional surveyors. In this case, they are a small number who possess the knowledge of the housing market and are part of an Institution that can expect their expectations to be published in the media and made available to other agents who are members.

The expectations are in the form of changes in the direction of price (P^e) between periods t and t+3 either as 'ups' ($P^u_{t,\,t+3}$) or 'downs' ($P^d_{t,\,t+3}$) over the following three months, although adjusted for the 'sames' ($P^s_{t,\,t+3}$) in the empirical study, that is $P^e_{t,\,t+3}$ based on previous values of $P^e_{t-i,\,t+3}$ that lie within the information set, I, at period t, namely

$$P^e_{t,\,t+3} = (P^e_{t-i,\,t+3}/I_t) \tag{1}$$

This group, however, is small relative to the majority of agents, which means that the distribution of expectations will initially occur slowly on publication to the media and the membership by the Institution, followed by a sudden increase in speed as the majority of estate agents, buyers and sellers convert to the change in predictions in an adaptive manner. This asserts that the followers of the surveyors will formulate expectations by the extent to which their prior predictions have now changed. For example, a simplified version could be

$$P^f_{t-j,\,t+3} = P^e_{t-j,\,t-3} + \lambda\,(P^e_{t-j,\,t+3} - P^e_{t-j,\,t-3}) \qquad (0 \leq \lambda \leq 1) \tag{2}$$

where λ equals the expectations coefficient of the 'followers', who translate the fraction of the extent that the current status has now changed for mortgage lenders, estate agents, buyers and sellers, adding or subtracting to the previous period's price growth to form the expected value in period t-j for t+3 which is absorbed into the current and lag values of the actual data produced by such institutions as the *Nationwide*. This is followed by a slow-down, as the remaining agents such as small developers fall into line, adapting to the majority's expectations of the future price.

This sequence of dispersion of the initial expectations formed by the surveyors and the degree of interdependence between agents implies a non-linear process of diffusion, captured by a dynamic logistic function, which according to Cramer (2003) is a natural counterpart of the regression method with an optimum lag to characterise the memory of the data sets, namely[1]:

1 The inner bracket of [3] can be interpreted as a constant plus a matrix, X, of regressors that represent a rectangular array of elements of order T, rows and K, columns, that is (T×K).

$$P^e_{t,\,t+3} = \left[1+\exp^{-(\beta_0 + \sum_{i=1}^{m} \alpha_i P^e_{t-i,t+3} + \sum_{j=0}^{m} \beta_j P^f_{t-j,t+3} + \varepsilon_t)} \right]^{-1} . \quad (3)$$

The expectations of the surveyors, $p^e_{t,\,t+3}$, therefore, will clearly be a function of previous predictions of itself, $p^e_{t-i,\,t+3}$ representing its historical formation (or memory) embodied in the information set, I_t of the current forecast at time 't'. This information will also be incorporated into the current, actual forward-looking growth rate as well as past data observations of mortgage lenders such as the *Nationwide* in the form of $p^f_{t-j,\,t+3}$ over time. The ε_t represents the disturbance term in the form of a remainder because the analysis denotes an average "screening device" of empirical evidence. In addition, the expectations will be related to the proportional coefficients, α_i, β_j and a constant, β_0.
The algebraic manipulation of expression [3] leads to the general statistical form of

$$LP^e_{t,t+3} = \beta_0 + \sum_{i=1}^{m} \alpha_i LP^e_{t-i,t+3} + \sum_{j=0}^{m} \beta_j P^f_{t-j,\,t+3} + \varepsilon_t \quad (4)$$

where $LP^e_{t,t+3}$ represents the logits: ln $(P^e_{t,t+3} / 1-P^e_{t,t+3})$. Expression [4] is the logistic format that captures the diffusion process represented by a Sigmoid (or S-shaped) curve. Once predictions are announced to the media by the *RICS* and made available to members, expectations will spread slowly because of the presence of uncertainty amongst numerous estate agents, but at an increasing rate until the point of inflexion is reached, after which the rate of change of conformity declines, affecting buyers' and sellers' expected prices as the process diffuses into the official data sets of the mortgage lenders like the *Nationwide*. The next section of this paper, therefore, takes this analysis of diffusion and attempts to explain the empirical regularities of expectations that lead to the actual figures of the housing market that give rise, in turn, to cycles of economic activity within the economy.

