Forecasting Manufacturing Output Growth Using Firm-Level Survey Data*

James Mitchell
NIESR

Richard J. Smith
University of Warwick and NIESR

Martin R. Weale
NIESR

This revision: January 2005

Abstract

Traditionally forecasts of macroeconomic aggregates are extracted from prospective qualitative survey data by relating official data on the aggregate to both the proportion of survey respondents who are "optimists" and the proportion who are "pessimists". But there is no reason to focus on these proportions to the exclusion of other possible means of aggregating and quantifying the underlying panel of respondent or firm-level survey responses. Accordingly in this paper we show how the panel of firm-level responses underlying these proportions can be exploited to derive forecasts of (aggregate) manufacturing output growth that do not lose information that may be contained in the pattern of individual responses. An application using firm-level prospective survey data from the Confederation of British Industry shows that the forecasts of manufacturing output growth derived using these "disaggregate" methods mark an improvement over the so-called "aggregate" methods based on use of the proportions data alone.

Keywords: Survey data; Forecasting; Quantification; Panel data

*Address for correspondence: 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; r.j.smith@warwick.ac.uk and m.weale@niesr.ac.uk. Financial support from the ESRC and DG-ECFIN is gratefully acknowledged. Smith also gratefully acknowledges financial support of his research from a 2002 Leverhulme Major Research Fellowship.

1 Introduction

The use of qualitative survey data as a complement to official macroeconomic data continues to be popular. Qualitative surveys usually ask not only about experiences in the recent past but also about prospects for the near future; they may therefore be helpful in providing a guide to short-term prospects in a way that official data cannot. In this paper we discuss a means of producing a forecast of official data using the disaggregated responses to a qualitative survey.

This type of survey typically involves ordered responses, namely respondents answer "up", "the same" or "down" to a range of questions including one about their expectations of their own output growth. The data are traditionally available in aggregate form rather than at the individual level. Published survey data typically report only the proportion of respondents who answered "up", "the same" or "down", or the balance of opinion ("up"'s minus "down"'s or the proportion of "optimists" less the proportion of "pessimists").[1] There have been numerous studies of the way in which these survey responses link to and anticipate official data for both output and price movements. There are two main, what we call "aggregate", approaches for linking these aggregate survey data (the proportions) to official data and deriving quantitative estimates of actual and expected output or price movements, the *probability method* of Carlson and Parkin (1975), and the *regression method* of Pesaran (1984, 1987).

A number of authors have looked at the performance of these methods and derivatives of them as ways of using questions about expected future output movements in individual firms to predict output growth. Entorf (1993) finds, however, looking at the question in the IFO survey about expected future business conditions (rather than the expected output of the respondent itself) that the proportion of respondents expecting business conditions to worsen is a better predictor of future output changes than is the balance statistic. Cunningham, Smith and Weale (1998) examining surveys from the United Kingdom also find that use of the balance statistic results in a loss of information. Smith and McAleer (1995) and Driver and Urga (2004) compare various approaches over long periods.

There have been a number of other studies looking at the performance of these prospective measures of economic performance, often published by the bodies which produce the indicators themselves, but in most cases they do not go beyond the question whether the indicators have some capacity to fit the data. Madsen (1993) studies the predictive power of production expectations in eight OECD countries. Hild (2002) uses the method of principal components to explore the inter-relationships between variables in the French survey while Bouton and Erkel-Rousse (2002) look at the information contained in qualitative data on the service sector for France. Gregoir and Lenglart (2000) use the survey to derive a co-incident indicator based on a two-state Markov process. Parigi and Schlitzer (1995) consider forecasts of the Italian business cycle. Pesaran and Weale (2004) offer a survey of the use of both qualitative and quantitative data in expectations measurement.

However, in these and similar studies interpretation and analysis of such data is usu-

[1] There is often a small number of "don't knows" which are usually ignored in the analysis of such surveys.

ally based on aggregation of individual responses in a way which may lose information contained in the pattern of individual responses and does not exploit the panel aspect of the data. Little attention to date has been paid to whether better signals of economic behaviour can be derived from analysis of the panel data set of individual or firm-level responses underlying the aggregate responses, or proportions.[2] One exception is Mitchell, Smith and Weale (2004) who consider how the retrospective survey responses of individual firms can be combined if the aim is to produce an early indication of official output data, based on the fact that survey data are published ahead of official data on output growth. They find that more accurate indicators are obtained when quantification proceeds in a manner which allows for a degree of heterogeneity across firms.

