

## **Output Expectations and Forecasting of U.K. Manufacturing**

### **Abstract**

This paper analyses both quarterly data from the Confederation of British Industry (C.B.I.) Survey on respondents' expectations of recent and forthcoming manufacturing output, and monthly Office of National Statistics (O.N.S.) figures on actual manufacturing output within the U.K. Quarterly output expectations of the C.B.I. manufacturers are explained from the monthly O.N.S. observations using a bounded rationality approach. The logistic formulation models the diffusion process across respondents. There is a backward-looking C.B.I. Survey perspective, explained by past O.N.S observations, and a forward-looking perspective, explained from future O.N.S. statistics. Also, the forecasting of monthly manufacturing output from earlier values along with the quarterly C.B.I. Survey information is examined, and tested against the alternative Pesaran/Thomas method.

The study provides econometric evidence for the validity of the logistic model, and shows that bounded rationality can explain the formation of predictions among business managers in the U.K. manufacturing sector. The emerging consensus from the literature, supported by this paper, is that the logistic format is a superior 'approximation' to the true data generating process compared with the earlier standard Anderson/Pesaran/Thomas approach. An adjustment to the Survey is used, which achieves perfect symmetry with 'up' and 'down' versions of the data. The benefits of this adjustment are tested in the forecasting Section.

**Key Words:** Manufacturing, Output, Expectations, Forecasting, Logistic, Bounded Rationality

**JEL Classification:** C10, D21

**By:** David Bywaters and D.Gareth Thomas\*

**E-mail Address:** [d.bywaters@herts.ac.uk](mailto:d.bywaters@herts.ac.uk)

\* The authors would like to thank Dr Tsitsianis and the anonymous referee for their constructive comments made on the content of the paper along with Matt Thomas for his help with the computing.

## **Introduction**

Expectations play a central rôle in macroeconomics (Evans and Honkapohja, 2001) because in an uncertain world, decision-making by the business community is based on their expectations concerning the future behaviour of key economic variables. As indicated by Branch (2004), the rational hypothesis has failed to explain the survey data on inflationary expectations, because of the knowledge required by economic actors on the structure of the economy. Even an ‘up-dating rule-of-thumb’ in the spirit of rational expectations can require substantial resources and knowledge of the economy, which is ‘on par’ with a skilled, experienced econometrician. It is this lack of ability and information for the formation of rational expectations that leads agents to adopt bounded rationality to explain the process of formation and diffusion of expectations (Sargent, 1993), which is the approach applied in this paper to output expectations.

This article presents a model of output expectation formation applied to information derived from the Confederation of British Industry (C.B.I.) Industrial Trends Survey. This data is explained in more detail in the next Section, and the following one is concerned with relevant methodology and functional form. There are a range of approaches to modelling expectations of output in the literature, and three papers have been published quite recently applying them specifically to C.B.I. figures (Mitchell et al, 2002; Driver and Urga, 2004 and Mitchell et al, 2005). Mitchell et al. (2005) provide an overview of the alternative depictions of output expectation formation. Driver and Urga (2004) deal only with backward-looking models, and therefore, exclude forecasting. This paper maintains the monthly disaggregation of the Government output statistics, which are used to explain the C.B.I. Survey expectations; the other recent articles use quarterly data. The articles by Mitchell et al (2002 and 2005) disaggregate by company respondent to the C.B.I. Survey, but unfortunately the ability to disaggregate in that way ceased in 1999, so there is no current forecasting application.

### ***The C.B.I. Data***

The Survey draws information from managers of at least one thousand, seven hundred manufacturing companies, who are asked the following questions,

*"Excluding seasonal variations, what has been the trend over the past four months and what are the expected trends for the next four months...?"*

This is the company Survey data on backward- and forward-looking expectations used in this article. According to Mitchell et al. (2005), page 488, the majority of researchers in this field:

*"treat the difference between the four-month period referred to in the survey and the quarterly frequency of our official data as being unimportant".*

The approach used in this study maintains the timing distinction between the quarterly survey and the monthly output figures, by comparing quarterly survey data with the monthly, non-seasonally adjusted government statistics. The respondents give an indication of output trends at the beginning of the month ( $m_0$ ), covering the past four months ( $m_0$  to  $m_{-4}$ ), which are published each January, April, July and October, which are ( $m+1$ ). By contrast, Official manufacturing output data to the end of the month, ( $m-1$ ), are made available initially two months later ( $m+1$ ), although subject to later correction (C.B.I., 1983).

The C.B.I. survey started in 1958 on the basis of three questionnaires per year until 1972 when it became quarterly. We have converted the three figures each year before 1972 into four by interpolation. The sources of the observations are the C.B.I. and the Datastream database. The empirical analysis in subsequent sections of this article suggests that these 1959 to 1971 observations are reliable, with no evidence of heteroskedasticity in the regressions.

