A Bayesian Indicator of Manufacturing Output from Qualitative Business Panel Survey Data*

James Mitchell
National Institute of Economic and Social Research

Richard J. Smith
Faculty of Economics, Cambridge University

Martin R. Weale
National Institute of Economic and Social Research

January 2006. PRELIMINARY

Abstract

Qualitative business survey data are used widely to provide indicators of economic activity ahead of the publication of official data. Traditional indicators exploit only aggregate survey information, namely the proportions of respondents who report "up" and "down". This paper considers disaggregate or firm-level survey responses. It derives an alternative Bayesian indicator of economic activity relating firms' categorical responses to official data using ordered discrete-choice models. An application to firm-level survey data from the Confederation of British Industry shows that the proposed indicator of manufacturing output growth can provide more accurate early estimates of manufacturing output growth than traditional indicators.

Keywords: Survey data, Indicators, Quantification, Forecasting; Forecast Combination

*Address for correspondence: James Mitchell, NIESR, 2 Dean Trench Street, Smith Square, London, SW1P 3HE, U.K. Tel: +44 (0) 207 654 1926. Fax: +44 (0) 207 654 1900. E-Mail: j.mitchell@niesr.ac.uk.
This paper replaces an earlier paper entitled "Aggregate versus Disaggregate Survey-Based Indicators of Economic Activity". We wish to thank J. Morrison and J. Wood of the CBI for their advice about the survey data. Financial support from the ESRC (RES-000-22-1009) is also gratefully acknowledged.

1 Introduction

Statisticians and economists are under considerable pressure to produce up-to-date estimates of the state of the economy. With interest rates now being set to regular timetables in all major countries, interest rate setters have a regular need for up-to-date information. Moreover, those working in financial institutions who want to anticipate the actions of interest rate setters also require efficient means of estimating the state of the economy.

In recent years, building on Stone (1947), a number of authors have developed methods of relating the state of the economy to an unobserved latent variable, the most important factor in a static [see Stock & Watson (2002)] or dynamic [see Forni et al. (2001)] factor representation of a large array of economic data. These methods can be used in real-time to provide estimates of the state of the economy, even if some data accrue only with a lag. Nevertheless, this approach suffers from the considerable disadvantage that the relationship between the latent variable and GDP at real prices, the conventional and well-understood measure of economic activity, is by no means clear. Thus other authors [see Mitchell, Smith, Weale, Wright & Salazar (2005)] produce monthly estimates of GDP itself in a manner designed always to be aligned against the most recent official quarterly data.

Whichever approach is adopted, it is sensible to make whatever use can be made of data as they accrue. Collection and publication of official data is subject to processing delays; thus in the United Kingdom the monthly index of industrial production (including manufacturing output) is published about thirty-seven days after the end of the month to which it relates. Eurostat legislation imposes a maximum delay of forty-five days in the European Union. Qualitative surveys about the state of the industrial sector, are, however, published with a much shorter lag and their publication is usually accompanied by some discussion of what can be learned from them about the most recent movements and short-term expected future movements in economic activity, at least in the sector to which the surveys relate. These surveys ask *inter alia* whether, after adjusting for normal seasonal movements, output has risen, stayed the same or fallen in recent months. The question thus arises how formally to convert the findings of such surveys into early estimates of movements in economic activity. The traditional approach to this question has been to take the aggregate findings of such surveys, the proportion of firms reporting that output has risen, stayed the same or fallen, and relate them to official output data. Approaches suggested have included the probability method [Carlson & Parkin (1975)] and the regression method [Pesaran (1984, 1987)], plus variants of these.[1] Collectively, we call these approaches "aggregate". The inclusion of the reported proportions in the factor models mentioned above can also be seen as a means of relating the aggregate survey data to other economic variables.

In this paper we are concerned with a question which arises with any survey but which has been little discussed in the context of surveys of business activity.[2] How should the responses of the individual firms be quantified and combined if the aim of the survey is to produce an early indication of official output data?[3] There is no reason to believe that working with the aggregate findings of the survey is the best way of doing this; it may well be that quantification in a manner which allows for a degree of heterogeneity among firms exploits the information more efficiently than do the traditional approaches and therefore allows more accurate inferences to be drawn about output movements.

We construct a new indicator by using ordered discrete choice models to link individual firms' categorical survey responses to official data and then infer, *via* Bayes' Theorem, the most likely or expected value for the official data given firms' categorical responses. The indicator is designed to give more weight to firms whose answers have a close link to the official data than to those whose experiences correspond only weakly or not at all. This can be seen as a variant of the forecast combination problem addressed by Bates & Granger (1969) and Granger & Ramanathan (1984), although the form of the problem is rather different. The approach is also a development of that applied by Mitchell, Smith & Weale (2005*b*) to prospective survey data.

Use of the proposed indicator is illustrated in an application to industrial survey data from the Confederation of British Industry (CBI). We find that in-sample it explains more of the variation in manufacturing output growth than traditional indicators constructed using "aggregate" data. The evidence is mixed on an out-of-sample basis. However, the out-of-sample analysis is limited to just eight quarters, due to data constraints, and further work when data are available is therefore required to re-examine the new indicator's

[1]See Pesaran & Weale (2005) for a survey.

[2]In other areas of econometrics the benefit of analysing individual as well as aggregate data is generally recognised. There has been limited previous work using individual responses to qualitative surveys [see Nerlove (1983); Horvath et al. (1992); McIntosh et al. (1989); Branch (2004); Souleles (2004)]. However, this work has focused on testing the nature of expectation formation.

[3]Mitchell et al. (2002) developed a semi-disaggregate model showing that, in linking the survey to official data, the performance could be enhanced if attention was paid not only to the responses of individual firms but also to the extent to which these responses had changed compared with the previous survey. Nevertheless, in contrast to the model developed in this paper, their approach is only semi-disaggregate as it is based on the proportions; it does not take account of the relative informational content of individual survey responses.

performance.

The plan of the remainder of this paper is as follows. Section 2 motivates the Bayesian indicator which exploits disaggregate survey data. Section 3 describes the CBI data. Section 4 illustrates the use of the proposed indicator in an application to firm-level industrial survey data from the CBI. Section 5 considers the indicator out-of-sample. Section 6 concludes.

2 Quantification Across Firms

2.1 Ordered Discrete Choice Models

Consider a survey that asks a sample of N_t manufacturing firms at time t whether their output growth, for example, has risen, not changed or fallen relative to the previous period. Crucially the number of firms in the sample is allowed to vary across t.

The categorical responses in the survey are assumed to be related to economy-wide manufacturing output growth x_t in the following manner. Let the actual output growth of firm i at time t, y_{it}, $(i = 1, ..., N_t)$, depend on x_t according to the linear model

$$y_{it} = \alpha_i + \beta_i x_t + \varepsilon_{it}, \tag{1}$$

$(t = 1, ..., T)$, where α_i and β_i are firm-specific time-invariant coefficients. The error term ε_{it} captures the component of firm-specific output growth y_{it} unanticipated by both firm i and the econometrician at time t. More precisely, we assume the conditional linear specification $E(y_{it}|\Omega_t^i) = \alpha_i + \beta_i x_t$ where Ω_t^i comprises information available to firm i at time t and includes x_t. Hence, $E(\varepsilon_{it}|\Omega_t^i) = 0$ and ε_{it} is uncorrelated with x_t rendering x_t weakly exogenous by assumption. The validation of this and other assumptions, for example, the absence of dynamics in x_t, is a necessary concomitant in any empirical application. Indeed the model (1) can be straightforwardly augmented to accommodate the endogeneity of and dynamic dependence in x_t. See Appendix A which also describes the diagnostic tests to be employed. In the following analysis it is further assumed that output growth x_t is stationary. A fixed-effects interpretation for (1) is provided by (B.21) in Appendix B.3.