In this paper we follow Mitchell, Smith and Weale (2004) in quantifying firms' categorical responses by constructing a "disaggregate" indicator of output growth that is built round ordered discrete choice models linking individual firms' categorical responses to official data and then inferring the most likely values for the official data given the categorical responses. For convenience we adopt their notation and lay-out in this paper. However, extending the work of Mitchell, Smith and Weale (2004), here we show how the prospective or forward looking individual responses can be combined to produce forecasts of output growth, that is, we consider how the qualitative prospective survey data published at time $t-1$ can be converted into a quantitative indicator of expected growth in period t.

The plan of this paper is as follows. Section 2 derives so-called "disaggregate" forecasts of manufacturing output growth that exploit firm-level forward-looking survey information. Section 3 considers an application to survey data from the Confederation of British Industry; the performance of the disaggregate forecasts is compared with traditional aggregate forecasts both in-sample and out-of-sample. Section 4 makes some concluding comments.

2 Firm-Level Quantification of Prospective Survey Data

We consider a survey that asks a sample of N_{t-1} manufacturing firms near the end of period $t-1$ whether their output growth is expected to rise, not change or fall over period t compared to period $t-1$. Our method of quantification of the prospective survey responses follows Mitchell, Smith and Weale (2004) in postulating an underlying relationship between firm specific output growth and the official data for aggregate output growth and then extends their approach to the case of forward-looking or prospective survey responses. Our approach is influenced by the fact that the number of respondents varies from period to period, although a reasonable number respond in the majority of the periods.

[2] There has been limited previous work using individual responses to surveys [see Nerlove, 1983; Horvath, Nerlove and Willson, 1992; McIntosh, Schiantarelli and Low, 1989; Branch, 2004; Souleles, 2004]. However, this work has focused on testing the nature of expectation formation.

The categorical responses in the survey are assumed to relate to observed official data for economy-wide manufacturing output growth x_t in the following manner. Let the actual output growth of firm i at time t, y_{it}, which may be known to firm i but is assumed unknown to the econometrician, depend on x_t according to the conditional linear model

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

$(t = 1, ..., T)$, where η_{it} is the difference between y_{it} and x_t anticipated by firm i, reflecting information private to firm i at time t that is not observed by the econometrician. This information may reflect firm or industry level influences. The random variable ε_{it} captures the component of firm-specific output growth y_{it} unanticipated by both firm i and the econometrician at time t. That is, $E(y_{it}|\Omega_t^i) = x_{it} = x_t + \eta_{it}$, where Ω_t^i comprises information available to firm i at time t and includes x_t. In the following analysis it is further assumed that output growth x_t is a stationary variable, an assumption supported by tests for a unit root in the level series of manufacturing output.

At the end of period $t-1$ firm i makes a prediction, y_{it}^*, of y_{it} based on macroeconomic information available to the firm, the relevant individual specific information set Ω_{t-1}^i, $i = 1, ..., N_{t-1}$. If this prediction is formed rationally

$$y_{it}^* = E(y_{it}|\Omega_{t-1}^i) = x_t^* + \eta_{it}, \tag{2}$$

where $x_t^* = E(x_t|\Omega_{t-1})$ is the economy wide rational expectation of x_t, Ω_{t-1} is the macroeconomic information set available to *all* firms at time $t-1$ and

$$x_t = x_t^* + \zeta_t, \tag{3}$$

where ζ_t is a white-noise macroeconomic shock unanticipated by firms such that $E(\zeta_t|\Omega_{t-1}) = E(\zeta_t|\Omega_{t-1}^i) = 0$.