EMPIRICAL ESTIMATION OF THE DIFFUSION MODEL

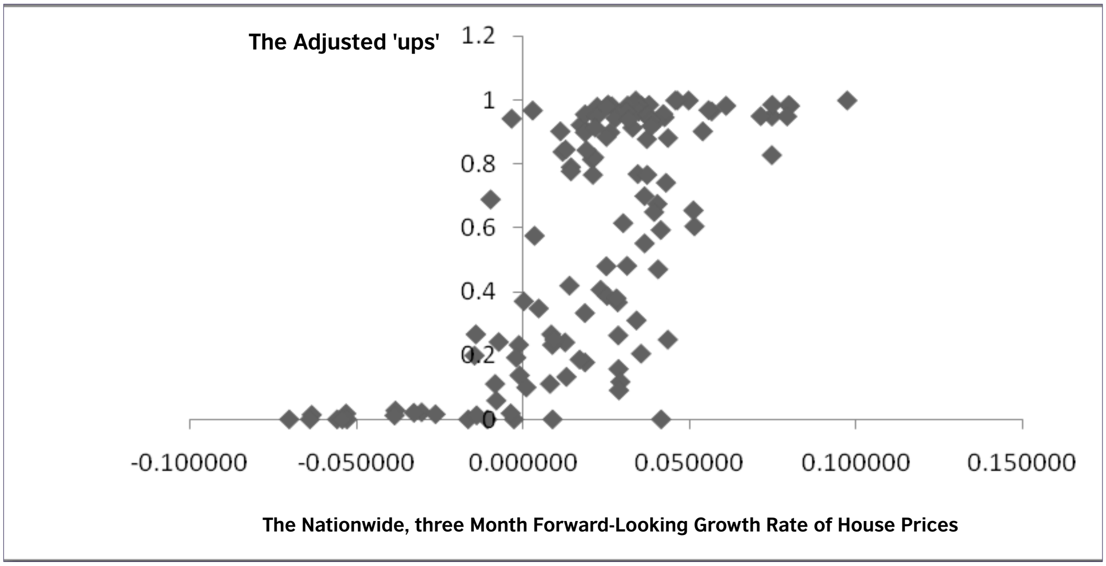
The logistic model of diffusion in [4] can be partly estimated using the *RICS* data on future price trends either in the form of the 'ups' ($P^u_{t,t+3}$) or 'downs' ($P^d_{t,t+3}$) to quantify $P^e_{t,t+3}$ within the non-linear process, but only if the 'sames' ($P^S_{t,t+3}$) are a constant proportion within the Survey, otherwise the end results will differ between 'ups' and 'downs'.

Given that the 'sames' have changed considerably over the time period, it is necessary to make an adjustment. The 'ups' and 'downs' can be normalised so they sum to one (or hundred), by calculating $Z=1-P^s_{t,t+3}$ to derive the adjusted variables $u_{t,t+3} = P^u_{t,t+3}/Z$ and $d_{t,t+3} = P^d_{t,t+3}/Z$.
Then, either normalised variable can be used in the empirical analysis to proxy $P^e_{t,t+3}$ within [4], because each gives perfectly symmetric results.
If the adjusted 'ups' are adopted and plotted against the three-month, forward-looking house prices derived from the *Nationwide* observations in a scatter diagram, as in Figure 3 overleaf, then it can be seen that the non-linear S-shaped path of diffusion manifests itself within the data sets in the form of a Sigmoid curve. This suggests that the theoretical discussion may well be along the correct lines and should be in the form of the logits within expression [4].
The difficulty, however, is that the expectation series of the adjusted 'ups' contain zero values, and therefore, cannot be logged in the form of [4]. To overcome this problem, the zeros have been put to 0.005 and the ones reduced to 0.995 so that the logistic variable can be derived and the statistical analysis continues along the theoretical lines discussed in the previous Section.
The statistical study employs the methodology of general-to-specific in order to derive a restricted form of expression [4] with the normalized 'ups' and the actual, forward-looking growth rates. This captures the 'memory' embodied in the data sets that drives the short-run dynamics of the lag