It is necessary that model (1) for firm-level growth y_{it} is coherent with the economy-wide outturn x_t. Let z_{it} denote (the level of) output of firm i at time t. From (1), after cross-multiplication and summation over $i = 1, ..., N_t$, $\sum_{i=1}^{N_t} \Delta z_{it} = \sum_{i=1}^{N_t} z_{it-1}\alpha_i +$

$\sum_{i=1}^{N_t} z_{it-1}\beta_i x_t + \sum_{i=1}^{N_t} z_{it-1}\varepsilon_{it}$, where Δ is the first difference operator. For coherency we therefore require that $\sum_{i=1}^{N_t}\Delta z_{it}/\sum_{i=1}^{N_t} z_{it-1} \xrightarrow{p} x_t$, $\sum_{i=1}^{N_t} z_{it-1}\alpha_i/\sum_{i=1}^{N_t} z_{it-1} \xrightarrow{p} 0$, $\sum_{i=1}^{N_t} z_{it-1}\beta_i/\sum_{i=1}^{N_t} z_{it-1} \xrightarrow{p} 1$ and $\sum_{i=1}^{N_t} z_{it-1}\varepsilon_{it}/\sum_{i=1}^{N_t} z_{it-1} \xrightarrow{p} 0$ $(N_t \to \infty)$.

Actual growth y_{it} of firm i at time t is unobserved but the survey contains data corresponding to whether output growth has risen, not changed or fallen relative to the previous period. To account for the ordinal nature of the responses, we use ordered discrete choice models based on the latent regression (1). Define the indicator variables

$$y_{it}^j = 1 \text{ if } \mu_{(j-1)i} < y_{it} \leq \mu_{ji} \text{ and 0 otherwise, } (j = 1, 2, 3), \tag{2}$$

corresponding to "down", "same" and "up", respectively, where $\mu_{0i} = -\infty$, μ_{1i}, μ_{2i} and $\mu_{3i} = \infty$ are firm-specific threshold parameters. We assume that the error terms ε_{it}, $(t = 1, ..., T)$, are logistic with common cumulative distribution function (c.d.f.) $F(z) = [1+\exp(-z)]^{-1}$, $-\infty < z < \infty$, $(i = 1, ..., N_t)$. The logistic distribution is similar in shape to the normal but has slightly heavier tails and is particularly convenient since it offers a closed form distribution function. The probabilistic foundation for the observation rule (2) is given by the conditional probability $P_{jit} = P_i(j|x_t, i)$ of observing the categorical response $y_{it}^j = 1$ for choice j at time t given the value of x_t and firm i

$$P_{jit} = F(\mu_{ji} - \alpha_i - \beta_i x_t) - F(\mu_{(j-1)i} - \alpha_i - \beta_i x_t), \ (j = 1, 2, 3). \tag{3}$$

As discrete choice models are only identified up to scale, including the intercept α_i in (1) necessitates setting, for example, the first threshold parameter μ_{1i} to zero to achieve identification. Consequently the decision probabilities (3) are invariant to multiplying (1) by an arbitrary constant. Assuming the errors ε_{it} are independently and identically distributed over time, the likelihood function for firm i is

$$L_i = \prod_{t=1}^{T} P_{1it}^{y_{it}^1} P_{2it}^{y_{it}^2} P_{3it}^{y_{it}^3}. \tag{4}$$

Under the above assumptions, maximisation of (4) yields consistent estimates $(T \to \infty)$ of α_i, β_i and μ_{ji} denoted by $\hat{\alpha}_i$, $\hat{\beta}_i$ and $\hat{\mu}_{ji}$ respectively. In addition, assuming firms are randomly sampled and that the model (1) is correctly specified, the errors ε_{it} are also independently distributed over firms. If we let $\varepsilon_{.t}$ denote the N_t-vector of ε_{it}'s and let $\Sigma_t = Var(\varepsilon_{.t})$, this implies that Σ_t is diagonal. In this case there is no efficiency loss involved in estimation of the discrete choice models firm-by-firm rather than as a system.

2.2 Inferring the Official Data: the Proposed Indicator

Given ordered logit model for each firm i, $i = 1, ..., N_t$, an estimator for x_t may be inferred from these survey data. As survey data are usually published ahead of the official data, this provides an early quantitative estimate of x_t. Since they are not subject to revision they must be assessed against near-final official data.

Let j_{it}, ($j_{it} = 1, 2, 3$), denote the survey response of firm i at time t, where 1, 2 and 3 correspond to "down", "same" and "up", respectively. We need to work out the density function of x_t conditional on the N_t firms' observed survey responses at time t, $\{j_{it}\}_{i=1}^{N_t}$. We denote this density function $f(x_t|\{j_{it}\}_{i=1}^{N_t})$.

Let $f(x_t)$ denote the prior density function of x_t. This density function can be conditioned on lagged values of x_t, say x_{t-1}, when x_t follows a dynamic process; see Appendix A. Also as Appendix A explains, for each firm one can consider ordered logit models augmented with x_{t-1} when x_{t-1} is statistically significant in the firm-level model (1); cf. (A.3) below. Below the exposition, without loss of generality, considers both the unconditional density for x_t and the case when x_{t-1} is statistically insignificant in the firm-level model.

Diagonality of Σ_t implies that conditional on x_t firms' categorical responses are independent across firms. That is, the joint conditional probability of observing the N_t firms' categorical responses, $\{j_{it}\}_{i=1}^{N_t}$, is given as the product of their marginal probabilities $P(j_{it}|x_t, i)$:

$$P(\{j_{it}\}_{i=1}^{N_t}|x_t) = \prod_{i=1}^{N_t} P(j_{it}|x_t, i). \tag{5}$$

Therefore, the joint conditional probability of observing response j across firms i, ($i = 1, ..., N_t$), is given as

$$P(\{j_{it}\}_{i=1}^{N_t}) = \int_{-\infty}^{\infty} \prod_{i=1}^{N_t} P(j_{it}|x_t, i) f(x_t) dx_t. \tag{6}$$

Bayes' Theorem states that:

$$f(x_t|\{j_{it}\}_{i=1}^{N_t}) = \frac{\prod_{i=1}^{N_t} P(j_{it}|x_t, i) f(x_t)}{P(\{j_{it}\}_{i=1}^{N_t})}. \tag{7}$$

Then we can define the indicator D_t. D_t is given as the Bayes estimator (under squared error loss) for x_t given $\{j_{it}\}_{i=1}^{N_t}$ which is the mean of the posterior density $f(x_t|\{j_{it}\}_{i=1}^{N_t})$:

$$D_t = E(x_t|\{j_{it}\}_{i=1}^{N_t}) = \int_{-\infty}^{\infty} x_t f(x_t|\{j_{it}\}_{i=1}^{N_t}) dx_t. \tag{8}$$

Given $f(x_t)$, all of the above integrals may be calculated by numerical evaluation. Estimators $\hat{P}(j_{it}|x_t, i)$ for $P(j_{it}|x_t, i)$ and, thus, $\hat{P}(j_{it}|i)$ for $P(j_{it}|i)$ are given by substitution of the estimators $\hat{\alpha}_i$, $\hat{\beta}_i$ and $\hat{\mu}_{ji}$, $(j = 0, ..., 3)$, in (3). Hence, a feasible Bayes estimator $D_t = \hat{E}(x_t|\{j_{it}\}_{i=1}^{N_t})$ may be obtained from (8) by numerical evaluation.

The indicator D_t considers all firms' responses, $i = 1, ..., N_t$, simultaneously. It is designed to give more weight to firms whose answers have a close link to the official data than to those whose experiences correspond only weakly or not at all. This can be seen as a variant of the forecast combination problem addressed by Bates & Granger (1969) and Granger & Ramanathan (1984). There are obviously any number of reasons why some firms might be more useful as indicators than others, ranging from the nature of the business that they run to the care they employ in completing the survey-return and it is intuitive that study of individual firms' performances provides valuable information lost in aggregation.