2.1 The Relationship between the Prospective Survey Data and the Official Data

We may re-express (2) as

$$y_{it}^* = \alpha_i + \beta_i x_t^* + \varepsilon_{it}^*, \tag{4}$$

where ε_{it}^* is a mean zero random variable. In (4), α_i and β_i are firm-specific time-invariant coefficients expressed in terms of (2) by defining $\eta_{it} = \alpha_i + (\beta_i - 1)x_t^* + \varepsilon_{it}^*$, $(i = 1, ..., N_{t-1}$, $t = 1, ..., T)$.[3] After substitution from (3)

[3] We may also write $E(y_{it}^*|\Omega_{t-1}) = \alpha + \beta x_t^*$ where $E(\alpha_i|\Omega_{t-1}) = \alpha$ and $E(\beta_i|\Omega_{t-1}) = \beta$. Let z_{it} denote (the level of) expected output of firm i at time t, so that $y_{it}^* = \Delta z_{it}/z_{it-1}$ where Δ denotes the first difference operator. From (4), $\sum_{i=1}^{N_{t-1}} \Delta z_{it} = \sum_{i=1}^{N_{t-1}} z_{it-1}\alpha_i + \sum_{i=1}^{N_{t-1}} z_{it-1}\beta_i x_t^* + \sum_{i=1}^{N_{t-1}} z_{it-1}\varepsilon_{it}^*$, after cross-multiplication and summation over $i = 1, ..., N_{t-1}$. For coherence we require $\sum_{i=1}^{N_{t-1}} \Delta z_{it}/\sum_{i=1}^{N_{t-1}} z_{it-1} \xrightarrow{p} x_t^*$, $\sum_{i=1}^{N_{t-1}} z_{it-1}\alpha_i/\sum_{i=1}^{N_{t-1}} z_{it-1} \xrightarrow{p} 0$, $\sum_{i=1}^{N_{t-1}} z_{it-1}\beta_i/\sum_{i=1}^{N_{t-1}} z_{it-1} \xrightarrow{p} 1$ and $\sum_{i=1}^{N_{t-1}} z_{it-1}\varepsilon_{it}^*/\sum_{i=1}^{N_{t-1}} z_{it-1} \xrightarrow{p} 0$ $(N_{t-1} \to \infty)$.

$$
\begin{aligned}
y_{it}^{*} &= \alpha_i + \beta_i(x_t - \zeta_t) + \varepsilon_{it}^{*} && (5)\\
&= \alpha_i + \beta_i x_t + \psi_{it}, && (6)
\end{aligned}
$$

where $\psi_{it} = \varepsilon_{it}^{*} - \beta_i \zeta_t$. Estimation of (6) thus needs to deal with the endogeneity of x_t. To do so we follow Smith and Blundell (1986) and Newey (1987) by supplementing (6) with the process assumed to govern the determination of x_t which yields a two equation simultaneous model.

We assume that x_t, measured relative to its mean, is generated by an autoregressive process which without loss of generality is taken to be first-order[4]

$$
x_t = \lambda x_{t-1} + u_t, \tag{7}
$$

$(t = 1, ..., T)$, where $|\lambda| < 1$ to ensure stationarity of output growth and u_t is an *i.i.d.* mean zero disturbance. We further assume that the bivariate distribution governing ψ_{it} and u_t is such that conditional on u_t we can write their dependence in the form

$$
\psi_{it} = \rho_i u_t + \nu_{it}, \tag{8}
$$

where ρ_i, $(i = 1, ..., N_{t-1})$, is a firm-specific parameter and ν_{it} is an *i.i.d.* mean zero disturbance distributed independently of u_t. Substitution in (6) generates the conditional model

$$
\begin{aligned}
y_{it}^{*} &= \alpha_i + \beta_i x_t + \rho_i u_t + \nu_{it} && (9)\\
&= \alpha_i + \beta_i \lambda x_{t-1} + (\rho_i + \beta_i) u_t + \nu_{it}.
\end{aligned}
$$

We consider two-step estimation along the lines suggested in Smith and Blundell (1986) and Newey (1987). First, estimate (7) by least squares which yields the consistent estimator $\widehat{\lambda}$ for λ (as $T \to \infty$) and the least squares residuals $\widehat{u}_t = x_t - \widehat{\lambda} x_{t-1}$, $(t = 1, ..., T)$. Secondly, estimate the parameters α_i, β_i, μ_{ji} and ρ_i in (9) after substituting $\widehat{u}_t$ for u_t (or $\widehat{\lambda}$ for λ and $\widehat{u}_t$ for u_t) as described below.