Respondents are asked to identify a trend over the past and future four months without seasonal variation, but they may instinctively contrast their current response with that of a year earlier. The Survey figures may be the result of a smoothing process (Cunningham, 1997), which could lead to serial correlation patterns in the regression analysis.

The form in which the study has used the monthly Official manufacturing output is as annual logarithmic changes. This was achieved by taking the logarithm of the Official estimate of manufacturing output each month, and subtracting the logarithm of output 12 months previous. This reflects the ‘overlapping’ yearly sequence of events, which, we argue, is the basis for managers' responses.

### **Output Expectations and Functional Form: A Logistic Model of Business Expectations**

Mitchell et al (2005) explain that there are two, fundamentally different approaches, to quantifying qualitative information such as the C.B.I. Survey data; these are the regression method, in this output expectations case the Pesaran/Thomas procedure, and the probability method, in this instance, the logistic (or logit) model. Their abilities to explain the Survey expectations cannot be directly compared, because the dependent quantitative variables are different, but their abilities to forecast the Government output data can be compared using Root Mean Squared Error (RMSE), the Akaike Information Criteria (AIC) and the Schwartz Bayesian Criteria (SBC).

Pesaran (1984) developed the Linear Anderson Method to explain Survey Expectations of prices. It was further adapted in Thomas (1995) to explain Output Survey Expectations. Driver and Urga (2004) use the Linear Anderson and, for some series, the Pesaran method. The logistic (or logit) model is an example of a probability method, related to the approach of Carlson and Parkin (1975). As Cramer (1991) explains, there can be a number of reasons for using a logistic formulation: one is a probability threshold explanation, described and used in Mitchell et al (2002); another is a diffusion explanation, where the results of a Survey are a ‘snapshot’ of a dynamic process of expectation formation by companies. A third explanation could be that the logistic function facilitates the ‘compression’ of a variable over an infinite range to one over a restricted scale.

To expand on the diffusion explanation, suppose that the business environment has been benign, most firms have expected 'up', and there has been positive growth in aggregate manufacturing output in the economy. Then some leading respondents of the 'pack' begin to expect recession, and change to 'down', reducing the proportion of 'ups'. The expected change in direction, up or down, denoted by  $E_k$ , should later be reflected in the Official data, represented by  $x$ , denoting actual manufacturing output growth. The expected change in direction,  $E_k$ , is related to the output growth rate by a proportional coefficient  $\beta_1$ , and a constant,  $\beta_0$ , to a linear approximation. The 'leaders of the pack' are later followed by more of 'the pack', and finally by the 'stragglers'.

The formation of the expectations is based on bounded rationality, meaning that respondents are rational in pursuit of self-interest, but are constrained by the complexity and cost of processing information. Agents are "intendedly rational, but only limitedly so" (Simon, 1961). The intended rationality of economic agents is constrained by the existence of imperfect, locally available asymmetrical information, and by the complexity and expense attached to processing it into meaningful expectations about the changing economic environment within limited time. There are limits on the skills of individuals to process information without error, and even with the advent of computers, there is finite computational capacity to store and process available information (Simon, 1986). Interdependent agents may seek to minimise the costs of expectation formation by relying on the expectations formed by others, such as naive agents with high information related costs, or with little or no resources available for gathering and processing data. Those entrepreneurs with the skills and resources needed to form expectations from the available imperfect information are likely to become trendsetters (or 'alpha') within the 'pack' of respondents. The latecomers are particularly slow in adopting the market expectations, because they face high costs in adapting to change. This sequence of dispersion and degree of interdependence implies a non-linear process which can

be captured by the logistic function. The alpha trend-setter and latecomer groups are small relative to the majority of entrepreneurs, which means that the dissemination of business expectations will be slow initially, with a sudden acceleration as the majority of entrepreneurs 'convert' to the change in predictions of  $E_k$ , which is followed by a slow-down as the stragglers reluctantly fall into line.

The proportions of 'ups' (U) or 'downs' (D) of the Survey can be used to measure this non-linear process, but unless the proportion of the 'sames' (S) is constant in the Survey, the results will be different. Other investigators have not always appreciated that the unadjusted 'ups' give quite different results to the 'downs' if the 'sames' alter. Historically the proportion of 'sames' has changed considerably at times, so some adjustment is clearly called for. The 'ups' and the 'downs' can be normalised so that they sum to 1 (or 100), by calculating  $Z=1-S$  to derive the adjusted variables of  $u=U/Z$  and  $d=D/Z$ <sup>1</sup>. Then either normalised variable can be used in the empirical analysis because they give perfectly symmetric results for  $E_k$ .

The non-linear diffusion process can be represented by the logistic function denoted as follows:

$$E_k = (1 + \exp^{-(\beta_0 + \beta_1 X)})^{-1} \quad [1]$$

where, as already indicated in the discussion,

- $E_k$  = the proportion of observations on the expected change in direction: up or down, in output of the k-th firm for the period;
- $X$  = the growth rates of backward- and forward-looking output derived from the Official time-series data;
- $\beta_0$  = a constant that fixes the curve on the time scale;
- $\beta_1$  = the rate of growth coefficient.