To illustrate this property consider $\beta_1 = 0$. This implies that firm 1's categorical survey responses offer no information about the official data. For this firm $P(j_{it}|x_t, i) = P(j_{it}|i)$. (7) is then given as

$$f(x_t|\{j_{it}\}_{i=1}^{N_t}) = \frac{\prod_{i=2}^{N_t} P(j_{it}|x_t, i) f(x_t)}{P(\{j_{it}\}_{i=2}^{N_t})}, \tag{9}$$

implying firm 1 receives no weight in the indicator D_t.

This approach can be contrasted with Mitchell, Smith & Weale (2005*a*, 2005*b*). Separately for each firm they calculate density functions for x_t conditional on the survey responses j_{it} and then take an average of these across firms. Let us outline the steps involved in computing this firm-level indicator.

The conditional probability of observing response j for firm i is $P(j_{it}|i) = \int_{-\infty}^{\infty} P(j_{it}|x_t, i) f(x_t) dx_t$. Bayes' Theorem states that

$$f(x_t|j_{it}, i) = \frac{P(j_{it}|x_t, i) f(x_t)}{P(j_{it}|i)}. \tag{10}$$

For firm i, the Bayes estimator (under squared error loss) for x_t given j_{it} is the mean of the posterior density $f(x_t|j_{it}, i)$:

$$E(x_t|j_{it}, i) = \int_{-\infty}^{\infty} x_t f(x_t|j_{it}, i) dx_t, \tag{11}$$

which takes one of three values depending on the observed sample response j_{it} of firm i at time t. For a firm whose answers are unrelated to movements in the official series, these mean estimates will be the same for each category of answer. This will be simply the mean growth rate of the official series. This follows from the fact that in a large time-series, if a firm responds at random the disaggregate method gives the same score (mean output growth) to all categorical responses; i.e. since $P(j_{it}|x_t, i) = P(j_{it}|i)$, $E(x_t|j_{it}, i) = m_1'$, where m_1' is the sample mean of $\{x_t\}$; see (11). In all other cases these estimates will provide some indication about the growth of the official series.

Assuming firms are independent, their indicator $\overline{D}_t$ of economic activity at time t is then given as[4]:

$$\overline{D}_t = (1/N_t) \sum\nolimits_{i=1}^{N_t} \hat{E}(x_t|j_{it}, i). \tag{12}$$

3 CBI Survey Data

The *Industrial Trends Survey* (ITS) of the CBI, which is conducted on a quarterly basis, gives qualitative opinion from UK manufacturing firms on past and expected trends in output, exports, prices, costs, investment intentions, business confidence and capacity utilisation. In our application we consider the following question:

- "Excluding seasonal variations, what has been the trend over the past four months with regard to volume of output?".

Firms can respond either "up", "same", "down" or "not applicable". This retrospective question provides the basis of deriving timely indicators of manufacturing output growth x_t. The number that answer "not applicable" is very small and ignored in later analysis. Although there is a one month overlap on each survey as firms are asked to report over a four month period four times a year, as the responses are qualitative this aspect of the data is viewed as unlikely to be important.

We consider a sample of $43,936$ responses from the ITS. The sample records the survey responses of, in total, 5002 firms over the period 1988q3 to 1997q3, although we do extend the sample to 1999q3 to analyse the out-of-sample performance of the alternative nowcasts. There are, on average, only 1183 firms in the sample at time t, with 8.7 time-series observations per firm. Many observations are missing as firms do not always respond to consecutive surveys. This prevents the construction of a panel data set with sufficient

[4] As well as equal weights, $(1/N_t)$, Mitchell, Smith & Weale (2005a, 2005b) consider weights based on firm-size.

time-series observations across all firms for the estimation of (1) without assuming some homogeneity in behaviour across firms. Quantification based on (1) requires sufficient time-series observations for a given firm for reliable parameter estimation.

In this application, we consider twenty observations to be satisfactory.[5] If, given i, the error terms ε_{it} are independent conditional on x_t, $(t = 1, ..., T)$, these observations need not be consecutive. Hence, firms that do not respond to at least twenty surveys are dropped from the sample used to derive D_t (and $\overline{D}_t$) indicators of manufacturing output growth. Since these firms are dropped there is a danger that the sample selection could induce bias in the D_t (and $\overline{D}_t$) indicator.[6] In any case, notwithstanding the implied theoretical properties of the indicators, their usefulness is determined by how well they perform in practice, both in-sample and out-of-sample, relative to the traditionally used quantification techniques employed on the aggregate survey data. This should, and does, serve as the main test of their value.

An alternative approach, that does not require some firms to be removed, is a random effects reformulation of (1) which imposes homogeneity restrictions across firms. Re-express (1) as $y_{it} = \alpha + \beta x_t + \zeta_{it}$, where $\zeta_{it} = (\alpha_i - \alpha) + (\beta_i - \beta)x_t + \varepsilon_{it}$ and $E(\alpha_i) = \alpha$, $E(\beta_i) = \beta$. Random effects estimation requires the evaluation of T-dimensional integrals which may be achieved by the use of the Geweke-Hajivassiliou-Keane simulator; see, for example, Keane (1994). In general, however, $E(\alpha_i|\Omega_t) \neq \alpha$, $E(\beta_i|\Omega_t) \neq \beta$ where Ω_t comprises information available to *all* firms at time t which includes x_t. That is, the fixed effects α_i and β_i are correlated with the outturn x_t, rendering traditional random effects panel-data estimators inconsistent through the presence of heterogeneity bias. Since our results indicate considerable heterogeneity across firms in the slope coefficients, we do not follow this approach here.

Over the period $1988q3 - 1997q3$ twenty non-consecutive time series observations are available for 643 manufacturing firms. This increases to 833 firms when the out-of-sample period is considered too. To give an impression of the nature of the survey responses,

[5]Of course, this choice is somewhat arbitrary and warrants further investigation *via* Monte-Carlo experiments. In related work we have taken an eclectic approach and when examining the performance of the $\overline{D}_t$ indicator considered a range of so-called "cut-off" values; see Mitchell, Smith & Weale (2005*a*). In practice $\overline{D}_t$ appears to behave similarly across a wide range of cut-off values.

[6]In Mitchell, Smith & Weale (2005*a*, 2005*b*) we tested for, and subsequently rejected, sample selection. The test involved comparing the performance of the aggregate indicators in the "included" and "excluded" samples. In the absence of sample selection, the included sample may be regarded as a random sample from the full-sample and inference from both included and excluded samples should be equivalent apart from sampling error. That is, indicators or statistics derived from both included and excluded samples should not differ significantly.

Figure 1 plots the percentage of these 643 firms that reported an "up", "same" or "down" response over the data period. It also plots the quarterly growth at an annual rate of (seasonally adjusted) manufacturing output. Visual inspection of the graph suggests that the survey responses track movements in manufacturing output growth at least in the sense that there appears to be more pessimism during recessions and more optimism in expansionary periods.