Anticipated growth y_{it}^{*} of firm i at time t is unobserved but the survey at time $t-1$ contains data corresponding to whether output growth is expected to rise, not change or fall in period t. 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 (9). Define the indicator variables

$$
y_{it-1}^{j} = 1 \text{ if } \mu_{(j-1)i} < y_{it}^{*} \leq \mu_{ji} \text{ and } 0 \text{ otherwise}, \ (j = 1, 2, 3), \tag{10}
$$

corresponding to "down", "same" and "up", respectively, where $\mu_{0i} = -\infty$, μ_{1i}, μ_{2i} and $\mu_{3i} = \infty$ are firm-specific threshold parameters.[5] We assume that the error terms ν_{it},

[4] Additional lagged terms in x_t can be included if necessary to render u_t serially uncorrelated.

[5] Discrete choice models are only identified up to scale; including the intercept α_i in (9) necessitates setting, for example, the first threshold parameter μ_{1i} to zero to achieve identification.

$(t = 1, ..., T)$, are logistic with common cumulative distribution function (c.d.f.) $F(z) = [1 + \exp(-z)]^{-1}$, $-\infty < z < \infty$, $(i = 1, ..., N_{t-1})$.[6] The probabilistic foundation for the observation rule (10) is given by the conditional probability $P(j|x_{t-1}, i, u_t)$ of observing the categorical response $y^j_{it-1} = 1$ for choice j for firm i at time $t-1$ given the values of x_{t-1} and u_t

$$\begin{aligned} P(j|x_{t-1}, i, u_t) &= F(\mu_{ji} - \alpha_i - \beta_i x_t - \rho_i u_t) - F(\mu_{(j-1)i} - \alpha_i - \beta_i x_t - \rho_i u_t) \\ &= F(\mu_{ji} - \alpha_i - \beta_i \lambda x_{t-1} - (\rho_i + \beta_i) u_t) \\ &\quad - F(\mu_{(j-1)i} - \alpha_i - \beta_i \lambda x_{t-1} - (\rho_i + \beta_i) u_t), \end{aligned} \tag{11}$$

$(j = 1, 2, 3)$. Given the assumption that the errors ν_{it} are independently and identically distributed over time, the likelihood function for firm i is

$$L_i = \prod\nolimits_{t=2}^{T} P_{1it-1}^{y^1_{it-1}} P_{2it-1}^{y^2_{it-1}} P_{3it-1}^{y^3_{it-1}}, \tag{12}$$

where $P_{jit-1} \equiv P(j|x_{t-1}, i, u_t)$, $(j = 1, 2, 3)$. Under the above assumptions, maximisation of (12) yields consistent estimates $(T \to \infty)$ of α_i, β_i, ρ_i and μ_{ji} denoted by $\widehat{\alpha}_i$, $\widehat{\beta}_i$, $\widehat{\rho}_i$ and $\widehat{\mu}_{ji}$, $(j = 0, ..., 3)$, respectively after substitution of $\widehat{u}_t$ for u_t (or $\widehat{\lambda}$ for λ and $\widehat{u}_t$ for u_t).[7] The estimation of u_t (or λ and u_t) will affect the asymptotic standard errors of these estimates. Alternatively rather than estimating *via* maximum likelihood, Bayesian methods such as Markov Chain Monte Carlo could be employed; see Albert and Chib (1993).

2.2 Inferring the Official Data

Given an ordered logit model for each firm i, an estimator for x_t may be inferred from the prospective survey data published at time $t-1$; we adapt the method of Mitchell, Smith and Weale (2004) which is based on the use of Bayes' Theorem. In so doing the qualitative prospective survey data at time $t-1$ are converted into a quantitative indicator of expected growth in period t available to users of the survey at $t-1$. This indicator can then be evaluated as a one-step ahead forecast of x_t.

Our initial interest centres on the conditional density $f(u_t|j, i, x_{t-1})$ for observing u_t given the survey response j for firm i at time $t-1$ and the time $t-1$ value of the official data. Let $f(u_t)$ denote the time-invariant probability density function (p.d.f.) of u_t, where $f(u_t) = f(u_t|x_{t-1})$ since the shocks u_t that hit the aggregate economy at t are assumed independent of x_{t-1}. This p.d.f. is also assumed normal with mean zero and variance $E(u_t^2)$.[8] Therefore, the conditional probability given x_{t-1} of observing response j for firm

[6] The logistic distribution is similar in shape to the normal but has slightly heavier tails. The logistic distribution is convenient since it offers a closed form distribution function.

