Aggregate versus Disaggregate Survey-Based Indicators of Economic Activity*

James Mitchell
NIESR

Richard J. Smith
NIESR and University of Warwick

Martin R.Weale
NIESR

This Revision: November 2004

Abstract

Qualitative 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 alternative *disaggregate* indicators 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 disaggregate indicators of manufacturing output growth provide more accurate early estimates of manufacturing output growth than traditional aggregate indicators.

Keywords: Survey data, Indicators, Quantification, Forecasting

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.

*Address for correspondence: James Mitchell, National Institute of Economic and Social Research, 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. We are grateful to an Editor of this Journal and two anonymous referees for their constructive comments and criticisms of an earlier draft. We would also like to thank G. Camba-Mendez and K. Lee for their helpful comments and J. Morrison and J. Wood of the CBI for their advice about the survey data. Financial support from the ESRC is also gratefully acknowledged. The Leverhulme Trust under a 2002 Major Research Fellowship also provided partial financial support for the second author. The usual disclaimer applies.

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 the paper by 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 and 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 *et al.*, 2004] 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. 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; see Mourougane and Roma (2002). 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 and Parkin, 1975], the regression method [Pesaran, 1984] and the reverse regression method [Cunningham *et al.*, 1998]. For an example of an application using Carlson and Parkin's approach, see Lee (1994). The regression approach is used, for example, by the Bank of England [see Britton *et al.*, 1999]. Smith and McAleer (1995) and Appendix A below compare these approaches. 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. How should the responses of the individual firms be combined if the aim of the survey is to produce an early indication of official output data? 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 "disaggregate" indicator built round ordered discrete choice models linking individual firms' categorical responses to official data. Then, given the categorical responses, we infer the most likely values for the official data. In contrast to Kaiser and Spitz (2000), who use a "pooled" ordered discrete choice model, our approach does not impose homogeneity among firms, an assumption we reject below. An alternative "semi-disaggregate" indicator based on grouping the firms according to their responses both at time t and $t-1$ is proposed in Mitchell *et al.* (2002a). We illustrate the use of the disaggregate indicators in an application to industrial survey data from the Confederation of British Industry. We find our disaggregate indicators of manufacturing output growth explain more of the variation in the outturn for output growth than traditional indicators constructed using "aggregate" data. They also offer more accurate estimates on an out-of-sample basis.

The plan of this paper is as follows. Section 2 motivates the disaggregate indicator. Section 3 considers an application, and section 4 makes some concluding comments.

2 Firm-Level Quantification

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 on x_t. See section 2.2 below 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 (A.21) in Appendix A.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 [see Amemiya 1985, Ch.9] based on the latent regression (1). Define the indicator variables

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

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.

2.2 Specification Tests

It is important to test the implied restrictions embodied in (1), particularly given that macroeconomic data are widely accepted to exhibit dependence. 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

$$x_t = \alpha_x + \beta_x x_{t-1} + u_t, \tag{5}$$

$(t = 1, ..., T)$, 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 (5) 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

$$\varepsilon_{it} = \rho_i u_t + \xi_{it}, \tag{6}$$

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 (6) in (1) generates the conditional dynamic model

$$\begin{aligned} y_{it} &= \alpha_i + \beta_i x_t + \rho_i u_t + \xi_{it} \\ &= \alpha_i^* + \beta_{i0}^* x_t + \beta_{i1}^* x_{t-1} + \xi_{it}, \end{aligned} \tag{7}$$

$(i = 1, ..., N_t)$, 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 (7) 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 and Blundell (1986) and Newey (1987). Firstly, (5) 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 (7) is estimated by ordered logit as in section 2.1 after substitution of $\hat{u}_t$ for u_t. 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 (7); see section 2.3 below. 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.

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 and Irish (1987), Machin and 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 simulated out-of-sample experiments to assess the forecasting performance of the selected models against benchmark forecasts in section 3.3.

2.3 Inferring the Official Data

Given an ordered logit model for each firm i, an estimator for x_t may be inferred from the 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. Our initial interest centres on the conditional density $f(x_t|j, i)$ for observing x_t given the survey response j for firm i.