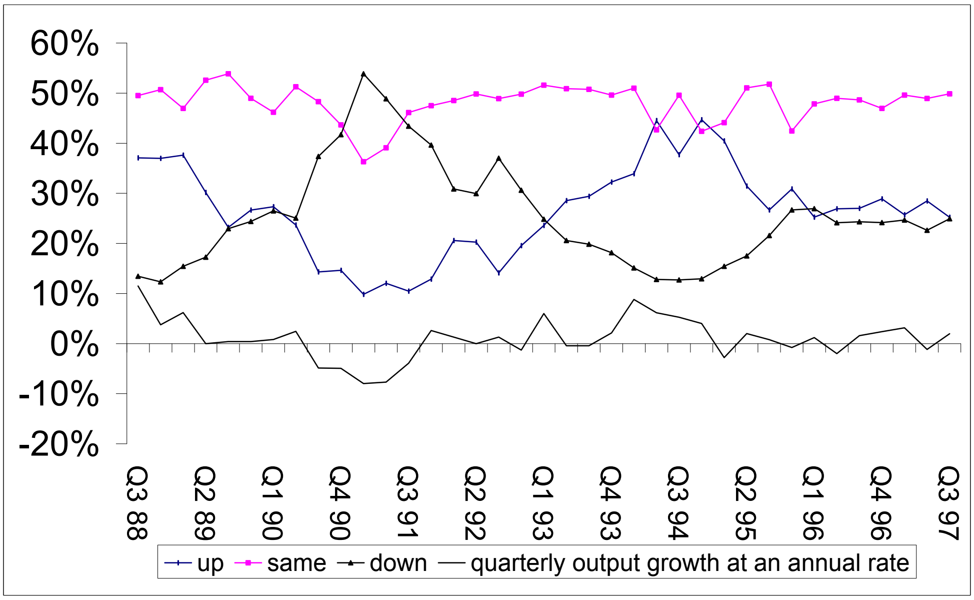

Figure 1: Unweighted percentage of firms reporting "up", "same" or "down"

4 Indicators of Manufacturing Output

4.1 Firm-Level Estimation of the Relationship Between Survey Responses and Manufacturing Output Growth

The indicator D_t (and $\overline{D}_t$) is based on firm-level estimation. To gain an impression of what the firm-level estimates look like, the degree of heterogeneity across firms and the empirical support for the chosen specification, for illustrative purposes we focus on those

643 manufacturing firms that over the period $1988q3 - 1997q3$ reply to at least twenty surveys.

We estimated ordered logit models for each of the 643 firms. A Wald test (fixed N_t) rejected the null hypothesis $\beta_i = \beta$ for all i with a p-value of 0.00. Firms, thus, appear to be heterogenous in terms of how they react to changes in the aggregate environment. Table 1 gives an impression of this heterogeneity. It displays the number of firms that have t-ratios for testing $\beta_i = 0$ in a specified range; firms are sorted by size as measured by sales volumes.

Table 1: t-ratios for $\widehat{\beta}$: the number of firms in a specified range with firms sorted by firm size

Firm Size		t-ratio (t_i)					
	Quintile	$t_i \leq -2$	$-2 < t_i \leq -1$	$-1 < t_i \leq 0$	$0 < t_i \leq 1$	$1 < t_i \leq 2$	$t_i > 2$
small	1	1	9	14	44	41	19
↓	2	0	5	20	32	42	29
	3	2	4	22	30	37	33
↓	4	1	6	28	45	33	25
big	5	0	8	17	37	36	33

Table 1 reveals the considerable variation across firms in how their survey responses relate to manufacturing output growth. There is no transparent relationship between firm size and the t-statistic t_i, and, thus, the ordered logit estimator $\hat{\beta}_i$. It is noteworthy that the survey responses of only 4 of the 643 firms individually have a significantly negative relationship with x_t with 139 significantly positive based on a one-sided test at the 0.025 level. This is consistent with our prior that, in general, we should expect a rise in manufacturing output growth to be associated with a rise in reported firm-specific output. However, as Table 1 suggests, the joint hypothesis of no positive relationship could not be rejected using a Bonferroni test (fixed N_t) based on these t-statistics.

The firm-level logit models were also subjected to the specification tests described in Appendix A. Both the test of the statistical significance of x_{t-1} in the firm-level model (A.3), $\rho_i = 0$, $(i = 1, ..., N_t)$, and a score test for misspecification are considered. This latter test is a joint test for omitted variables (specifically x_{t-1} and powers of $\widehat{\beta}_i x_t$ to test for incorrect functional form), neglected conditional heteroskedasticity and asymmetry in the distribution of the error terms ε_{it}; see Murphy (1996). Results are presented using Bonferroni adjusted critical values. Table 2 reports the proportion of times, across the

643 firms, these tests did not reject. The results are supportive of the chosen specification (1) for the firm-level model.

Table 2: Specification tests for ordered discrete choice models. Proportion of times the specification tests were not rejected

$\rho_i = 0$	1.000
Score	0.964

4.2 The Density Function $f(.)$

It remains to specify $f(.)$. The assumption that x_t is stationary is supported by tests for a unit root in the level series of manufacturing output. In addition one cannot reject the null hypothesis that $\beta_x = 0$ using a long-run of data from the early 1970s.[7] The sample observations for x_t have sample mean and sample standard deviation equal to 1.12 and 4.32 respectively. The sample exhibits little skewness, $\sqrt{b_1} = m_3/m_2^{3/2} = 0.046$, where m_i is the i-th sample central moment, but there is some evidence of kurtosis, $b_2 = m_4/m_2^2 = 3.571$, indicating that $f(.)$ may have thicker tails than the normal distribution. A modified version of the Jarque-Bera test that is robust against serial correlation and conditional heteroscedasticity in x_t [see Bai & Ng (2005)] does not reject the normality of $f(.)$ with a p-value of 0.790. We also experimented with the Pearson family of density functions, but the performance of the disaggregate indicators D_t (and $\overline{D}_t$) was not affected.

4.3 Comparing the Performance of the Indicators

We compare the performance of the indicators D_t and $\overline{D}_t$ against that of four traditional quantification techniques employed on aggregate proportions: the balance statistic [BAL], the probability method of Carlson & Parkin (1975) [CP], the regression approach of Pesaran (1984, 1987) [P] and the reverse-regression approach of Cunningham et al. (1998) [CSW] based on the logistic distribution. See Appendix B for a review.

The CP aggregate indicator is identified up to a scaling parameter [see Appendix B]. Following CP we chose this parameter to ensure that the mean of the indicator is

[7] Rejection is consistent with the fall in volatility, and persistence, of economic activity observed across the G7 in the last 25 years; e.g. see Stock & Watson (2003). Of course macroeconomic time-series often exhibit structural instablities or breaks. If the data generating process changes over the sample period then estimation over a partial or rolling window of the available observations might be felicitous.

equal to the mean of the outturn over the sample-period. This does not imply that the indicator is unbiased in the statistical sense. In contrast, the regression and reverse-regression indicators are unbiased since they implicitly estimate the scaling parameter through regression-based methods.

Table 3 summarises the performance of the indicators. For completeness, Table 4 assesses performance including the out-of-sample period too. Tables 3 and 4 reveal that the new indicators provide more accurate early estimates of output growth than traditional indicators employed on the aggregate proportions. Regardless of how the indicators are scaled, the higher correlation of the D_t and $\overline{D}_t$ indicators indicates that a stronger signal about the official data may be recovered from them than the aggregate data.

The two indicators D_t and $\overline{D}_t$ exhibit a similar correlation against the outturn for manufacturing output growth. However, D_t performs better than $\overline{D}_t$ with respect to the root mean squared error [RMSE] criterion. This is because despite the sample mean of $\overline{D}_t$ approximately estimating that of the outcomes x_t correctly, it appears too smooth. It displays too little volatility as compared with the outturn x_t. This feature has been observed elsewhere for alternative indicators; see, for example, Cunningham (1997). Less volatility is observed because the scale is incorrect. An explanation for this finding arises from consideration of those firms whose responses are poorly correlated with actual output growth. In the extreme case where responses are uncorrelated with output, inclusion of these firms reduces the standard deviation of the $\overline{D}_t$ indicator but does not affect its correlation with output growth. This follows from the fact that in a large time-series, if a firm responds at random the firm-level disaggregate method gives the same score (mean output growth) to all categorical responses; i.e. $E(x_t|j_{it}, i) = m'_1$. For these firms therefore there is no contribution to the variance of the aggregate $\overline{D}_t$. Excess smoothness of $\overline{D}_t$ may thus be viewed as due to the presence of firms in the sample whose responses contain little or no signal about output growth.

However, D_t does not suffer from this problem since, as indicated above, it is designed to give more weight to firms whose answers have a close link to the official data than to those whose experiences correspond only weakly or not at all. Therefore while D_t has a similar, indeed slightly improved, correlation against x_t, it is not too smooth. D_t better picks up the scale of x_t. This is evidenced by a higher standard deviation and lower RMSE than $\overline{D}_t$.