structure prevailing in the housing market. It arises from the interaction of demand and supply determining the future direction of the rate of growth of monthly house prices, specifically

$$Lu_{t,t+3} = \beta_0 + \sum_{i=1}^{24} \alpha_i Lu_{t-i,t+3} + \sum_{j=0}^{24} \beta_j P^f_{t-j,\, t+3} + \varepsilon_t \qquad (5)$$

Figure 1 S-Shaped Curve



where $Lu_{t,t+3}$ represents ($u_{t,t+3}$ /1-$u_{t,t+3}$) formed from the adjusted 'ups', $P^f_{t,t+3}$ denotes the three-month forward-looking growth rate, measured by the *Nationwide's* official, unadjusted house price series, P_t, in the form of ($\ln P_{t+4}$ - $\ln P_{t+1}$).

The general form of [5] was simplified by imposing restrictions and removing insignificant variables, using formal t-statistics, while lowering the standard error of the regression and improving performance against the selection criteria. The restricted model overleaf in Table 1 represents the forward-looking, diffusion model. The R^2 is the multiple correlation coefficient with the adjusted one denoted by $\bar{R}^2$, σ is the estimated standard error of the regression, DW is the Durbin-Watson statistic, AIC is the Akaike Information and SBC is the Schwartz Bayesian Criterions, LL represents the log-likelihood, RRS denotes the residual sum of squares and T is the number of observations used in the estimation.

Table 1 The Estimated Diffusion Model with Nationwide Data Set

DEPENDENT VARIABLE $\Delta Lu_{t,t+3}$ • TIME PERIOD FROM 1999 M:10 TO 2009 M:9			
Regressor	**Coefficient**	**Standard error**	**T-Ratio**
Costant	-0.19006	0.14760	-1.2877
$Lu_{t-5,t+3}$	-0.14490	0.031623	-4.5821
$Lu_{t-12,t+3}$	0.19323	0.03851	5.7427
$P^f_{t,t+3}$	10.6318	3.0228	3.5171
$P^f_{t-14,t+3}$	-25.9521	5.4485	4.2199
$P^f_{t-24,t+3}$	18.8206	4.4599	4.2199
D_2	3.0910	0.95660	3.2312
D_3	2.7173	0.96878	2.8049

Dummy variables: D_2, for 2004; M9 = -1, otherwise zero; D_3, for 2008: M12 = -1, otherwise zero.

R^2 = 0.45714, $\bar{R}^2$ = 0.42321, σ = 0.93988, DW =1.8521, AIC = -166.6930, SBC = -177.8430, LL = -158.6930, RRS = 98.9385, T = 120.

A:χ_{12} = 14.8243, B:χ_1 = 0.19905, C:χ_2 = 0.023373, D:χ_1 = 0.94464.

If the estimated equation is left as $Lu_{t,t+3}$, then the R^2 = 0.92378.

A: Lagrange multiplier test of residual serial correlation, compared with twelve months previous.
B: Ramsey's Reset test for functional form, using the square of the fitted values.
C: Lagrange multiplier test of normality, based on a test of skewness and kurtosis of residuals.
D: Lagrange multiplier test of heteroscedasticity, based on the regression of squared residuals on squared fitted values.