[7] As the parameters α_i, β_i, ρ_i and μ_{2i} are only identified up to scale, the decision probabilities (11) are invariant to multiplying (9) by an arbitrary constant.

[8] These assumptions are supported by empirical tests of the type considered in footnote 10 below.

i is $P(j|i,x_{t-1}) = \int_{-\infty}^{\infty} P(j|x_{t-1},i,u_t)f(u_t)du_t$. Bayes' Theorem states that

$$f(u_t|j,i,x_{t-1}) = \frac{P(j|x_{t-1},i,u_t)f(u_t)}{P(j|i,x_{t-1})}. \tag{13}$$

For firm i, the Bayes estimator (under squared error loss) for u_t given j_{t-1} and x_{t-1} is the mean of the posterior density $f(u_t|j_{t-1},i,x_{t-1})$:

$$E(u_t|j,i,x_{t-1}) = \int_{-\infty}^{\infty} u_t f(u_t|j,i,x_{t-1})du_t, \tag{14}$$

which at time t takes one of three values depending on the observed sample response j of firm i at time $t-1$ and the value of lagged output growth x_{t-1}. Given $f(u_t)$, all of the above integrals may be calculated by numerical evaluation.

Estimators $\widehat{P}(j|x_{t-1},i,u_t)$ for $P(j|x_{t-1},i,u_t)$ and, thus, $\widehat{P}(j|i,x_{t-1})$ for $P(j|i,x_{t-1})$ are given by substitution of the estimators $\widehat{\alpha}_i$, $\widehat{\beta}_i$, $\widehat{\rho}_i$, $\widehat{\mu}_{ji}$, $(j=0,...,3)$, and $\widehat{\lambda}$ in (11), *viz.*

$$\begin{aligned}\widehat{P}(j|x_{t-1},i,u_t) &= F(\widehat{\mu}_{ji} - \widehat{\alpha}_i - \widehat{\beta}_i\widehat{\lambda}x_{t-1} - (\widehat{\rho}_i + \widehat{\beta}_i)u_t) \\ &\quad -F(\widehat{\mu}_{(j-1)i} - \widehat{\alpha}_i - \widehat{\beta}_i\widehat{\lambda}x_{t-1} - (\widehat{\rho}_i + \widehat{\beta}_i)u_t).\end{aligned} \tag{15}$$

Hence, a feasible Bayes estimator $\widehat{E}(u_t|j,i,x_{t-1})$ may be obtained from (14) by numerical evaluation.

Let j_{it-1}, $(j_{it-1} = 1,2,3)$, denote the prospective survey response of firm i published at time $t-1$. To create a disaggregate indicator for economic activity at time $t-1$, by the law of iterated expectations the conditional expectation of x_t given all firms' survey responses j_{it-1}, $(i=1,...,N_{t-1})$, and x_{t-1}

$$E(x_t|\{j_{it-1}\}_{i=1}^{N_{t-1}}, x_{t-1}) = \sum\nolimits_{i=1}^{N_{t-1}} H_{it-1}E(x_t|j_{it-1},i,x_{t-1}), \tag{16}$$

where H_{it-1} is the exogenous sample probability of observing firm i at time $t-1$. Now, from (7),

$$E(x_t|j_{it-1},i,x_{t-1}) = \lambda x_{t-1} + E(u_t|j_{it-1},i,x_{t-1}). \tag{17}$$

Hence, assuming firms are independent, we define the parametric indicator

$$\widehat{x}_t^D = \sum\nolimits_{i=1}^{N_{t-1}} w_{it-1}\widehat{E}(x_t|j_{it-1},i,x_{t-1}), \tag{18}$$