Let $f(x_t)$ denote the time-invariant density function of x_t. Therefore, the conditional probability of observing response j for firm i is $P(j|i) = \int_{-\infty}^{\infty} P(j|x_t, i)f(x_t)dx_t$. Bayes' Theorem states that

$$f(x_t|j, i) = \frac{P(j|x_t, i)f(x_t)}{P(j|i)}. \tag{8}$$

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

$$E(x_t|j, i) = \int_{-\infty}^{\infty} x_t f(x_t|j, i)dx_t, \tag{9}$$

which takes one of three values depending on the observed sample response j_{it} of firm i at time t. Given $f(x_t)$, all of the above integrals may be calculated by numerical evaluation.

Estimators $\hat{P}(j|x_t, i)$ for $P(j|x_t, i)$ and, thus, $\hat{P}(j|i)$ for $P(j|i)$ 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 $\hat{E}(x_t|j, i)$ may be obtained from (9) by numerical evaluation.

To create a disaggregate indicator D_t of economic activity at time t, from the law of iterated expectations the conditional expectation of x_t given all firms' survey responses j_{it}, $(i = 1, ..., N_t)$,

$$E(x_t|\{j_{it}\}_{i=1}^{N_t}) = \sum\nolimits_{i=1}^{N_t} H_{it}E(x_t|j_{it}, i), \tag{10}$$

where H_{it} is the exogenous sample probability of observing firm i at time t. Hence, assuming firms are independent, we define the parametric indicator

$$D_t = \sum\nolimits_{i=1}^{N_t} w_{it}\hat{E}(x_t|j_{it}, i), \tag{11}$$

where $w_{it} > 0$ is the weight given to firm i at time t and $\sum_{i=1}^{N_t} w_{it} = 1$. If firms constitute a random sample, then equal weights are appropriate since all firms are equally likely in the sample. However, if firms are drawn according to some stratified sampling process, then the weights w_{it} should reflect stratum weights; for example, if strata are defined by firm size, then firms should be size-weighted.

An alternative non-parametric disaggregate indicator ND_t for the conditional expectation $E(x_t|\{j_{it}\}_{i=1}^{N_t})$ which avoids the assumption of a parametric structure for $f(x_t|j, i)$ *via* (8) may be based on the conditional empirical distribution function. Define the indicator function $1(x_t \leq x, j_{it} = j|i) = 1$ if $x_t \leq x$ and $j_{it} = j$ and 0 otherwise, $(j = 1, 2, 3)$. Let $T_i^j = \sum_{s=1}^{T} y_{is}^j$ which is the number of times firm i gives response j in the survey; hence, T_i^j/T is the sample proportion of responses j for firm i, $(j = 1, 2, 3)$. The conditional empirical distribution function of x_t given reponse j for firm i is given by $\hat{F}(x|j, i) = \sum_{t=1}^{T} 1(x_t \leq x, j_{it} = j|i)/T_i^j$, $(j = 1, 2, 3)$, which assigns equal weight to each sample value. As $T \to \infty$ and, thus, $T_i^j \to \infty$, $T_i^j/T \xrightarrow{p} P(j|i)$ and $\sum_{t=1}^{T} 1(x_t \leq x, j_{it} = j|i)/T \xrightarrow{p} F(x, j|i)$ if, given firm i, x_t and j_{it} may be regarded as stationary random variables with joint conditional c.d.f. $F(x, j|i)$. Hence, $\hat{F}(x|j, i) \xrightarrow{p} F(x|j, i) = F(x, j|i)/P(j|i)$, the conditional c.d.f. of x_t given response j

and firm i. Therefore, the mean of $\hat{F}(x|j,i)$, $\sum_{s=1}^{T} y_{is}^{j} x_s / T_i^j$, is a consistent estimator for $E(x_t|j,i)$. A nonparametric disaggregate (ND_t) indicator, a discrete version of (11), is therefore defined as

$$ND_t = \sum\nolimits_{i=1}^{N_t} w_{it} \sum_{s=1}^{T} y_{is}^{j_{it}} x_s / T_i^{j_{it}}. \tag{12}$$

3 An Application: CBI Survey Data

The *Industrial Trends Survey* (ITS) of the Confederation of British Industry (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 $1997q3$ (37 quarters). 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 time-series observations across all firms for the estimation of (1) without assuming some homogeneity in behaviour across firms. Firm-level quantification requires sufficient time-series observations for a given firm for reliable parameter estimation. Given the static specification (1), however, observations need not be consecutive.