---

<sup>1</sup> In fact, the normalized 'ups' were tested against the unadjusted 'rises' using the Thomas (1995) procedure. This method was chosen because of the simplicity, allowing a straightforward comparison using the AIC and SBC statistics.

Defining  $z_t = \beta_0 + \beta_1 X_t$ , where  $E_k$  varies in an 'S' shape over time, a sigmoid path (Cramer, 1991), and re-writing [1] to ease the manipulation:

$$E_k = (1 + \exp^{-z_t})^{-1}. \quad [2]$$

It can be shown that  $0 < E_k < 1$  as  $-\infty < z_t < +\infty$ . This is a necessary condition for  $E_k$  to be a probability density function. It is simple to show by algebraic manipulation that [2] implies

$1 - E_k = (1 + \exp^{+z_t})^{-1}$ , with the so-called odds ratio equal to<sup>2</sup>:

$$\frac{E_k}{1 - E_k} = \frac{(1 + \exp^{+z_t})}{(1 + \exp^{-z_t})} = e^{z_t}. \quad [3]$$

Taking the natural log of each side of [3] to derive the logits:

$$\ln\left(\frac{E_k}{1 - E_k}\right) = \ln(\exp^{z_t}) = z_t \quad [4]$$

and substituting to obtain:

$$\ln\left(\frac{E_k}{1 - E_k}\right) = \beta_0 + \beta_1 X. \quad [5]$$

According to Cramer (1991), this logistic model of [5] can give empirically identical results to the probit model, given appropriate choice of the parameters. The logistic is used by Mitchell et al (2005 and 2002) because it captures the required characteristics that link the output expectation ( $E_k$ ) with “the growth rate...” in the official measure of output growth,  $X$ . In their words on page 133:

*“As the growth rate rises from a low negative value, one would expect both the proportions reporting no change and those reporting a rise in their volume of output to increase. Beyond a particular point the steady state proportion reporting no change should be expected to start falling again, while the proportion reporting a rise should continue to increase. The proportion reporting a fall should be expected to be declining function of the growth rate.”*

This statement clearly indicates that the logistic format captures the dynamic properties of diffusion as an “approximating model” of the true, evolving data-generating process (Hansen,

---

<sup>2</sup> The Taylor expansion in conjunction with the odds format justifies this approximation leading to [5] with a remainder (or a form of error-term), which is the goal of model selection in the approach of Hansen (2005). See page 10 of Cramer (1991) for an exposition of the mathematics.

2005)<sup>3</sup>. The quote above from Mitchell et al (2002) is immediately followed by: *We can see that the logit models have the expected properties.*

### **Empirical estimation of the backward-looking model of business expectation diffusion**

A logistic model of expectation diffusion is estimated here using the C.B.I. data on past output trends. In this econometric model,  $E_k$ , is adjusted to  $u_{q,b}$ , the proportion of firms that at the time of the Survey ( $q$ ) had a ‘rise’ in their output over the previous four months ( $b$ ), adjusted for ‘sames’. Equation [5] then becomes:

$$\ln\left(\frac{u_{q,b}}{1-u_{q,b}}\right) = \beta_0 + \beta_1 X. \quad [6]$$

The dependent variable is the adjusted proportion of managers in the Survey who think output rose in the previous four months (backward-looking). The independent variable is the Official change in manufacturing output. The econometric model of equation [6] indicates the extent to which a logistic model fits the observations. The estimated parameters of  $\beta_0$  and  $\beta_1$  are expected to be significant and positive, because an increase in the percentage of managers who have expected “ups” in output should be correlated with a rise in the rate of growth of manufacturing output. The inclusion of the constant term in the regression corrects any average bias of the official data (Cunningham, 1997).

The statistical study employs the Hendry methodology of general-to-specific estimation on the more general form of equation [6] to discover the dynamics and lag structure, as follows:

$$lu_{q,b} = \beta_0 + \sum_{i=1}^5 \alpha_i lu_{(q-i,b)} + \sum_{j=0}^{11} \beta_j X_{m-j} + c_1 d_1 + \varepsilon_q, \quad [7]$$

where  $lu_{(q-i,b)}$  equals  $\ln\left(\frac{u_{q-i,b}}{1-u_{q-i,b}}\right)$ ,  $X_{m-j}$  denotes the annual growth in the official, non-seasonally adjusted output series based on month  $m$ , for earlier months  $(m-j)$  and  $d_1$  is one

---

<sup>3</sup> The first glimpse of the logistic model applied to the C.B.I. Survey data in the public domain was in a footnote in Cunningham (1997).