The diagnostic statistics in Table 1, A to C, suggest a statistically well-defined model. The restrictions increased the AIC and SBC statistics along with a fall in standard error of the equation and the residual sum of squares, when compared with the general model. The dynamic process embodied in the equation suggests a complex diffusion over a time span of twenty-four periods. Also, the summation of the values on the official coefficients is an overall positive one. The estimates of the dependent variable have a vital function to play in acting as an error correction process, stretching back twelve months. The analysis in the next section of the paper considers the forecasting properties of the model.

ECONOMETRIC ESTIMATION OF THE FORECASTING MODELS

The forecasting of future growth in house prices using the *RICS* Survey is causally quite different to the explanation of expectations in the previous section. Simple reversal of a regression equation in these circumstances is not possible. The forward-looking model, however, does indicate which dependent and which explanatory variables might be of particular interest. Given the empirical information, the dependent variable investigated was the log change, ($\ln P_{t+4} - \ln P_{t+1}$) for $P^f_{t,t+3}$ explained by previous price changes in growth and logistic variables, where the restricted form was estimated overleaf in Table 2.

To make use of this model for forecasting requires the construction of an intermediate model of $P^f_{t-1,t+3}$ clearly signalling the complexity underlying the *RICS* Surveys, which reflects the complicated dynamics of the housing market when it comes to expectation formulations that could be the initial trigger of booms and slumps. Thus, policy instruments should be focusing on manipulating forward-looking expectations that develop in the housing market *via* the information set.

The model in Table 2, in terms of $P^f_{t,t+3}$ can be compared with the Pesaran (1994) procedure for generating expectations. This estimation of the Pesaran format was over the time period of 1999: M4 to 2010: M1 led to a well-defined, backward-looking statistical representation of the form:

$$P^f_{t,t-3} = -\underset{(0.0017731)}{0.016367} + \underset{(0.015736)}{0.061555}\, u_{t,t-3} + \hat{V}_t,$$

$$\begin{aligned}\hat{V}_t = {} & \underset{(0.058945)}{0.10086}\hat{V}_{t-1} - \underset{(0.10000)}{0.61791}\hat{V}_{t-3} + \underset{(0.10156)}{0.53420}\hat{V}_{t-4} - \underset{(0.096773)}{0.3805}\hat{V}_{t-6} + \underset{(0.082777)}{0.22747}\hat{V}_{t-7} + \underset{(0.054187)}{0.13304}\hat{V}_{t-11} \\ & + \underset{(0.054278)}{0.13872}\hat{V}_{t-14} + \underset{(0.056993)}{0.16986}\hat{V}_{t-23} + U_t.\end{aligned} \tag{6}$$

$R^2 = 0.89533$, $\bar{R}^2 = 0.88748$, $\hat{\sigma} = 0.010388$, DW $=2.0103$, T $= 130$, $\hat{S} = 0.030969$, AIC $= 404.4637$, SBC $= 390.1260$,

A:$\chi^2_{12} = 16.2345$, B:$\chi^2_1 = 0.43669$, C:$\chi^2_2 = 2.4723$, D:$\chi^2_1 = 2.4283$.

Table 2 The Estimated Forecasting Model

DEPENDENT VARIABLE $P^f_{t,t+3}$ • TIME PERIOD FROM 2000 M:10 TO 2009 M:10			
Regressor	**Coefficient**	**Standard Error**	**T-Ratio**
Costant	0.0006726	0.0016712	0.40248
$Lu_{t-10,t+3}$	0.0019744	0.0007623	2.5871
$Lu_{t-14,t+3}$	-0.0015691	0.0006082	-2.5800
$Lu_{t-20,t+3}$	-0.0025680	0.0008909	-2.8825
$Lu_{t-21,t+3}$	0.0031341	0.0012170	2.5753
$Lu_{t-22,t+3}$	-0.0033007	0.0012081	-2.7322
$Lu_{t-23,t+3}$	0.0052421	0.0011823	4.4337
$Lu_{t-24,t+3}$	-0.0026064	0.0008559	-3.0453
$P^f_{t-1,t+3}$	1.2850	0.056202	22.8639
$P^f_{t-3,t+3}$	-0.92930	0.11365	-8.1766
$P^f_{t-4,t+3}$	0.85450	0.11366	7.5181
$P^f_{t-6,t+3}$	-0.82729	0.11804	-7.0083
$P^f_{t-7,t+3}$	0.74494	0.12164	6.1242
$P^f_{t-9,t+3}$	-0.34255	0.086768	-3.9479
$P^f_{t-11,t+3}$	0.25862	0.099226	2.6064
$P^f_{t-12,t+3}$	-0.27439	0.089240	-3.0748
$P^f_{t-17,t+3}$	0.13897	0.056942	2.4406
$P^f_{t-22,t+3}$	0.18639	0.078668	2.3693
$P^f_{t-24,t+3}$	-0.015496	0.078388	-2.0000