In the absence of rules guiding the choice of how many time-series observations are necessary, we accordingly take an eclectic approach and consider a range of values for the minimum number of observations for the inclusion of a particular firm when examining the performance of the disaggregate indicators. In the parametric approach we consider a range from 7 to 37 observations as estimation is unreliable if not infeasible for less than 7 observations. For the purposes of illustration, since consistency of the estimators for α_i, β_i and μ_{ji} in the firm-level models is predicated on $T \to \infty$, rather than examination of the performance of the parametric disaggregate indicator irrespective of its theoretical properties, when subjecting the firm-level models to specification and hypothesis tests, we focus on those 643 firms that reply to at least 20 surveys. Although this choice is arbitrary, similar results were obtained for other values although the power of these tests should be expected to decline as fewer observations are considered per firm. For the nonparametric approach the observational range 1 to 37 is considered.

Exclusion of firms that do not respond to at least a given number of surveys may induce a selection bias in the disaggregate indicators. The resultant "included sample" obtained after dropping firms with less than a given number of time-series observations may no longer be considered as a random sample; firms may fail to respond to surveys for self-selection reasons. Consequently, inference may be biased; no bias results only if firms omitted from the included sample, the "excluded sample", may be regarded as *missing at random* or *ignorable* [see Griliches, 1986]. In particular, indicators or statistics derived from both included and excluded samples should not differ significantly. To test for the possibility of selection bias we tested the correlation of three traditional aggregate indicators, reviewed in Appendix A, with the outturn for output growth. In all cases, there was no evidence of a statistically significant difference between the performance of these aggregate indicators in the included and excluded samples; for further details, see Mitchell *et al.* (2002b). This result is consistent with the view that the included sample may be regarded as a random sample, and that inference from it should be unbiased. Therefore if disaggregate indicators outperform traditional aggregate indicators we can conclude that this improvement is due to disaggregation *per se*, and is not a consequence of sample selection. We also used forecast encompassing tests to examine whether the aggregate indicators derived from excluded firms add information *vis-à-vis* the disaggregate indicators. Again, there was little evidence to suggest that dropping firms led to any informational loss.

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; see Hsiao (2003). 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 and 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.

To give an impression of the nature of the survey responses, Figure 1 plots the percentage of the 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.

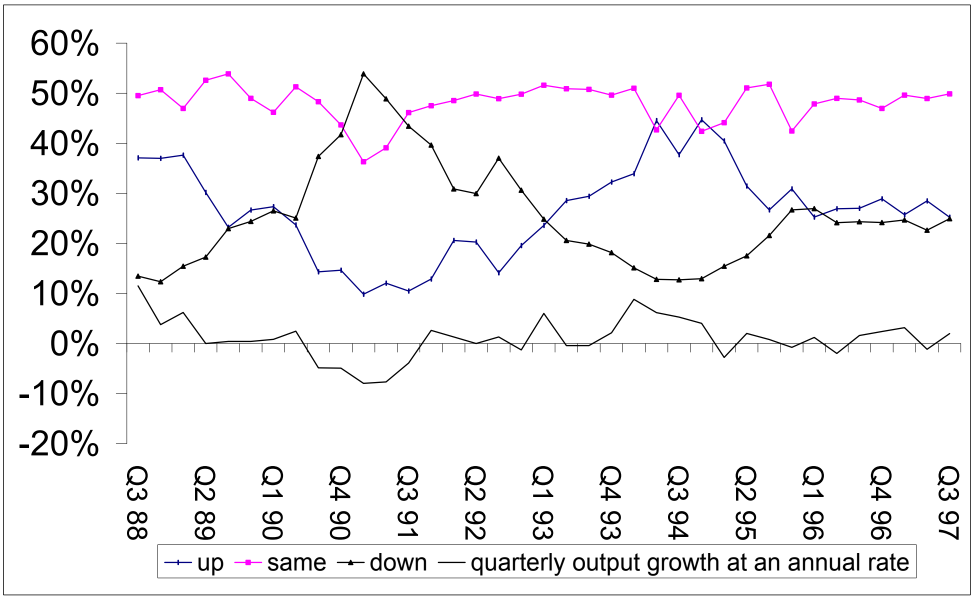

Figure 1: Unweighted percentage of firms reporting "up", "same" or "down"

3.1 Disaggregate indicators

3.1.1 The Density Function $f(.)$

The parametric disaggregate indicator D_t for manufacturing output growth x_t requires a time-invariant density function for x_t, $f(.)$, to be specified; see (8). The assumption that x_t is stationary is supported by tests for a unit root in the level series of manufacturing output. The sample observations for x_t have sample mean and sample standard deviation equal to 1.023 and 4.057 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 and Ng, 2003] does not reject the normality of $f(.)$ with a p-value of 0.790. Specifically the Bai-Ng test for normality employs a long run covariance matrix estimate obtained using the kernel method with prewhitening suggested by Andrews and Monahan (1992). Results were robust to using alternative estimators of the long run covariance matrix, such as the method of den Haan and Levin (2000) that uses a data-dependent prewhitening procedure. We also experimented with the Pearson family of density functions, but the performance of the disaggregate indicator was not affected; see Mitchell *et al.* (2002b) for further details.