Table 3: Indicator Performance: In-sample 1988q3 − 1997q3

	Mean	Stand. Dev.	Corr.	RMSE
Outturn for Manuf. Output Growth	1.1213	4.3182		
BAL	1.1213	2.8472	0.6593	3.2024
CP	1.1213	192.6461	0.6623	187.2309
P	1.1213	2.9921	0.6928	3.0712
CSW	1.1213	6.3224	0.6830	4.5552
D_t	1.2128	4.9649	0.8453	2.6209
$\overline{D}_t$	1.1317	0.4374	0.8318	3.9079

Table 4: Indicator Performance: In-sample 1988q3 − 1999q3

	Mean	Stand. Dev.	Corr.	RMSE
Outturn for Manuf. Output Growth	1.0856	4.0450		
BAL	1.0856	2.3602	0.5834	3.2484
CP	1.0856	4.8105	0.5668	7.7585
P	1.0856	2.3711	0.5861	3.2406
CSW	1.0856	6.7042	0.6033	5.2866
D_t	0.8817	4.8537	0.8200	2.7576
$\overline{D}_t$	1.0765	0.3653	0.8117	3.7125

5 Out-Of-Sample Analysis

Given the improved in-sample fit between the survey responses and official data using D_t rather than traditional indicators, we examine whether this superiority extends out-of-sample. To evaluate how accurate survey-based early estimates of output growth would have been out-of-sample an experiment designed to mimic "real-time" application of the different quantification approaches is undertaken. We are nevertheless assessing their performance against near-final rather than initial official data.

The out-of-sample analysis is conducted over the 8 periods, 1997q4 − 1999q3. Unfortunately it was not possible to extend the out-of-sample analysis beyond 1999 since in 1999q4 the CBI moved to a new survey processing platform that involved changing the participant identification numbers making it very difficult to match firms pre- and post-December 1999. Results presented below for the D_t and $\overline{D}_t$ indicators allow "new" firms to enter the sample during the out-of-sample period. Similar results were obtained when we focused on those firms who had given at least twenty survey responses during the in-sample period ending in 1997q3.

The out-of-sample analysis involves computing the indicators using both survey and official data from 1988q3 to 1997q3, as outlined above, and then using these in-sample estimates to infer output growth in 1997q4 given knowledge of the survey data in 1997q4, but crucially not the official data on output growth. Given that survey data are published ahead of official data this provides an early estimate of output growth. Data from 1988q3 to 1997q4 are then used along with survey data in 1998q1 to infer output growth in 1998q1, involving re-estimation of both the ordered logit models and the moments of $f(.)$. This recursive process is repeated until finally output growth in 1999q3 is inferred using survey and official data 1988$q3 - 1999q2$, plus survey data in 1999q3.

The results of this recursive exercise are summarised in Table 5.

Table 5: Indicator Performance: Out-of-sample, 1997$q4 - 1999q3$

	RMSE
BAL	3.5625
CP	2301.9
P	4.2547
CSW	8.7415
D_t	3.9550
$\overline{D}_t$	2.3502

Table 5 provides mixed evidence about the performance of the D_t and $\overline{D}_t$ indicators relative to the traditional indicators. We note that the poor performance of the Carlson-Parkin indicator is due to the recursively computed estimates of the scaling parameter $\hat{\lambda}$; see (B.7) below. If full-sample information (data up to 1999q3) is used to estimate λ the RMSE of CP reduces to 7.9543. Table 5 shows that $\overline{D}_t$ but not D_t beats the traditional indicators. But the better performance of $\overline{D}_t$ is explained simply by its smoothness. Output growth is far less volatile in the out-of-sample than in-sample period. Accordingly, a forecast close to the unconditional mean of x_t, like $\overline{D}_t$, performs well. It is therefore important for future work to re-consider the out-of-sample performance of the D_t indicator with different sample periods and different data-sets.

6 Concluding Comments

This paper develops an efficient means of extracting quantitative signals from qualitative survey data about the business environment. The approach is statistically coherent, being derived from the application of Bayes Theorem to a statistical model of the individual qualitative responses to the survey. Unlike methods based on aggregate results it takes account of the relative informational content of each individual survey response. An improved means of extracting the underlying signal from qualitative categorical data ahead of the publication of official data means that economic policy setting can be more firmly based. The method developed is applicable to other qualitative surveys. In addition a similar approach could be used to address the question on expected future output growth.

In an application to survey data from the CBI, in-sample the proposed indicator outperformed traditional indicators in terms of anticipating movements in manufacturing output growth. Out-of-sample the evidence was mixed; but data availability seriously limited the feasibility of the out-of-sample experiment. It is therefore important for future work to re-consider the performance of the proposed indicator in different contexts.

A Appendix A. Dynamics and Specification Tests

It is important to test the implied restrictions embodied in (1), particularly given that macroeconomic data are widely accepted to exhibit dependence over time. Model (1) may be viewed as a restricted form of a more general formulation that allows for both endogeneity of and dynamic dependence in the official data x_t.

This generalization consists of (1) augmented by a stationary autoregressive process governing the determination of output growth x_t. Let x_t follow the first order autoregressive process

$$x_t = \alpha_x + \beta_x x_{t-1} + u_t, \tag{A.1}$$

($t = 1, ..., T$), where $|\beta_x| < 1$ and u_t is an *i.i.d.* zero mean disturbance. Additional lagged terms in x_t may be included in (A.1) if x_t is generated by a higher order autoregressive process. We assume that conditional on u_t the dependence between ε_{it} and u_t takes the linear form

$$\varepsilon_{it} = \rho_i u_t + \xi_{it}, \tag{A.2}$$

where ρ_i is a firm-specific parameter and ξ_{it} is an *i.i.d.* disturbance distributed as logistic and independently of u_t, ($i = 1, ..., N_t$).

Substitution of (A.2) in (1) generates the conditional dynamic model

$$\begin{aligned} y_{it} &= \alpha_i + \beta_i x_t + \rho_i u_t + \xi_{it} \\ &= \alpha_i^* + \beta_{i0}^* x_t + \beta_{i1}^* x_{t-1} + \xi_{it}, \end{aligned} \tag{A.3}$$

($i = 1, ..., N_t$), where the firm-specific coefficients $\alpha_i^* = \alpha_i - \rho_i \alpha_x$, $\beta_{i0}^* = \beta_i - \rho_i$ and $\beta_{i1}^* = -\rho_i \beta_x$. A test of $\rho_i = 0$, ($i = 1, ..., N_t$), or the exclusion of the error term u_t in (A.3) jointly tests for the absence of dynamics and the weak exogeneity of x_t in (1). A simple two-step test of $\rho_i = 0$ may be formulated similarly to the procedures described in Smith & Blundell (1986) and Newey (1987). Firstly, (A.1) is estimated by least squares which yields the consistent estimates ($T \to \infty$), $\hat{\alpha}_x$ and $\hat{\beta}_x$, for α_x and β_x, and the residual $\hat{u}_t = x_t - \hat{\alpha}_x - \hat{\beta}_x x_{t-1}$, ($t = 1, ..., T$). Secondly, the augmented model (A.3) is estimated by ordered logit as in section 2 after substitution of $\hat{u}_t$ for u_t. Analogously to (3) P_{jit} is now defined as

$$P_{jit} = F(\mu_{ji} - \alpha_i - \beta_i x_t - \rho_i u_t) - F(\mu_{(j-1)i} - \alpha_i - \beta_i x_t - \rho_i u_t), \; (j = 1, 2, 3). \tag{A.4}$$

Finally, the hypothesis $\rho_i = 0$ may be assessed by a t-test based on the resultant estimate of ρ_i. Failure to reject $\rho_i = 0$ supports the use of (1) while its rejection implies that the official data should be inferred using the augmented conditional model (A.3); see section 2.2 above. To mitigate the effects of an inflated Type I error when testing $\rho_i = 0$ across i, $(i = 1, ..., N_t)$, Bonferroni adjusted critical values are used.