where $w_{it-1} > 0$ is the weight assigned to firm i at time $t-1$, $\sum_{i=1}^{N_t} w_{it-1} = 1$ and, from (17), $\widehat{E}(x_t|j_{it-1},i,x_{t-1}) = \widehat{\lambda}x_{t-1} + \widehat{E}(u_t|j_{it-1},i,x_{t-1})$, $(i=1,...,N_{t-1})$. 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-1} 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 $\widehat{x}_t^{ND}$ is an estimator for the conditional expectation $E(x_t|\{j_{it-1}\}_{i=1}^{N_{t-1}})$ which may be based on the conditional empirical distribution function. Define the indicator function $I(x_t \leq x, j_{it-1} = j|i) = 1$ if $x_t \leq x$ and $j_{it-1} = j$ and 0 otherwise, $(j = 1, 2, 3)$. Let $T_i^j = \sum_{s=2}^{T} y_{i,s-1}^j$ which is the number of times firm i gives response j in the survey; hence, $T_i^j/(T-1)$ is the sample proportion for response j by firm i, $(j = 1, 2, 3)$. The conditional empirical distribution function of x_t given response j for firm i is given by $\hat{F}(x|j,i) = \sum_{s=2}^{T} I(x_s \leq x, j_{is-1} = j|i)/T_i^j$, $(j = 1, 2, 3)$, which assigns equal weight to each sample value. Assuming $T_i^j \to \infty$ and, thus, $T \to \infty$, $T_i^j/T \xrightarrow{p} P(j|i)$, the probability of observing response j for firm i, and $\sum_{s=2}^{T} I(x_s \leq x, j_{is-1} = j|i)/(T-1) \xrightarrow{p} F(x, j|i)$, if, given firm i, x_s and j_{is-1} 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 $\hat{F}(x|j,i)$, $\sum_{s=2}^{T} y_{i,s-1}^j x_s/T_i^j$, is a consistent estimator for $E(x_t|j,i)$. A nonparametric disaggregate indicator is therefore defined as

$$\widehat{x}_t^{ND} = \sum\nolimits_{i=1}^{N_{t-1}} w_{it-1} \sum_{s=2}^{T} y_{i,s-1}^{j_{it-1}} x_s / T_i^{j_{it-1}}. \tag{19}$$

We note here, and subsequently in section 3.1.2, that although the disaggregate indicators in practice have a good correlation with the official data they show much less volatility. Less volatility is observed because the scale is incorrect. One explanation for this is based on those firms whose prospective responses are poorly correlated with actual output growth. In the extreme case where responses are uncorrelated with output, the inclusion of these reduces the standard deviation of the indicator but does not affect its correlation with output growth. Excess smoothness of the disaggregate indicators can then be explained by the presence of firms in the sample of survey responses whose responses contain little signal about output growth and are essentially 'noise'. To reconcile this incompatibility in volatility between the outturn and the indicators for manufacturing output growth, note that the outturn is the signal recovered from the survey data plus a residual error component. Rescaling the indicators through linear regression on the outturn is one simple method of obtaining an indicator which tracks output growth as closely as possible. The effects of this regression will also be taken into account in our subsequent out-of-sample analysis. Specifically we align the disaggregate indicators with the official data by regressing the outturn x_t on the indicator as follows

$$x_t = \varphi_0 + \varphi_1 \widehat{x}_t^k + \xi_t; \text{ for } k = D, ND. \tag{20}$$

In fact, for the parametric indicator $\widehat{x}_t^D$ we consider the following unrestricted form of (20) that should better pick up the dynamic nature of x_t

$$x_t = \varphi_0 + \varphi_1^* x_{t-1} + \varphi_2^* \widehat{u}_t^D + \xi_t, \tag{21}$$

where $\widehat{u}_t^D = \sum_{i=1}^{N_{t-1}} w_{it-1} \widehat{E}(u_t|j_{it-1}, i, x_{t-1})$. Note that when $\varphi_1^* = \widehat{\lambda}\varphi_2^*$ only the magnitude of $\widehat{x}_t^D$ is affected by this re-scaling.

2.3 Producing Out-Of-Sample Forecasts from the Qualitative Survey Data

The section above provides a means of linking the prospective survey responses to official data on output growth. The techniques discussed there can be used to quantify these forward-looking survey data in-sample. However, to be made operational out-of-sample we need to accommodate the fact that the official data for output growth are published with a lag.