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

The parametric disaggregate indicator 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 section 2.2. Both the test of $\rho_i = 0$, $(i = 1, ..., N_t)$, 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-corrected 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

3.1.3 Comparing the Performance of the Aggregate and Disaggregate Indicators

We compare the performance of the disaggregate indicators against that of three traditional quantification techniques employed on aggregate proportions: the probability method of Carlson and Parkin (1975) [CP], the regression approach of Pesaran (1984) [P] and the reverse-regression approach of Cunningham *et al.* (1998) [CSW] based on the logistic distribution. See Appendix A for a review.

The CP aggregate indicator is identified up to a scaling parameter [see Appendix A]. Following CP we chose this parameter to ensure that the mean of the indicator is 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. Note that the scale of the disaggregate indicators is identified directly by x_t, in the parametric case through the form taken by $f(x_t)$ in (8).

To evaluate fairly the performance of the alternative aggregate (and disaggregate) indicators, mean squared error (MSE) criteria are inappropriate as MSE depends on scale. We therefore summarise the performance of indicators of manufacturing output growth by examining their estimated correlation coefficients (corr.) with the outturn for manufacturing output growth. Correlation informs us about the informational content of the indicator series, a high value indicating that a strong signal about the outturn may be recovered from the indicator regardless of how the indicator has been scaled, and whether MSE is high or low. We considered both unweighted and weighted aggregate and disaggregate indicators; the weights, based on firms' sales volumes, are those used by the CBI in aggregating firms' responses. Weighting the proportions in the aggregate case, or the firm-level quantified series in the disaggregate case, unambiguously leads to worse indicators. However, this may not be true for all possible weighting schemes. We focus here on the unweighted estimates, but do present some representative results for the weighted estimates.

Figure 2 plots the correlation of the disaggregate indicators against manufacturing output growth as a function of the minimum number of observations considered per firm. It also shows how many firms are in a given sample. Only 22 firms reply to all 37 surveys. Regardless of how many observations per firm are considered, while both aggregate and disaggregate indicators are positively correlated with manufacturing output growth, the disaggregate indicators provide more accurate early estimates of output growth than traditional aggregate indicators. Also, the nonparametric disaggregate indicator ND_t appears to be better than its parametric counterpart D_t. An assessment of the performance of the aggregate and disaggregate indicators on a sectoral basis was also made. Grouping firms into seven industrial sectors based on their 1980 Standard Industrial Classification (SIC) code, we found that: the disaggregate indicators always explain more variation in sectoral output growth than the aggregate indicators; the nonparametric disaggregate indicator is always superior to the parametric disaggregate indicator; the sectoral indicators do not explain as much of the variation in the outturn as the indicators for (overall) manufacturing output examined above reflecting increased volatility at the sectoral level. See Mitchell *et al.* (2002b) for further details.

Figure 2 also shows that the correlation of the disaggregate indicators with the outturn declines as more observations per firm are considered. The fewer firms are dropped the better the fit of the disaggregate indicator. This is, at least in part, a consequence of over-fitting. A simple example illustrates this. If there are just T firms in the sample and they each reply once but at different points in time, then the nonparametric disaggregate indicator will fit the official data perfectly. Given this danger, it is therefore important to consider the behaviour of the indicators on an out-of-sample basis; see section 3.3.