The macroeconomic dynamics determined by β_x are distinct from dynamics in the firm-level models determined by ρ_i. It is therefore important to test the statistical significance of both parameters. When inferring the official data in section 2.2 when $\beta_x \neq 0$ we should consider $f(x_t \mid x_{t-1})$ rather than $f(x_t)$; in addition when $\rho_i \neq 0$ we should consider $P(j_{it}|x_t, x_{t-1}, i)$ rather than $P(j_{it}|x_t, i)$. The indicator D_t would then be given as $\int_{-\infty}^{\infty} x_t f(x_t|x_{t-1}, \{j_{it}\}_{i=1}^{N_t}) dx_t$.

Other implicit assumptions in (1) include linearity, homoskedasticity and that ε_{it} is distributed as logistic. Additional score or Lagrange Multiplier tests of misspecification appropriate for the ordered logit model should be employed to ascertain the empirical validity of (1); see, for example, Chesher & Irish (1987), Machin & Stewart (1990) and Murphy (1996).

While it is important if undertaking structural inference to ensure the model adequately explains the data, it is well known that there is little reason to expect a good in-sample fit to translate into good forecasts. We therefore undertake out-of-sample experiments to assess the forecasting performance of the selected models against traditional forecasts in section 5.

B Appendix B: Aggregate Quantification Techniques: A Review

Consider a survey that asks a sample of firms, for example, whether output growth x_t was "down", "same" or "up" relative to the previous period. Since the proportion of respondents who replied "down", "same" or "up" sum to unity the survey contains two pieces of independent information at time t. Let U_t and D_t denote the proportion of firms that reported an output rise and fall.

Although quantification of categorical survey responses is to some extent arbitrary, since survey responses are a firm's subjective assessment of the expected or actual behaviour of x_t, at the aggregate level quantitative measures of the expected or observed movement of x_t can be derived given certain assumptions. In this appendix four alternative methods of quantification are reviewed:

- the balance statistic and the probability approach of Carlson & Parkin (1975);
- the regression approach of Pesaran (1984, 1987);
- the reverse-regression approach of Cunningham et al. (1998) and Mitchell et al. (2002).

Although motivated in different ways, the four approaches are shown to share a common foundation. Our discussion compares the latter two methods to the probability approach and draws on Pesaran (1987) and Mitchell et al. (2002). For alternative reviews and extensions of the probability and regression approaches, see Pesaran & Weale (2005).

B.1 The Balance Statistic and the Probability Approach

This approach was first used by Theil (1952) to motivate the use of the "balance statistic" $U_t - D_t$ [see Anderson (1952)] as a method of quantification. The balance statistic, up to a scalar factor, provides an accurate measure of *average* output growth x_t if the percentage change in output of firms reporting a fall and the percentage change for firms reporting a rise are constant over time. The probability approach relaxes this restrictive assumption.

The probability method of quantification assumes that the response of firm i concerning economy-wide manufacturing output growth x_t is derived from a subjective probability density function for x_t, $f_i(.|i)$, which may differ in form across firms and is conditional on information available to firm i at time t; the dependence of $f_i(.|i)$ on t is suppressed in the discussion.

The responses of firm i are classified as follows. Let $x_{it} = \int x f_i(x|i)dx$ denote the mean of $f_i(.|i)$.

- "up" is observed if $x_{it} \geq b_{it}$;
- "down" is observed if $x_{it} \leq -a_{it}$;
- "same" is observed if $-a_{it} < x_{it} < b_{it}$,

where the threshold parameters a_{it} and b_{it} are both positive.

Assume that firms are independent and that the structure of $f_i(.|i)$ is the same and known for all firms; that is, $f_i(.|i) = f(.|i)$. Consequently, $x_{it} = \int x f(x|i)dx$ can be regarded as an independent draw from an aggregate density $f(x) = \int f(x|i)F(di)$, where

$F(.)$ denotes the distribution function of firms i; the density $f(.)$ is conditional on aggregate information available to all firms at time t, the dependence on which is again suppressed. Assume $f(.)$ has mean x_t.

Furthermore, if the response thresholds are symmetric and are fixed both across firms i and time t, that is, $a_{it} = b_{it} = \lambda$, then

$$D_t \xrightarrow{p} P(x_{it} \leq -\lambda) = F_t(-\lambda), \tag{B.1}$$

$$U_t \xrightarrow{p} P(x_{it} \geq \lambda) = 1 - F_t(\lambda), \tag{B.2}$$

where $F_t(.)$ is the cumulative distribution function obtained from $f(.)$ where, now, we indicate explicitly the dependence on time t. Then, as x_{it} is an unbiased predictor for x_t, we can estimate x_t given a particular value for λ and a specific form for the aggregate distribution function $F_t(.)$.

B.1.1 Carlson and Parkin's Method

The traditional approach of Carlson & Parkin (1975) assumes that $f(.)$ is a normal density function with mean x_t and variance σ_t; alternative densities $f(.)$ may be also considered; see Batchelor (1981) and Mitchell (2002).

From (B.1) and (B.2), the estimator for x_t is given as the solution to the equations

$$D_t = \Phi(\frac{-\lambda - \hat{x}_t}{\hat{\sigma}_t}), \tag{B.3}$$

$$1 - U_t = \Phi(\frac{\lambda - \hat{x}_t}{\hat{\sigma}_t}), \tag{B.4}$$

where $\Phi(.)$ is the standard normal cumulative distribution function. Using (B.3) and (B.4) to solve for $\hat{x}_t$ and $\hat{\sigma}_t$,

$$\hat{\sigma}_t = \frac{2\lambda}{\Phi^{-1}(1 - U_t) - \Phi^{-1}(D_t)}, \tag{B.5}$$

where $\Phi^{-1}(.)$ denotes the inverse standard normal cumulative distribution function. Thus,

$$\hat{x}_t = \lambda \left(\frac{\Phi^{-1}(1 - U_t) + \Phi^{-1}(D_t)}{\Phi^{-1}(1 - U_t) - \Phi^{-1}(D_t)} \right), \tag{B.6}$$

which leaves only λ undetermined. In the literature λ has been calculated in various ways. Carlson and Parkin assume unbiasedness over the sample period, $t = 1, ..., T$; that is, λ

is estimated as

$$\hat{\lambda} = \left(\sum\nolimits_{t=1}^{T} x_t\right) / \sum\nolimits_{t=1}^{T} \left(\frac{\Phi^{-1}(1-U_t) + \Phi^{-1}(D_t)}{\Phi^{-1}(1-U_t) - \Phi^{-1}(D_t)}\right). \tag{B.7}$$

Since λ is constant over time, its role is merely to scale $\hat{x}_t$.

B.2 The Regression Approach

Let aggregate output x_t be a weighted average of firms' output x_{it}, $(i = 1, ..., N_t)$,

$$x_t = \sum\nolimits_{i=1}^{N_t} w_i x_{it}, \tag{B.8}$$

where w_i is the weight assigned to firm i. Assuming (B.8) holds for the sample of firms under consideration, and categorising firms according to whether they reported an "up" or a "down", (B.8) can be rewritten as

$$x_t = \sum\nolimits_{i=1}^{N_t} w_i^+ x_{it}^+ + \sum\nolimits_{i=1}^{N_t} w_i^- x_{it}^- \tag{B.9}$$

where x_{it}^+ is x_{it} if firm i reports an "up" and 0 otherwise, likewise, x_{it}^- equals x_{it} if firm i reports a "down" and 0 otherwise and w_i^+ and w_i^- the associated weights. The survey does not provide exact quantitative information on x_{it}^+ and x_{it}^-. Following Anderson, if, up to a mean zero disturbance ξ_{it}, $x_{it}^+ = \alpha$ and $x_{it}^- = -\beta$, $\alpha, \beta > 0$, then

$$\begin{aligned} x_t &= \alpha \sum\nolimits_{i=1}^{N_t} w_i^+ - \beta \sum\nolimits_{i=1}^{N_t} w_i^- + \xi_t && \text{(B.10)} \\ &= \alpha U_t - \beta D_t + \xi_t, && \text{(B.11)} \end{aligned}$$

where $\xi_t = \sum_{i=1}^{N_t} w_i \xi_{it}$ and U_t and D_t are the (appropriately weighted) proportions of firms that reported an output rise and fall respectively. The unknown parameters α and β can be estimated *via* a linear (or non-linear) regression of x_t on U_t and D_t. The fitted values from this estimated regression then provide the quantified retrospective survey response estimator for x_t. To ensure the fitted values are unbiased estimates for x_t, an intercept is also included in the regression to allow for the possibility that ξ_t has a time-invariant non-zero mean. For periods of rising and variable changes in x_t, Pesaran extends this basic model to allow for an asymmetric relationship between x_t and x_{it}.