dummy variable<sup>4</sup>. The general form was simplified by imposing restrictions and removing insignificant variables, using formal ‘t’ tests, while lowering the standard error of the regression, and improving performance against the selection criteria. The following restricted model represents the backward-looking, bounded model:

$$\begin{aligned}
 lu_{q,b} = & 0.022224 + 0.71289lu_{(q-1,b)} + 0.22264 lu_{(q-4,b)} - 0.12114 lu_{(q-5,b)} + 2.3458 X_m + 4.1839 X_{m-1} \\
 & (0.025122) \quad (0.053786) \quad (0.068278) \quad (0.060161) \quad (0.63537) \quad (0.70134) \\
 & + 4.3235X_{m-2} - 1.5105X_{m-4} - 1.8239X_{m-5} - 1.9933X_{m-6} - 2.3070X_{m-7} + 1.2432d_1 + \varepsilon_q, \\
 & (0.77554) \quad (0.84940) \quad (0.8239) \quad (0.67353) \quad (0.73985) \quad (0.21368)
 \end{aligned}
 \tag{8}$$

$R^2 = 0.87958$   $\bar{R}^2 = 0.87155$   $\hat{\sigma} = 0.29906$   $DW = 2.0634$   $T = 177$ ,  $\hat{S} = 0.83443$   $AIC = -43.2782$   $SBC = -62.3351$ ,  
 $A: \chi^2_4 = 3.1570$   $B: \chi^2_1 = 1.0056$   $C: \chi^2_2 = 0.93191$   $D: \chi^2_1 = 0.0040099$ ,

where bracketed figures are standard errors.  $R^2$  is the multiple correlation coefficient with the adjusted one denoted by  $\bar{R}^2$ ,  $DW$  is the Durbin-Watson statistic,  $\hat{\sigma}$  is the estimated standard error of the regression,  $T$  is the number of observations used for estimation and  $\hat{S}$  is the standard deviation of the dependent variable,  $AIC$  is the Akaike Information Criterion and  $SBC$  is the Schwartz Bayesian Criterion<sup>5</sup>.

The diagnostic statistics in expression [8] are shown above<sup>6</sup> and suggest a well-defined model statistically. The restrictions increased the AIC and the SBC statistics when compared with the general model<sup>7</sup>. The dynamic process embodied in the equation suggests a complex diffusion over a time span of seven periods, where the first three values are compared with the following four negative lags. Also, lagged estimates of the dependent variable have an important rôle to play in acting as a error-adjustment process, lagged by one and four quarters, which imply the need for a lag of 5 quarters too, using operator algebra.

<sup>4</sup> The dummy variable refers to important events which led to short-run, ‘unusual’ changes at 1998: Q4 (-1) and 2004: Q4 (+1).

<sup>5</sup> Coefficient stability was checked for compared with the shorter period since 1972, but the results are not reported in detail.

<sup>6</sup> A: Lagrange Multiplier test of the residual serial correlation, B: Ramsey’s Reset test using the square of the fitted values, C: Based on a test of Skewness and Kurtosis of residuals, D: Based on the regression of the squared residual on squared fitted values.

<sup>7</sup> In this study, the AIC and SBC statistics are based on the maximum likelihood statistics, and therefore, it is the maximum values that are of interest (or reduced negative numbers).

## Empirical estimation of the forward-looking model of business expectation diffusion

The C.B.I. Survey forward-looking data requires similar alignment with the timing of the Official manufacturing figures as the backward-looking observations did, where in this case managers are expected to identify future output trends. We assume the ‘next’ four-months are compared with the same four-month period in the previous year, taking annual logarithmic differences. There is also a need for the normalization adjustment using the ‘same’ responses as for the backward analysis. The representation of bounded expectations now takes the general form, after some preliminary elimination of insignificant output variables:

$$lu_{q,f} = \beta_0 + \sum_{i=1}^{16} \alpha_i lu_{(q-i,f)} + \sum_{j=7}^{19} \beta_j X_{m+j} + c_1 d_1 + \varepsilon_q, \quad [9]$$

where  $lu_{(q-i,f)}$  equals  $\ln\left(\frac{u_{q-i,f}}{1-u_{q-i,f}}\right)$  the future adjusted “ups”,  $X_{m+j}$  denotes the forward growth in the official unadjusted output series and  $d_1$  is one dummy variable<sup>8</sup>. A more specific model for the period 1963: Q4 to 2003: Q3 was then derived by imposing restrictions to eliminate statistically insignificant variables:

$$\begin{aligned} \Delta lu_{q,f} = & 0.069693 - 0.32351 lu_{(q-1,f)} + 0.16166 lu_{(q-4,f)} - 0.21634 lu_{(q-5,f)} + 0.17889 lu_{(q-6,f)} \\ & (0.045321) \quad (0.045522) \quad (0.059264) \quad (0.063813) \quad (0.055659) \\ & + 3.3253 X_{m+8} + 2.7626 X_{m+9} + 3.5413 X_{m+10} - 3.3723 X_{m+13} + 1.3795 d_{1t} + \varepsilon_q, \\ & (0.92303) \quad (0.89646) \quad (1.0325) \quad (0.89795) \quad (0.16988) \end{aligned} \quad [10]$$

$$R^2 = 0.62152, \bar{R}^2 = 0.59881, \hat{\sigma} = 0.39067, DW = 2.1700, T = 160, \hat{S} = 0.61678, AIC = -81.4830, SBC = -96.8589.$$

$$A: \chi_4^2 = 3.3788, B: \chi_1^2 = 0.46967, C: \chi_2^2 = 2.2022, D: \chi_1^2 = 0.20101.$$