We found that despite the sample mean of the disaggregate indicators approximately estimating that of the outcomes x_t correctly, all the disaggregate indicators appear too smooth; for an example focussing on the behaviour of the disaggregate indicators with 20 observations see Table 3. For comparative purposes Table 3 also presents results for the aggregate indicators and examines the performance of the weighted estimates. Table 3 shows that the disaggregate indicators display too little volatility as compared with the outturn. 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 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 disaggregate method gives the same score (mean output growth) to all categorical responses; i.e. since $P(j|x_t, i) = P(j|i)$, $E(x_t|j, i) = m_1'$, where m_1' is the sample mean of $\{x_t\}$; see (9). For random firms the contribution to the aggregate therefore has no variance. Excess smoothness of the disaggregate indicators may thus be viewed as due to the presence of firms in the sample whose responses contain little or no signal about output growth. To reconcile this incompatibility in volatilities, note that the outturn is the signal recovered from the survey data plus a residual error component. We therefore re-scale the indicators through regression on the outturn as a simple method of obtaining an indicator which tracks output growth more closely. The effects of this regression are taken into account in our subsequent out-of-sample analysis. An alternative approach might be to identify those firms whose responses contained little or no signal and then to exclude these firms, or give them a lower weight, when defining the disaggregate indicator.

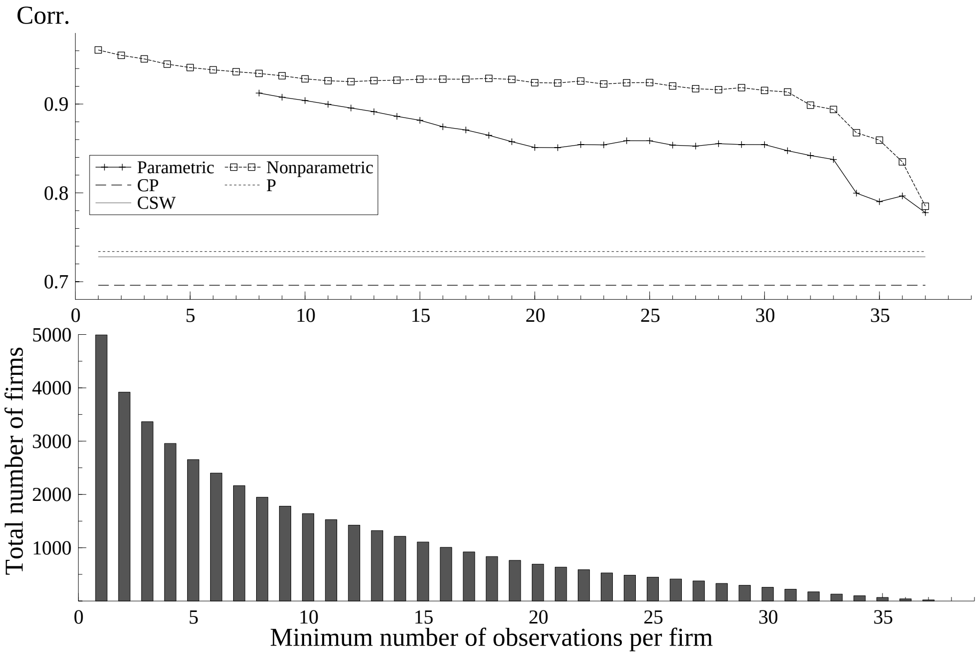

Figure 2: In-sample performance of the aggregate and disaggregate indicators. Correlation of the indicators against manufacturing output growth as a function of the minimum number of observations considered per firm

3.2 A Graphical Representation of the Disaggregate and Aggregate Indicators

To illustrate the superiority of disaggregate over aggregate indicators we consider the most informative aggregate and disaggregate indicators, that is, the unweighted P aggregate indicator and the unweighted re-scaled nonparametric disaggregate indicator ND_t. Figure 3 plots these two indicators together with the outturn for manufacturing output growth.

Figure 3 indicates that the disaggregate indicator picks up the dip of 1994 much better than the aggregate indicator. Furthermore, despite the poorer performance of both indicators in the late 1990s the disaggregate indicator is again closer to the outturn.

3.3 Out-Of-Sample Analysis

Given the improved in-sample fit between the survey responses and official data using disaggregate rather than aggregate indicators, we examine whether the superiority of the disaggregate indicators extends out-of-sample. To evaluate how accurate survey-based early estimates of output growth would have been out-of-sample an experiment

Table 3: Aggregate and Disaggregate Indicator Performance

		Mean	Stand. Dev.	Corr.
Outturn for Manuf. Output Growth		1.023	4.057	
CP	unweighted	1.023	175.748	0.696
	weighted	1.023	88.501	0.648
P	unweighted	1.023	2.978	0.734
	weighted	1.023	2.752	0.678
CSW	unweighted	1.023	5.574	0.728
	weighted	1.023	6.244	0.650
D_t	unweighted	1.031	0.427	0.851
	weighted	1.054	0.481	0.810
ND_t	unweighted	1.009	0.640	0.923
	weighted	1.008	0.685	0.905

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 that meant it became very difficult to match firms pre- and post-December 1999. Results presented below for the disaggregate indicators focus on those firms who had given at least one survey response during the in-sample period ending in $1997q3$. Similar results were obtained when "new" firms were allowed to enter the sample during the out-of-sample period.