From (17), the estimator for the expectation of x_t conditional on j_{it-1}, i and x_{t-1} is given by

$$\widehat{E}(x_t|j_{it-1}, i, x_{t-1}) = \widehat{\lambda}x_{t-1} + \widehat{E}(u_t|j_{it-1}, i, x_{t-1}). \tag{22}$$

But out-of-sample x_{t-1} is unknown, although we do know j_{it-1} since the survey data are published ahead of the official data.[9] We can however make use of the value generated as in (18). We denote this value by $\hat{x}^D_{t-1|t-2}$ to emphasise that it is calculated only using information up to time $t-2$ using the logistic equations (9) whereas $\hat{x}^D_{t-1}$ was computed from estimates of the logistic equations including period $t-1$. We denote the density function of x_{t-1} conditional on $\hat{x}^D_{t-1|t-2}$ as $g(x_{t-1}|\hat{x}^D_{t-1|t-2})$. We cannot estimate this directly from the individual logistic equations because the joint density function of combinations of these is unknown. Instead we explore the time-series relationship between x_{t-1} and $x^D_{t-1|t-2}$. We can accept the hypothesis that $E(x_{t-1} - \hat{x}^D_{t-1|t-2}) = 0$ and $x_{t-1} - \hat{x}^D_{t-1|t-2}$ is normally distributed; we therefore assume the variance to be $E(x_{t-1} - \hat{x}^D_{t-1|t-2})^2$.[10]

The second term on the right hand side of (22), using the forecast $\hat{x}^D_{t-1|t-2}$ for x_{t-1}, is then given as

$$\widehat{E}(u_t|j_{it-1}, i, \hat{x}^D_{t-1|t-2}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u_t f(u_t|j_{it-1}, i, x_{t-1}) g(x_{t-1}|\hat{x}^D_{t-1|t-2}) du_t dx_{t-1}. \tag{23}$$

An alternative approach would be to consider firm-specific forecasts for x_{t-1} instead of using our aggregate indicator $\hat{x}^D_{t-1|t-2}$, and examine $\widehat{E}(x_{t-1}|j_{it-2}, i, x_{t-2})$; see (17). But given that our aim is the production of an aggregate forecast for x_t again it seems natural to base this on an aggregate forecast $\hat{x}^D_{t-1|t-2}$ for x_{t-1}.

Similarly to (21), the out-of-sample forecasts are re-scaled based on the in-sample estimates of φ_0, φ^*_1 and φ^*_2, denoted $\widehat{\varphi}_0$, $\widehat{\varphi}^*_1$ and $\widehat{\varphi}^*_2$, so that the re-scaled out-of-sample

[9] For example, imagine it is January. While the survey has been published indicating what firms expect to happen in the first quarter, not only are official data for manufacturing output growth not available for this quarter but they are also not yet available for the last quarter of the previous year.

[10] For example, over the in-sample period (using in-sample estimates) $1988q4 - 1997q4$, see below, we test the normality of $g(.)$ using a modified version of the Jarque-Bera test that is robust against serial correlation and conditional heteroscedasticity in $x_{t-1} - \hat{x}^D_{t-1|t-2}$; see Bai and Ng (2003). We find that normality is not rejected with a p-value of 0.202. There is also no statistical evidence for non-zero bias in $(x_{t-1} - \hat{x}^D_{t-1|t-2})$, again using a robust estimator for the standard error. Unbiasedness is not rejected with a p-value of 0.114.

forecast, $\widehat{\hat{x}}_t^D$, is given as

$$\widehat{\hat{x}}_t^D = \widehat{\varphi}_0 + \widehat{\varphi}_1^* \hat{x}_{t-1|t-2}^D + \widehat{\varphi}_2^* \sum_{i=1}^{N_{t-1}} w_{it-1} \widehat{E}(u_t | j_{it-1}, i, \hat{x}_{t-1|t-2}^D). \tag{24}$$

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 are the expected trends for the next four months with regard to volume of output?".

Firms can respond either "up", "same", "down" or "not applicable".[11] This prospective question provides the basis for a one-quarter ahead (leading) indicator of manufacturing output growth x_t.

We consider a sample of 43,936 firm-level responses to the ITS from October 1988 to October 1997, although we do extend the sample to October 1999 to analyse the out-of-sample performance of the alternative forecasts; see section 3.2.[12] Responses from the ITS are available in January, April, July and October. The prospective responses are assumed to relate to the period $1988q4 - 1997q4$, although essentially published one quarter earlier. For example, the January survey represents information formed before any first quarter information became available to firms; but the prospective question asks firms for their expectation of what will happen during this first quarter.