The diffusion of expectations is explained here by the future output from months eight to thirteen, augmented by the error adjustment mechanism from previous expectations. The estimated parameters in the equation are statistically significant and the model explains approximately sixty per cent of the change in the quarterly diffusion process of expectations of manufacturing output. This means the predictions of the model as a rise in the percentage of

<sup>8</sup> In the form of 1973: Q2 (1), 1974: Q1 (-2) and Q2 (1), accounting for the Barber Boom, the three-day week in manufacturing industry caused by the miners’ strike and the effect of the General Election of that year.

managers predicting growth in manufacturing output in the next four months is correlated with a subsequent overall increase in the rate of growth of output in the O.N.S. data. This empirical representation is well defined statistically on inspection of the diagnostic statistics in [10] with the AIC and the SBC improving compared with the (unreported) general model.

These estimation results, in particular the significant lagged expectation variables, provide evidence that the expectations formation and diffusion among managers is really an error correction process. A further step in the analysis is to impose a precise error adjustment formula in order to capture these short-run dynamics. Using the data through from 1960 to 2004, the co-integrating regression for these variables of interest is:

$$\begin{aligned}
 lu_{q,f} &= 0.62569 + 11.9369X_m + U_q, \\
 &\quad (0.055443) \quad (1.1789) \\
 R^2 &= 0.36546, \quad DW = 0.84596.
 \end{aligned}
 \tag{11}$$

The Durbin-Watson statistic is reported, as it is sufficiently above zero to suggest that the residuals from the regression may indeed be stationary for co-integration. The  $R^2$  is relatively low, because the estimation is a function of changes. Furthermore, to complement the conclusion for stationarity in the residuals,  $U_q$ , from [11], Dickey-Fuller (DF) and the Augmented Dickey-Fuller (ADF) statistics were calculated and tests for co-integration performed. Concluding that there is a long-run relationship existing between the variables, the error-correction model is estimated using the residuals from expression [11] over the period 1960:Q2 to 2004:Q4:

$$\begin{aligned}
 \Delta lu_{q,f} &= 3.3213QX_m - 0.21140U_{q-1} - 0.85609d_1 + 2.6469d_2 + 1.2296d_3 + 1.4247d_4 + \varepsilon_q, \\
 &\quad (0.83628) \quad (0.047697) \quad (0.29921) \quad (0.32321) \quad (0.29861) \quad (0.29938) \\
 R^2 &= 0.5307, \quad \bar{R}^2 = 0.51958, \quad \hat{\sigma} = 0.42171, \quad DW = 1.9887, \quad T = 179, \quad \hat{S} = 0.60842,
 \end{aligned}
 \tag{12}$$

where  $d_i, i=1,2,3,4$  are four dummy variables<sup>9</sup>. The quarterly change in  $X_m$ ,  $X_m - X_{m-3}$ , is denoted by  $QX_m$ . All the diagnostic tests suggest a well-defined statistical model. The sample size was reduced, so that stability could be ‘checked’, by starting at 1972: Q1, suggested by Driver and Urga (2004) as this is the point when the Survey became an actual quarterly response, which led to:

$$\Delta u_{q,f} = 3.1737QX_m - 0.20393U_{q-1} + 2.6721d_2 + 1.2288d_3 + 1.4301d_4 + \varepsilon_q.$$

(1.0575)      (0.057140)      (0.35076)      (0.31456)      (0.31570)

[13]

$$R^2 = 0.56311, \quad \bar{R}^2 = 0.54935, \quad \hat{\sigma} = 0.44403, \quad DW = 2.0616, \quad T = 132, \quad \hat{S} = 0.66144.$$

The coefficients are stable when comparing [12] and [13]. These error correction models are much more constrained than the specific forward-looking model of [10], almost to the point of over simplification. Expression [10] is statistically the better model, as shown by the lower standard error of the regression, but [12] and [13] add another perspective on the process of expectation formation.