The out-of-sample analysis involves computing the aggregate and disaggregate 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 $1998q1$. This recursive process is repeated until finally output growth in $1999q3$ is inferred using survey and official data $1988q3 - 1999q2$, plus survey data in $1999q3$. Both aggregate and disaggregate out-of-sample estimates are re-scaled by recursively regressing their in-sample counterparts against the outturn for output growth. In this way no *ex post* information about output growth is used when quantifying the survey data in real-time. As is standard when evaluating forecasts, the performance of the indicators is evaluated in terms of their root MSE against the outturn.

The results of this recursive exercise are summarised in Figure 4. Figure 4 plots the root MSE of the aggregate (computed using the full sample) and the disaggregate indicators, both unweighted, against the outturn for manufacturing output growth. The performance of the disaggregate indicators is evaluated as a function of the minimum

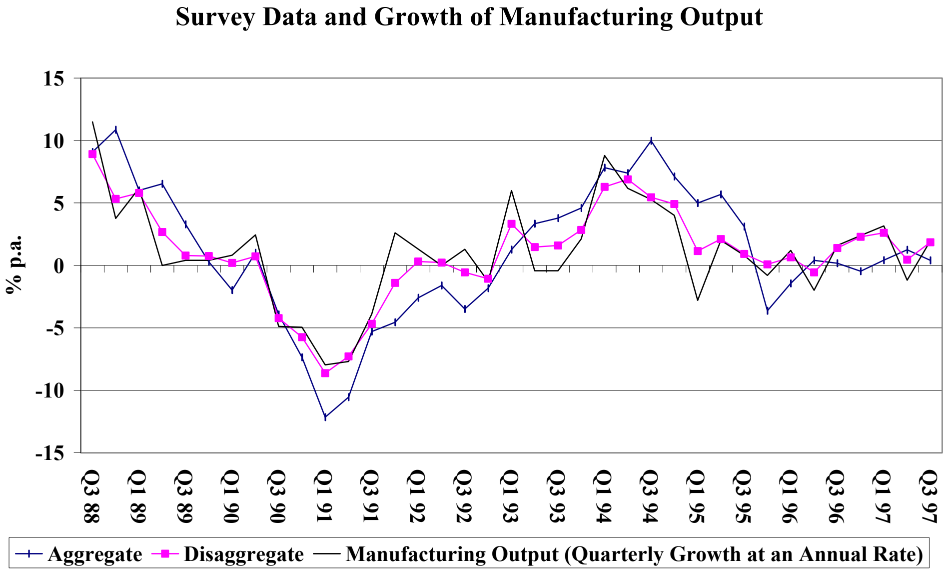

Figure 3: Explanatory power of disaggregate and aggregate indicators

number of observations considered per firm.

Firstly, both the parametric and nonparametric disaggregate indicators produce more accurate forecasts than the aggregate indicators. Moreover, there is evidence to suggest that these improvements are statistically significant using small sample corrected Diebold-Mariano tests; see Harvey *et al.* (1997). This finding is encouraging and further motivates the use of disaggregate survey-based indicators. Secondly, the performance of the parametric (nonparametric) disaggregate indicator is best when 25 (7) observations are considered per firm. However, out-of-sample analysis, particularly with small samples, is always sensitive to the sample period chosen. This is particularly so in this application where output growth is far less volatile in the out-of-sample than in-sample period; output growth has a standard deviation of 4.1% for the period $1988q3 - 1997q3$ as compared with 2.6% for $1997q4 - 1999q3$. Experimentation with a 16 rather than 8 quarter out-of-sample period did, however, deliver similar results.