B.2.1 Relating the Regression Approach to the Probability Approach

Suppose that x_{it} is a random draw from a uniform density function $f(.)$ with mean x_t and range $2q$, $q > 0$; that is,

$$\begin{aligned} f(x) &= (2q)^{-1} \text{ if } x_t - q \leq x \leq x_t + q, \\ &= 0 \text{ otherwise,} \end{aligned} \tag{B.12}$$

with corresponding cumulative distribution function

$$\begin{aligned} F_t(x) &= (2q)^{-1}[x - (x_t - q)] \text{ if } x_t - q \leq x \leq x_t + q \\ &= 0 \text{ if } x < x_t - q \\ &= 1 \text{ if } x > x_t + q. \end{aligned} \tag{B.13}$$

From (B.2) and (B.1),

$$U_t = \frac{q + \hat{x}_t - \lambda}{2q}, \tag{B.14}$$

$$D_t = \frac{q - \hat{x}_t - \lambda}{2q}, \tag{B.15}$$

An estimate of output growth x_t may then be written as a function of the balance statistic; *viz.*

$$\hat{x}_t = q(U_t - D_t), \tag{B.16}$$

which provides an alternative justification for the use of the balance statistic.

A generalisation of (B.16) is obtained by relaxing the assumption that the "no change" interval is symmetric; that is, replace $(-\lambda, \lambda)$ by $(-a, b)$. Hence, (B.14) and (B.15) become

$$U_t = \frac{q + \hat{x}_t - b}{2q}, \tag{B.17}$$

$$D_t = \frac{q - \hat{x}_t - a}{2q}. \tag{B.18}$$

Then the estimator for x_t is

$$\hat{x}_t = \alpha U_t - \beta D_t, \tag{B.19}$$

which is equivalent to the estimator for x_t in (B.11) based on U_t and D_t for the single time period t, where the two scaling parameters are defined as

$$\alpha = \frac{2q(q-a)}{2q-a-b}, \ \beta = \frac{2q(q-b)}{2q-a-b}. \tag{B.20}$$

B.3 The Reverse-Regression Approach

Cunningham et al. (1998) and Mitchell et al. (2002) relate survey responses to official data by relating the proportions of firms reporting rises and falls to the official data. Under the assumption that (after revisions) official data offer unbiased estimates of the state of the economy this avoids biases caused by measurement error in the data.

Let the categorical survey response of firm i at time t be determined by the firm-specific unobserved continuous random variable y_{it}^* which is related to economy-wide manufacturing output growth x_t through the linear representation

$$y_{it} = x_t + \eta_{it} + \varepsilon_{it}. \tag{B.21}$$

which may be expressed in terms of (1) by defining $\eta_{it} = \alpha_i + (\beta_i - 1)x_t, (i = 1, ..., N_t,$ $t = 1, ..., T)$. In (B.21), η_{it} is the difference between y_{it} and x_t anticipated by firm i while ε_{it} is an unanticipated component, that is, $E(y_{it}|i) = x_{it} = x_t + \eta_{it}$.

The retrospective survey data provide firm level categorical information on the individual-specific random variable y_{it} *via* the discrete random variable y_{it}^j, $j = 1, 2, 3$, where

$$y_{it}^j = 1 \text{ if } c_{j-1} < y_{it} \leq c_j \text{ and 0 otherwise,} \tag{B.22}$$

where $c_0 = -\infty$ and $c_3 = \infty$, $j = 1, 2, 3$ with the intervals (c_0, c_1), (c_1, c_2) and (c_2, c_3) corresponding to "down", "same" and "up" respectively. Note that the thresholds c_j are invariant with respect to firm i and time t. Defined in terms of the error terms in (B.21), the observation rule (B.22) becomes

$$y_{it}^j = 1 \text{ if } c_{j-1} - x_t < \eta_{it} + \epsilon_{it} \leq c_j - x_t \text{ and 0 otherwise.} \tag{B.23}$$

A probabilistic foundation may be given to the observation rule (B.23) by letting the scaled error terms $\{\sigma(\eta_{it} + \epsilon_{it})\}$, $\sigma > 0$, possess a common and known cumulative distribution function $F(.)$, $i = 1, ..., N_t$, which is parameter free and assumed time-invariant. Then,

$$P(y_{it}^j = 1|x_t) = F(\mu_j - \sigma x_t) - F(\mu_{j-1} - \sigma x_t), \tag{B.24}$$

where $\mu_j = \sigma c_j$, $j = 1, 2, 3$.

B.3.1 Motivating the Regression Formulation

Let the survey proportion of firms that give response j at time t be denoted by $P_t^j = \sum_{i=1}^{N_t} y_{it}^j / N_t$, $j = 1, 2, 3$. As $P_{jt} = P(y_{it}^j = 1|x_t) = F(\mu_j - \sigma x_t) - F(\mu_{j-1} - \sigma x_t)$, $E(P_t^j|x_t) = P_{jt}$. If we further assume that $F(.)$ is symmetric, then $P_{1t} = F(\mu_1 - \sigma x_t)$ and $P_{3t} = F(-(\mu_2 - \sigma x_t))$. Hence, we may define the non-linear regressions

$$\begin{aligned} P_t^1 &= D_t = F(\mu_1 - \sigma x_t) + \xi_t^1, \\ P_t^3 &= U_t = F(-(\mu_2 - \sigma x_t)) + \xi_t^3. \end{aligned} \tag{B.25}$$

Assuming that the survey responses of firms are independent given x_t,

$$N_t^{1/2} \begin{pmatrix} \xi_t^1 \\ \xi_t^3 \end{pmatrix} \xrightarrow{d} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} F_t^1(1 - F_t^1) & -F_t^1 F_t^3 \\ -F_t^1 F_t^3 & F_t^3(1 - F_t^3) \end{pmatrix} \right), \tag{B.26}$$

where $F_t^1 = F(\mu_1 - \sigma x_t)$ and $F_t^3 = F(-(\mu_2 - \sigma x_t))$. Restricting attention to categories $j = 1$ and $j = 3$ only results in no loss of information since $\sum_{j=1}^3 P_t^j = 1$.

If $F(.)$ is strictly monotonic, the non-linear regressions (B.25) may be simplified by taking Taylor series approximations to $F^{-1}(D_t)$ and $F^{-1}(U_t)$ about $F(\mu_1 - \sigma x_t)$ and $F(-(\mu_2 - \sigma x_t))$ respectively yielding the *asymptotic* ($N_t \to \infty$) linear regression models

$$\begin{aligned} F^{-1}(D_t) &= \mu_1 - \sigma x_t + u_t^1, \\ F^{-1}(U_t) &= -\mu_2 + \sigma x_t + u_t^3, \end{aligned} \tag{B.27}$$

where

$$\begin{aligned} u_t^1 &= (f_t^1)^{-1} \xi_t^1 + o_p(N_t{}^{-1}), \\ u_t^3 &= (f_t^3)^{-1} \xi_{t,3} + o_p(N_t{}^{-1}), \end{aligned} \tag{B.28}$$

and $f_t^1 = f(\mu_1 - \sigma x_t)$, $f_t^3 = f(-(\mu_2 - \sigma x_t))$ and the density function $f(z) = dF(z)/dz$.