$R^2 = 0.94189$, $\bar{R}^2 = 0.93014$, $\sigma = 0.0086278$, DW =2.05484, AIC = 351.5020, SBC = 326.0218,
LL = 370.5020, RRS = 0.0066251, T = 109.

A:$\chi_{12} = 9.0114$, B:$\chi_1 = 0.12997$, C:$\chi_2 = 3.6514$, D:$\chi_1 = 0.34384$.

$P^f_{t,t-3}$ is the backward-looking price growth, that is the log change in ($\ln P_t - \ln P_{t-3}$) with $u_{t,t-3}$ the past, adjusted 'ups'. The inclusion of the past residuals $\hat{V}_t$ was tested as instruments using the Hausman (1978) test, which could not reject the null-hypothesis of no error-correcting forecasts. Thus, it is necessary to include the past residuals as instruments for the self-correcting mechanism of forecasting errors. Furthermore, the diagnostic tests clearly indicate a well-defined statistical format.

The expected growth in house prices over the next three months using the adjusted ‘ups’, $u_{t,t+3}$ is utilized in the following equation to create expectations from equation [6] to aid comparison:

$$P^f_{t,t+3} = - 0.016367 + 0.061555\, u_{t,t+3} + 1.008667V_t - 0.61791V_{t-2} + 0.5342V_{t-3} - 0.38085V_{t-5} + 0.22747V_{t-6} + 0.13304V_{t-10} - 0.13872V_{t-13} + 0.16969V_{t-22}. \quad (7)$$

These are compared with the fitted-values of $P^f_{t,t+3}$ derived from the equation in Table 2.

According to Hansen (2005), in addition to the use of model selection statistics, the root mean squared forecast error (RMSFE) can be relevant in comparing models. In the case of equation [7], the RMSFE led to a value of 0.021323. Using the fitted-values of $P^f_{t,t+3}$ from Table 2, the RMSFE was 0.0069912. The logistic function is shown here to have a lower RMSFE, compared with the Pesaran method.

Given the complexity that has developed from the Hendry methodology, it was decided to experiment with the lag length of the dependent variable. For example, when the lag length was put to twelve months into the future, that is log change of $(\ln P_{t+12} - \ln P_{t+1}) = P^f_{t,t+12}$, Table 3 overleaf was derived from the estimation. This suggests that the Survey data set contains more than just three months of information, but could well contain a yearly sequence of events. Given the analysis and the empirical evidence found in this paper, the missing link for policy is to influence forward-looking house price expectations *via* the fundamentals of the information set. This should be the next step in the research, to expose the rudiments of the information set.