4 Concluding Comments

Using a panel of firm level survey responses obtained from the CBI disaggregate indicators for output growth are derived using ordered discrete choice models relating firms' categorical survey responses to a quantitative measure of economic activity. Considerable

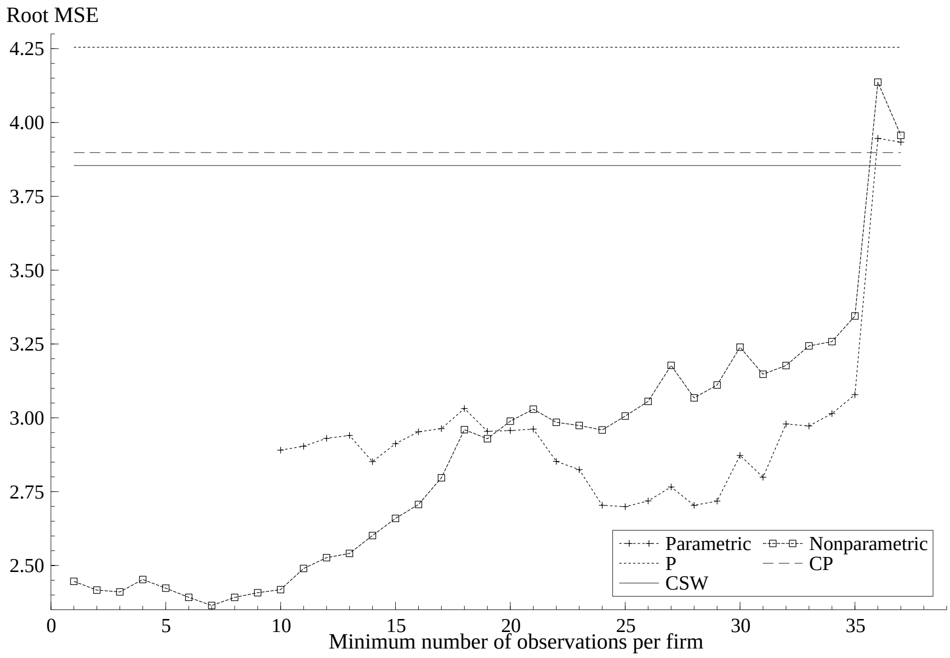

Figure 4: Out-of-sample performance of the aggregate and disaggregate indicators. Root MSE of the indicators against manufacturing output growth as a function of the minimum number of observations considered per firm

heterogeneity across firms is present concerning how their responses relate to the measure of economic activity. The disaggregate indicators outperformed traditional aggregate indicators in terms of anticipating movements in both manufacturing output growth (and sectoral output growth) and both in-sample and out-of-sample.

References

[1] Amemiya, T. (1985), *Advanced Econometrics*, Basil Blackwell.

[2] 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.

[3] Andrews, D.W.K. and J. Monahan (1992), "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator", *Econometrica*, 60, 953-966.

[4] Bai, J. and S. Ng (2003), "Tests for skewness, kurtosis and normality for time series data", *Journal of Business and Economic Statistics*. Forthcoming.

[5] Batchelor, R. (1981), "Aggregate expectation under the stable laws", *Journal of Econometrics*, 16, 199-210.

[6] Batchelor, R. (1982), "Expectations, output and inflation", *European Economic Review, 17, 1-25.*

[7] Britton, E., J. Cutler and A. Wardlow (1999), "The Bank's use of survey data", *Bank of England Quarterly Bulletin*, May, 177-182.

[8] Carlson, J.A. and M. Parkin (1975), "Inflation expectations", *Economica*, 42, 123-138.

[9] Chesher, A. and M. Irish (1987), "Residual analysis in the grouped and censored normal linear model", *Journal of Econometrics*, 34, 33-61.

[10] Cunningham, A.W.F. (1997), "Quantifying survey data", *Bank of England Quarterly Bulletin*, August, 292-300.

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

[12] den Haan, W.J. and A.T. Levin (2000), "Robust covariance matrix estimation with data-dependent VAR prewhitening order", *NBER Technical Working Paper 255.*

[13] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2001), "Coincident and Leading Indicators for the Euro Area", *Economic Journal*, 111, 62-85.

[14] Griliches, Z. (1986), "Economic data issues". In *Handbook of Econometrics, Volume III* (ed. Griliches, Z. and M.D. Intriligator), Elsevier.

[15] Harvey, D. I. and S. J. Leybourne and P. Newbold (1997), "Testing the Equality of Prediction Mean Square Errors", *International Journal of Forecasting*, 13, 273–281.

[16] Hsiao, C. (2003), *Analysis of Panel Data* (Second edition), Cambridge University Press.

[17] Kaiser, U. and A. Spitz (2000), "Quantification of qualitative data using ordered probit models with an application to a business survey in the German service sector", paper presented at the 25th CIRET conference, Paris, 10-14th October 2000.