### **Econometric estimation of the forecasting model of future O.N.S. output**

The forecasting of future output using the Survey is causally quite different to the explanation of expectations in the previous Sections. Simple reversal (or inversion) 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 of the forward-looking model, the dependent variable investigated was the change in  $X_{m+8}$ , explained by the relevant previous output changes, and

---

<sup>9</sup> They refer to important historical events which led to short-run, ‘unusual’ changes:  $d_1 = 1966: Q4 (-1)$  and  $1969: Q4 (1)$ ,  $d_2 = 1974: Q1(-1)$  and  $1974: Q2 (1)$ ,  $d_3 = 1979: Q2 (1)$  and  $1979: Q3 (-1)$ ,  $d_4 = 1980: Q3 (-1)$  and  $1983: Q2 (1)$ . The events of 1979, 1980 are the ‘deep-seated’ recession effects and the Falklands war created uncertainty in 1983.

logistic variables from available Survey data, giving rise to the following restricted empirical formulation in Table [1]:

<b>Table [1]:</b> Dependent variable is $\Delta X_{m+8} = X_{m+8} - X_{m+5} + X_{m-4} - X_{m-7}$ , 1962:Q3 2004:Q3		
$Lu_{q-i, f}$	Coefficient	Standard Deviation
(-0)	0.0070994	0.0036385
(-1)	-0.0059482	0.0036084
(-5)	0.0097090	0.0037925
(-6)	-0.0156160	0.0038233
(-7)	0.0174550	0.0040754
(-8)	-0.0129460	0.0038257
$X_{m+j}$ (+7)	0.1679400	0.0596550
(+6)	0.32088	0.055398
(+5)	-0.40679	0.066924
(+2)	-0.11491	0.063138
(+1)	-0.15765	0.062644
$X_{m-j}$ (-0)	0.15862	0.062790
(-4)	0.37200	0.065157
(-6)	0.27799	0.060513
(-7)	-0.55133	0.067936
(-11)	-0.18076	0.066333
(-12)	0.12655	0.060178
(-13)	-0.089129	0.060285
(-14)	-0.27206	0.061403
(-15)	0.18809	0.056620
(-16)	-0.29360	0.069189
(-18)	0.13613	0.061049
(-19)	0.46340	0.065137
(-20)	-0.22250	0.059332
(-21)	-0.11110	0.058729
(-23)	0.15476	0.058507
(-24)	0.18781	0.056418
$\hat{\beta}_0$	0.002138	0.002617
$\hat{C}_1$	0.069284	0.015870
$\hat{C}_2$	0.080316	0.018181
$\hat{C}_3$	0.070736	0.017241
$\hat{C}_4$	-0.046745	0.016819
$\hat{C}_5$	0.080514	0.020284
$R^2 = .81254 \quad \bar{R}^2 = .76843 \quad \hat{\sigma} = 0.020492 \quad DW = 1.9710 \quad T = 169 \quad \hat{S} = 0.042583$		
<b>Test Statistics</b>	LM Version	F Version
A: Serial Correlation	Chi-sq(4)=5.3622	F(4,132)=1.0814
B: Functional Form	Chi-sq(1)=0.41409	F(1,135)=0.33160
C: Normality	Chi-sq(2)=0.15167	not applicable
D: Heteroskedasticity	Chi-sq(1)=0.000138	F(1,167)=0.0001125

where  $C_{i,i=1,2,3,4,5}$  are five dummy variables<sup>10</sup>. The model shows the astonishing complexity underpinning the Survey. To make practical use, however, of this model for forecasting requires a ‘cascade’ of intermediate  $\Delta X$  models, which provide forecasts of  $X_{m+7}$ ,  $X_{m+6}$ ,  $X_{m+5}$ ,  $X_{m+2}$  and  $X_{m+1}$  along with  $x_m$ . These models are illustrated by the following inverted ‘error correction’ for  $X_m$ , even though a more general form would be preferable in practice:

<sup>10</sup> The form is: 1972: Q3 (1) and 1973: Q2 (-1), 1978: Q3 (1) and 1979: Q2 (-1), 1978: Q4 (1) and 1979: Q3 (-1), 1979: Q1 (1) and 1979: Q4 (-1), as well as 1979: Q2 (1) and 1980: Q1 (-1).

$$Q(X_m) = 0.021917 \Delta l u_{q, f+1} + 0.29455 Q(X_m)_{-3} - 0.55414 U_{q-1} + 0.10139 d_1 + 0.11192 d_2 + 0.084128 d_3 + \varepsilon_q$$

$$(0.032774) \quad (0.047122) \quad (0.055807) \quad (0.018723) \quad (0.018845) \quad (0.018686) \quad [14]$$

$$R^2 = 0.62612 \quad \bar{R}^2 = 0.61512 \quad \hat{\sigma} = 0.026167 \quad DW = 1.8325, \quad T = 176, \quad \hat{S} = 0.042179, \quad AIC = 388.5321, \quad SBC = 379.0206.$$

where  $d_i, i=1,2,3$  are three dummy variables<sup>11</sup>. In fact, this equation shows that the explanation of expectation formation can be reversed indirectly to explain future output using Ordinary Least Squares, although the ‘classical’ assumptions underpinning regression estimation are not satisfied for this type of reversed analysis.