Table 3 A Forecasting Model, Twelve Months Ahead

DEPENDENT VARIABILE $\Delta P^f_{t,t+12}=P^f_{t,t+12}-P^f_{t-1,t+12}$ • TIME PERIOD FROM 2000 M:10 TO 2009 M:1			
Regressor	**Coefficient**	**Standard Error**	**T-Ratio**
Costant	-0.0010070	0.0020347	-0.49489
$Lu_{t-1,t+3}$	-0.0031463	0.0008732	-3.6031
$Lu_{t-2,t+3}$	0.0047452	0.0009147	5.1877
$Lu_{t-6,t+3}$	-0.0016784	0.0006789	-2.4720
$Lu_{t-12,t+3}$	-0.0035910	0.0008312	-4.3202
$Lu_{t-13,t+3}$	0.0026312	0.0009408	2.7969
$Lu_{t-15,t+3}$	0.0019840	0.0006026	3.2922
$Lu_{t-23,t+3}$	0.0012278	0.0004519	2.7169
$P^f_{t-1,t+12}$	0.15290	0.021685	7.0508
$P^f_{t-6,t+12}$	-0.17978	0.045207	-3.9769
$P^f_{t-10,t+12}$	-0.21174	0.098920	-2.1405
$P^f_{t-11,t+12}$	-0.45308	0.12454	-3.6380
$P^f_{t-12,t+12}$	0.75898	0.10445	7.2662
$P^f_{t-22,t+12}$	-0.59035	0.10937	-5.3979
$P^f_{t-23,t+12}$	0.49213	0.10297	4.7796
D_4	0.037020	0.0091427	4.0492

Dummy variable: D_4, for 2007: M12 = -1, otherwise zero.

R^2 = 0.76156, $\bar{R}^2$ = 0.71898, σ = 0.0085057, DW = 1.8118, AIC = 357.5253, SBC = 306.6840,
LL = 343.5253, RRS = 0.008057, T = 100.
A:χ_{12} = 18.2972, B:χ_1 = 0.045388, C:χ_2 = 1.0760, D:χ_1 = 0.26786.

CONCLUSIONS/SUMMARY

The paper has focused on the process of the formation of expectations of house prices underlying the *RICS* Survey. The study suggests that there is a diffusion process, which is captured by the logistic model. The empirical results from the estimation provide evidence supporting the logistic specification. This is in line with models of bounded rationality, where decision-making is uncertain, self-fulfilling, complex, and costly. The majority of agents follow the few, namely the Chartered Surveyors.

The policy implications discussion in the previous section suggests that the expectations of house prices could well be playing a pivotal part in the cycles of economic activity, and, therefore, the forecasting model could be used as an early warning system of forthcoming 'swings and roundabouts' within the economy.

References

Carroll, C.D. (2003) Macroeconomic Expectations of Households and Professional Forecasters, *Quarterly Journal of Economics*, Vol.118, February, pp 269-298.

Cramer, J.S. (2003) *Logit Models from Economics and Other Fields*, Cambridge University Press.

Garino, G. and Sarno, L. (2004) Speculative Bubbles in U.K. House Prices: Some New Evidence, *Southern Economic Journal*, Vol.70, Issue 4, pp 777-795, April.

Hansen, B.E. (2005) Challenges for Econometric Model Selection, *Econometric Theory*, 21, pp 60-68.

Hausman, J. (1978) Specification Tests in Econometrics, *Econometrica*, 46, pp 1251-1271.

Meen, G. (2000) Housing Cycles and Efficiency, *Scottish Journal of Political Economy*, Vol.47, (2), pp 114-140.

Mitchell, J., Smith, R. and Weale, M. (2002) Quantification of Qualitative Firm-Level Survey Data, *Economic Journal*, Vol.112, pp 117-135.

Mitchell, J., Smith, R. and Weale, M. (2005) forecasting Manufacturing Output Growth, using Firm-Level Survey Data, *The Manchester School*, Vol.73, No.4, pp 479-499.

Muellbauer, J. and Murphy, A. (2008) Housing Markets and the Economy: the assessment, *Oxford Review of Economic Policy*, Vol. 24, (1), pp 1-33.

Pesaran, M.H. (1984) Expectations Formations and Macroeconometric Modelling, in Malgrange and Muet (eds.) *Contemporary Macroeconometric Modelling*, Basil Blackwell, pp 27-55.