Since x_t is observed, feasible and asymptotically efficient estimation of (B.27) is achieved by generalised least squares (or minimum chi-squared) estimation given the structure of the variance-covariance matrix of u_t^1 and u_t^3.

B.3.2 Estimation of x_t

Estimates of the official (economy-wide) macroeconomic data x_t may be derived from the estimated regressions. Consider the inverse regression model (B.27) and let

$$\hat{x}_t^1 = \frac{\hat{\mu}_1 - F^{-1}(D_t)}{\hat{\sigma}},\ \hat{x}_t^3 = \frac{\hat{\mu}_2 + F^{-1}(U_t)}{\hat{\sigma}}. \tag{B.29}$$

where $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\sigma}$ denote the coefficient estimates. Both $\hat{x}_t^1$ and $\hat{x}_t^3$ are consistent estimators of x_t. A reconciled estimator for x_t is obtained using the variance-covariance matrix of $\hat{x}_t^1$ and $\hat{x}_t^3$ [see Cunningham et al. (1998) and Stone et al. (1942)]. Note that when there is a poor statistical relationship between the survey proportions and x_t, σ will be small and the implied indicator becomes very volatile; see (B.29).

B.3.3 Relating the Reverse-Regression Approach to the Probability Approach

Let $F_t(x) = F((x - x_t)/\sigma_t)$ with $F(.)$ symmetric. From (B.1) and (B.2) with an asymmetric interval for "same" $(-a, b)$, cf. (B.3) and (B.4), equate

$$1 - U_t = F(\frac{b - \hat{x}_t}{\hat{\sigma}_t}), \tag{B.30}$$

$$D_t = F(\frac{-a - \hat{x}_t}{\hat{\sigma}_t}). \tag{B.31}$$

From the symmetry of $F(.)$,

$$U_t = F(\frac{-b + \hat{x}_t}{\hat{\sigma}_t}). \tag{B.32}$$

Hence,

$$F^{-1}(U_t) = \frac{-b + \hat{x}_t}{\hat{\sigma}_t}, \tag{B.33}$$

$$F^{-1}(D_t) = \frac{-a - \hat{x}_t}{\hat{\sigma}_t}. \tag{B.34}$$

Therefore, in comparison with (B.27), $\mu_1 = -a/\sigma_t$, $\mu_2 = b/\sigma_t$ and $\sigma = 1/\sigma_t$.

References

Anderson, O. (1952), 'The business test of the IFO-Institute for Economic Research, Munich, and its theoretical model', *Review of the International Statistical Institute* **20**, 1–17.

Bai, J. & Ng, S. (2005), 'Tests for skewness, kurtosis and normality for time series data', *Journal of Business and Economic Statistics* **23**, 49–60.

Batchelor, R. A. (1981), 'Aggregate expectations under the stable laws', *Journal of Econometrics* **16**, 199–210.

Bates, J. M. & Granger, C. W. J. (1969), 'The combination of forecasts', *Operational Research Quarterly* **20**, 451–468.

Branch, W. A. (2004), 'The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations', *Economic Journal* **114**, 592–621.

Carlson, J. & Parkin, M. (1975), 'Inflation expectations', *Economica* **42**, 123–138.

Chesher, A. & Irish, M. (1987), 'Residual analysis in the grouped and censored normal linear model', *Journal of Econometrics* **34**, 33–61.

Cunningham, A. W. F. (1997), 'Quantifying survey data', *Bank of England Quarterly Bulletin* **August**, 292–300.

Cunningham, A. W. F., Smith, R. J. & Weale, M. R. (1998), Measurement errors and data estimation: the quantification of survey data, *in* I. Begg & S. G. B. Henry, eds, 'Applied Economics and Public Policy', Cambridge University Press, Cambridge.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2001), 'Coincident and leading indicators for the euro area', *Economic Journal* **111**(471), 62–85.

Granger, C. W. J. & Ramanathan, R. (1984), 'Improved methods of combining forecasts', *Journal of Forecasting* **3**, 197–204.

Horvath, B., Nerlove, M. & Wilson, D. (1992), A re-interpretation of direct tests of forecast rationality using business survey data, *in* K. H. Oppenlander & G. Poser, eds, 'Business Cycle Analysis by Means of Economic Surveys, Part 1', Avebury, Aldershot.

Keane, M. (1994), 'A computationally practical simulation estimator for panel data', *Econometrica* **62**, 95–116.

Machin, S. J. & Stewart, M. B. (1990), 'Unions and the financial performance of British private sector establishments', *Journal of Applied Econometrics* **5**, 327–350.

McIntosh, J., Schiantarelli, F. & Low, W. (1989), 'A qualitative response analysis of uk firms' employment and output decisions', *Journal of Applied Econometrics* **4**, 251–264.

Mitchell, J. (2002), 'The use of non-normal distributions in quantifying qualitative survey data on expectations', *Economics Letters* **76**, 101–107.

Mitchell, J., Smith, R. & Weale, M. (2002), 'Quantification of qualitative firm-level survey data', *Economic Journal* **112**, C117–C135.

Mitchell, J., Smith, R. & Weale, M. (2005*a*), Aggregate versus disaggregate survey-based indicators of economic activity. National Institute of Economic and Social Research Discussion Paper No. 194 (Revised).

Mitchell, J., Smith, R. & Weale, M. (2005*b*), 'Forecasting manufacturing output growth using firm-level survey data', *The Manchester School* **73**, 479–499.

Mitchell, J., Smith, R., Weale, M., Wright, S. & Salazar, E. (2005), 'An indicator of monthly GDP and an early estimate of quarterly GDP growth', *Economic Journal* **115**, F108–F129.

Murphy, A. (1996), 'Simple LM tests of mis-specification for ordered logit models', *Economics Letters* **52**, 137–141.

Nerlove, M. (1983), 'Expectations, plans, and realizations in theory and practice', *Econometrica* **51**, 1251–1279.

Newey, W. K. (1987), 'Efficient estimation of limited dependent variable models with endogenous explanatory variables', *Journal of Econometrics* **36**, 231–250.

Pesaran, M. H. (1984), Expectations formation and macroeconomic modelling, *in* P. Malgrange & P. Muet, eds, 'Contemporary Macroeconomic Modelling', Blackwell, Oxford, pp. 27–55.

Pesaran, M. H. (1987), *The limits to rational expectations*, Basil Blackwel, Oxford.

Pesaran, M. H. & Weale, M. (2005), Survey expectations, *in* G. Elliott, C. W. J. Granger & A. Timmermann, eds, 'Handbook of Economic Forecasting', North-Holland, North Holland, p. Forthcoming.

Smith, R. J. & Blundell, R. W. (1986), 'An exogeneity test for a simultaneous equation Tobit model with an application to labor supply', *Econometrica* **54**, 679–685.

Souleles, N. S. (2004), 'Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the Michigan consumer sentiment surveys', *Journal of Money, Credit and Banking* **36**, 39–72.

Stock, J. H. & Watson, M. W. (2003), Understanding changes in international business cycle dynamics. NBER Working Paper No. 9859.

Stock, J. & Watson, M. (2002), 'Macroeconomic forecasting using diffusion indexes', *Journal of Business and Economic Statistics* **20**(2), 147–162.

Stone, J. R. N. (1947), 'On the interdependence of blocks of transactions', *Supplement to The Journal of the Royal Statistical Society* **9**, 1–32.

Stone, J. R. N., Champernowne, D. A. & Meade, J. E. (1942), 'The precision of national income estimates', *Review of Economic Studies* **9**, 111–25.

Theil, H. (1952), 'On the time shape of economic microvariables and the Munich business test', *Revue de l'Institute International de Statistique* **20**, 105–120.