The predictive models can now be tested. Clearly, the absence of auto-correlation among the residuals is one of the necessary conditions for the forecasts, from the proposed models, to be boundedly rational. For an unbiased and efficient forecast of  $Q(X_m)$  in [14], the constant should be zero and the coefficient should be equal to one. The empirical study, over the time span of 1961: Q1 to 2004: Q4 derived the following estimates:

$$Q(X_m) = 0.0003443 + 0.99992 Q(\hat{X}_m) + \hat{V}_t$$

$$(0.019495) \quad (0.05859) \quad [15]$$

where  $Q(\hat{X}_m)$  is the fitted-values from [14] with  $\hat{V}_t$  free of serial-correlation. Both hypotheses under the null are statistically accepted, so the depiction is unbiased and efficient. Following the same procedure for the model in Table [1], the estimation resulted in:

$$\Delta(X_{m+8}) = 0.000 + 1.0000 \Delta(\hat{X}_{m+8}) + \hat{V}_t$$

$$(0.000) \quad (0.037169) \quad [16]$$

These forecasts appear to have the required, desirable properties.

The model in [14], in terms of  $\hat{X}_m$ , can also be compared with the Pesaran/Thomas (1995) procedure for generating expectations. This estimation of the Pesaran/Thomas format could not cope with the ‘full’ sample period because of heteroskedasticity, although the time scale between

---

<sup>11</sup> The format is as follows: 1978: Q2 (1) and 1979: Q1 (-1), 1980: Q1 (1) and Q2 (-1) as well as 1975: Q2 (-1) and 1989: Q1 (1).

1977: Q4 to 2004: Q4 led to a well-defined, backward-looking statistical representation of the form<sup>12</sup>:

$$\begin{aligned} X_m &= -0.061111 + 0.1389U_{q,b} - 0.10349d_1 + \hat{V}_t, \\ &\quad (0.0084985) \quad (0.015736) \quad (0.016435) \\ \hat{V}_t &= 0.40667\hat{V}_{t-1} - 0.25315\hat{V}_{t-2} - 0.22318\hat{V}_{t-4} + U_t. \\ &\quad (0.093351) \quad (0.096448) \quad (0.092188) \end{aligned} \quad [17]$$

$$R^2 = 0.60125, \bar{R}^2 = 0.5808, \hat{\sigma} = 0.027662, DW = 1.9160, T = 110, \hat{S} = 0.042790, AIC = 235.65, SBC = 227.55,$$

$$A: \chi_4^2 = 3.1773, B: \chi_1^2 = 0.010217, C: \chi_2^2 = 0.92905, D: \chi_1^2 = 0.073408.$$

$X_m$  is the backward-looking output growth<sup>13</sup> with  $U_{q,b}$ , the past, adjusted ‘ups’. The inclusion of the past residuals,  $\hat{V}_t$ , were tested as “instruments” using the Hausman (1978) test, which led to an F-statistic of 15.61, rejecting 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 forecast errors. The model in [17] was tested against the use of the unadjusted, past ‘ups’ (in [18] below), which led to a reduction in the AIC and SBC statistics of 227.2332 and 221.8323 respectively, showing the adjustment is superior. The diagnostic tests clearly indicate a ‘well-defined’ statistical format, even for the unadjusted data, but showing less explanatory power. The unadjusted, backward-looking model is as follows:

$$\begin{aligned} X_m &= -0.053502 + 0.25399U_{q,ab} - 0.10283d_1 + \hat{V}_t, \\ &\quad (0.0094409) \quad (0.036919) \quad (0.017787) \\ \hat{V}_t &= 0.36633\hat{V}_{t-1} + U_t. \\ &\quad (0.093351) \end{aligned} \quad [18]$$

$$R^2 = 0.51813, \bar{R}^2 = 0.504498, \hat{\sigma} = 0.030121, DW = 1.87140, T = 110, \hat{S} = 0.042790, AIC = 227.23, SBC = 221.83,$$

$$A: \chi_4^2 = 7.2478, B: \chi_1^2 = 0.43103, C: \chi_2^2 = 3.9049, D: \chi_1^2 = 0.28232. \quad U_{q,ab} \text{ is the data set representing the unadjusted, C.B.I. ‘ups’}.$$

<sup>12</sup> It is suggested by Driver and Urga (2004) in their article, that it is not possible to derive output expectations. The difficulty arises because they assume that the process is non-linear, when it can be captured by a linear format, as outlined in Thomas (1995).

<sup>13</sup> Plus a dummy variable to remove unusual events at the following points in the data set: 1979: Q1 (1), 1980: Q2 (1) and 1981: Q1 (1), which arose from recession affecting output growth.