[18] Keane, M.P. (1994), "A Computationally Practical Simulation Estimator for Panel Data", *Econometrica*, 62, 95-116.

[19] Lee, K.C. (1994), "Formation of price and cost inflation expectations in British manufacturing industries: a multisectoral analysis", *Economic Journal*, 104, 372-386.

[20] Machin, S.J. and M.B. Stewart (1990), "Unions and the financial performance of British private sector establishments", *Journal of Applied Econometrics*, 5, 327-350.

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

[22] Mitchell, J., R.J. Smith and M.R. Weale (2002a), "Quantification of qualitative firm-level survey data", *Economic Journal*, 112, C117-C135.

[23] Mitchell, J., R.J. Smith and M.R. Weale (2002b), "Aggregate versus Disaggregate Survey-Based Indicators of Economic Activity", National Institute of Economic and Social Research Discussion Paper No. 194.

[24] Mitchell, J., R.J. Smith, M.R. Weale, S. Wright and E.L. Salazar (2004), "An Indicator of Monthly GDP and an Early Estimate of Quarterly GDP Growth", *Economic Journal*. Forthcoming.

[25] Mourougane, A. and M. Roma (2002), "Can confidence indicators be useful to predict short term real GDP growth?", European Central Bank Working Paper No. 133.

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

[27] Newey, W.K. (1987), "Efficient estimation of limited dependent variable models with endogenous explanatory variables", *Journal of Econometrics*, 36, 231-250.

[28] Pesaran, M.H. (1984), "Expectations formation and macroeconomic modelling". In *Contemporary Macroeconomic Modelling* (eds. Malgrange, P. and P. Muet), Blackwell.

[29] Pesaran, M.H. (1987), "The limits to rational expectations", Basil Blackwell.

[30] Smith, J. and M. McAleer (1995), "Alternative procedures for converting qualitative response data to quantitative expectations: an application to Australian manufacturing", *Journal of Applied Econometrics*, 10, 165-185.

[31] Smith, R.J. and R.W. Blundell (1986), “An exogeneity test for a simultaneous equation Tobit model with an application to labor supply”, *Econometrica*, 54, 679-685.

[32] Stock, J.H. and M.W. Watson (2002), “Macroeconomic forecasting using diffusion indexes”, *Journal of Business and Economic Statistics*, 20, 147–162.

[33] Stone, J.R.N. (1947), “On the interdependence of blocks of transactions”, *Supplement to The Journal of the Royal Statistical Society*, 9(1), 1-32.

[34] Stone, J.R.N., D.A. Champernowne and J.E. Meade (1942), “The precision of national income estimates”, *Review of Economic Studies*, 9, 111-25.

[35] 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.

[36] Wren-Lewis, S. (1985), “The quantification of survey data on expectations”, *National Institute Economic Review*, 113, 39-49.

Appendix A: 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 three alternative methods of quantification are reviewed:

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

Although motivated in different ways, the three 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.* (2002a). For alternative reviews and extensions of the probability and regression approaches, see Wren-Lewis (1985) and Smith and McAleer (1995).

A.1 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, $x_{it} = \int xf(x|i)dx$ 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

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

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

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 $F_t(.)$.

A.1.1 Carlson and Parkin's Method

The traditional approach of Carlson and 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 (A.1) and (A.2), the estimator for x_t is given as the solution to the equations

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

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

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

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

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

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

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

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

For alternative approaches, see *inter alia* Batchelor (1981, 1982), Pesaran (1984), and Wren-Lewis (1985). Since λ is constant over time, its rôle is merely to scale $\hat{x}_t$.

A.2 The Regression Approach

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

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

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

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

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

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

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}.

A.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,

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

with corresponding cumulative distribution function

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

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

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

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

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

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

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

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

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

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

Then the estimator for x_t is

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

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

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

A.3 The Reverse-Regression Approach

Cunningham *et al.* (1998) and Mitchell *et al.* (2002a) 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

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

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 (A.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^j_{it}, $j = 1, 2, 3$, where

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

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 (A.21), the observation rule (A.22) becomes

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

A probabilistic foundation may be given to the observation rule (A.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,

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

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

A.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

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

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

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

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 (A.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

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

where

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

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 (A.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.

A.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 (A.27) and let

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

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 (A.29).

A.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 (A.1) and (A.2) with an asymmetric interval for "same" $(-a, b)$, cf. (A.3) and (A.4), equate

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

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

From the symmetry of $F(.)$,

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

Hence,

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

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

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