According to Hansen (2005), in addition to use of the model selection statistics, the AIC and SBC indicated above, the root mean-squared error, or in this case the root mean squared forecast error (RMSFE) can be relevant in comparing models. The expected change in output over the next four months,  $U_{q,f}$ , the adjusted ‘ups’, are utilized in the following equation to create output expectations from [17] to aid comparison:

$$\hat{X}_m^e = -0.061111 + 0.1389U_{q,f} + 0.40667V_t - 0.25315V_{t-1} - 0.23318V_{t-3}. \quad [19]$$

These are compared with the fitted-values,  $\hat{X}_m$ , derived from expression [14]. In the case of [19], the RMSFE led to a value of 0.031912. For  $\hat{X}_m$ , from [14], the RMSFE was 0.024595. The logistic function is shown here to have the lower RMSFE, compared with the Pesaran/Thomas method.

Future manufacturing output is important for explaining developments in such factors as manufacturing capital expenditure and exports along with domestic production, and therefore, affects overall aggregate demand, the balance of payments and the level of spare capacity in the U.K. economy. The forecasting model described should influence both the Bank of England’s decisions on its interest rate policy and the U.K. Treasury’s choice of fiscal policy.

## **Conclusion**

The paper has focused on the process of formation of business expectations underlying the C.B.I. Survey. The study argues for the adoption of a bounded rationality interpretation in order to explain business expectation diffusion in the context of limited information, complexity and expense, suggesting a non-linear diffusion process of expectations, captured by the logistic model. It predicts that an increase in the percentage of managers expecting growth in manufacturing output will be associated with a rise in the Official rate of growth of output, estimated with observations on both backward- and forward-looking output expectations.

The use of the C.B.I. Survey data for forecasting future output is also illustrated. The results of the estimations provide evidence supporting the logistic specification, which has been extensively tested. This article has shown how to adjust the raw ‘ups’ and ‘downs’ statistics with the ‘sames’ to achieve symmetry in results.

Other recent research in this area, Mitchell et al (2002 and 2005) have focused on disaggregating the Survey observations by the respondent. Unfortunately, the cessation of availability of that data in 1999 has reduced its relevance. This study uses output data disaggregated by months, although the Survey expectations are quarterly.

What all the research activity based on the C.B.I. Industrial Trends Survey indicates is that it remains an important source of information for the U.K. Government, the Bank of England and British businesses as well as academics. Attention should continue to be focused on improving forecasting performance through research of disaggregated data sets, the functional form and the formation of Survey expectations. Indeed, the discussion in this study not only applies to the functional form and the disaggregation, but also relates to the formation of expectations by providing some empirical support for the process of diffusion. This is in line with models of bounded rationality, where decision-making is uncertain and complex, and where entrepreneurs make decisions with one eye on their businesses, and the other firmly fixed on the expectations of the ‘alphas’ of the pack.

## References

- Branch, W.A. (2004) The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations, *Economic Journal*, 114, July, pp 592-61.
- Carlson, J. A., and Parkin, M. (1975) Inflation Expectations *Economica* Vol. 42, pp 123 to 138.
- C.B.I (1983) *Twenty-Five Years of 'Ups' and 'Downs'*, Confederation of British Industry, London.
- Cramer, J.S. (1991) *The Logit Model: An introduction for Economists*, Routledge, Chapman and Hall Inc., New York.
- Cunningham, A. (1997) Quantifying Survey Data, *Bank of England Quarterly Bulletin*, August, pp 292-300.
- Driver, C. and Urga, G. (2004) Transforming Qualitative Survey Data: Performance Comparisons for the U.K., *Oxford Bulletin of Economics and Statistics*, Vol.66, pp 71-90.
- Evans, G.W. and Honkapohja, S. (2001) *Learning and Expectations in Macroeconomics*, Princeton University Press.
- Hansen, B.E (2005) Challenges for econometric Model Selection, *Econometric theory*, Vol. 21, pp 60-68.
- MacKinnon, J.G. (1991) Critical Values for Co-integration Tests, in *Long Run Equilibrium Relationships*, Engle, R.F. and Granger, C.W.J. (eds.), Oxford University Press, Oxford, pp 267-76.
- Mitchell, J., Smith, R.J., and Weale, M.R. (2002) Quantification of Qualitative Firm-Level Survey Data *Economic Journal* Vol. 112 (March) pp C117-C135.
- Mitchell, J., Smith, R.J., and Weale, M.R. (2005) Forecasting Manufacturing Output Growth, Using Firm-Level Survey Data, *The Manchester School*, Vol. 73, No.4, pp 479-99.

Pesaran, M.H. (1984) Expectations Formation and Macroeconometric Modelling, in *Contemporary Macroeconomic Modelling*, Malgrange, P. and Muet, P.A. (eds.), Blackwell, Oxford.

Sargent, T.J. (1993) *Bounded Rationality in Macroeconomics*, Clarendon Press, Oxford.

Simon, H. (1961) *Administrative Behaviour*, Second Edition, Original Publication 1947, Macmillan, New York.

Simon, H. (1986) Rationality in Psychology and Economics, *Journal of Business*, 59(4), pp 209-224.

Thomas, D.G. (1995) Output Expectations within Manufacturing Industry, *Applied Economics*, 27, pp 403-408.