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 (9) for all firms. In contrast to traditional aggregate quantification techniques, firm level quantification requires sufficient time-series observations for a given firm for reliable parameter estimation.

[11] The number of firms that answered "not applicable" is very small and is ignored in later analysis. Along with most other authors who have written on the subject, we treat the difference between the four-month period referred to in the survey and the quarterly frequency of our official data as being unimportant.

[12] Unfortunately it was not possible to extend the out-of-sample analysis beyond 1999. In December 1999 the CBI moved to a new survey processing platform that involved changing the participant identification numbers. This means it is no longer straightforward to match firms pre and post December 1999 which is necessary to construct the panel data set of survey responses. We thank Jamie Morrison and Jonathan Wood of the CBI for their advice about these data.

In this application, we consider 20 observations to be satisfactory.[13] If, given i, the error terms ν_{it} are independent conditional on x_{t-1} and u_t, $(t = 1, ..., T)$, these observations need not be consecutive. Hence, firms that do not respond to at least 20 surveys are dropped from the sample used to derive disaggregate indicators of manufacturing output growth. Since these firms are dropped there is a danger that the sample selection could induce bias in the disaggregate indicators.[14] In any case, notwithstanding the implied theoretical properties of the disaggregate indicators, their usefulness is determined by how well they perform in practice, both in-sample and out-of-sample, relative to the traditionally used aggregate quantification techniques. This should, and does, serve as the main test of their value.

A possible alternative to our disaggregate approach, that avoids the need to drop data for some firms, is to pool the data by imposing homogeneity restrictions across firms and then exploit traditional panel-data estimators; see Hsiao (2003). However, we do not follow this approach here since our results, see Table 1 below, indicate considerable heterogeneity across firms in their slope coefficients; therefore imposing a common slope coefficient would result in heterogeneity bias.

Over the period $1988q3 - 1997q3$, twenty non-consecutive time series observations are available for 698 manufacturing firms. The number of firms who respond to the survey varies from one period to the next. Both unweighted and weighted indicators are considered; the weights, based on firms' sales volumes, are those used by the CBI in aggregating firms' responses. We have no *a priori* reason to believe that one set of weights is more applicable than another.

To give an impression of the nature of the survey responses, Figure 1 plots the percentage of the 698 firms that expected output growth to "go up", "stay the same" or "go down" over the data period. It also plots the subsequent outturn for quarterly growth at an annual rate of (seasonally adjusted) manufacturing output. Visual inspection of

[13] 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 disaggregate indicators considered a range of so-called "cut-off" values; see Mitchell, Smith and Weale (2004). Since in practice the disaggregate indicator appears to behave similarly across a wide range of cut-off values we confine attention here to a cut-off value of 20.

[14] We did consider the following test for sample selection. Let "included sample" denote those firms with at least 20 time series observations. Let "excluded sample" denote those firms in the full-sample omitted from the included sample. 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. We considered the correlation of traditional aggregate indicators, reviewed in the Appendix, 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. This is consistent with the view that the included sample may be regarded as a random sample, and that inference from it should be unbiased. This implies that if disaggregate indicators outperform traditional aggregate indicators we can conclude that this improvement is due to disaggregation *per se*, and is not the consequence of using a different sample. We also supplemented the above by using forecast encompassing tests to examine whether the aggregate indicators derived from the excluded firms add information *vis-à-vis* the disaggregate indicators. Again, there was little evidence to suggest that dropping those firms with fewer than 20 observations led to any informational loss.

the graph suggests that the survey responses track movements in manufacturing output growth, particularly in the sense that the firms were right to be pessimistic during the recessionary period at the beginning of the 1990s. Indeed given this, and the fact that the proportions do not appear to track output growth as well after this recession, we should expect the in-sample predictive performance of aggregate methods that use these proportions to be flattered when analysed over this recessionary period. The explanatory power of the aggregate proportions will deteriorate when the recession of the early 1990s is not considered. A striking aspect of the survey is the number of firms which expect no change in their output. Nerlove (1983) comments on the fact that prospective output growth is much more concentrated on "no change" than are reports about what has (retrospectively) happened to output. This is obviously consistent with a situation where substantial deviations from the initial expectation are the result of shocks which were not forecast.

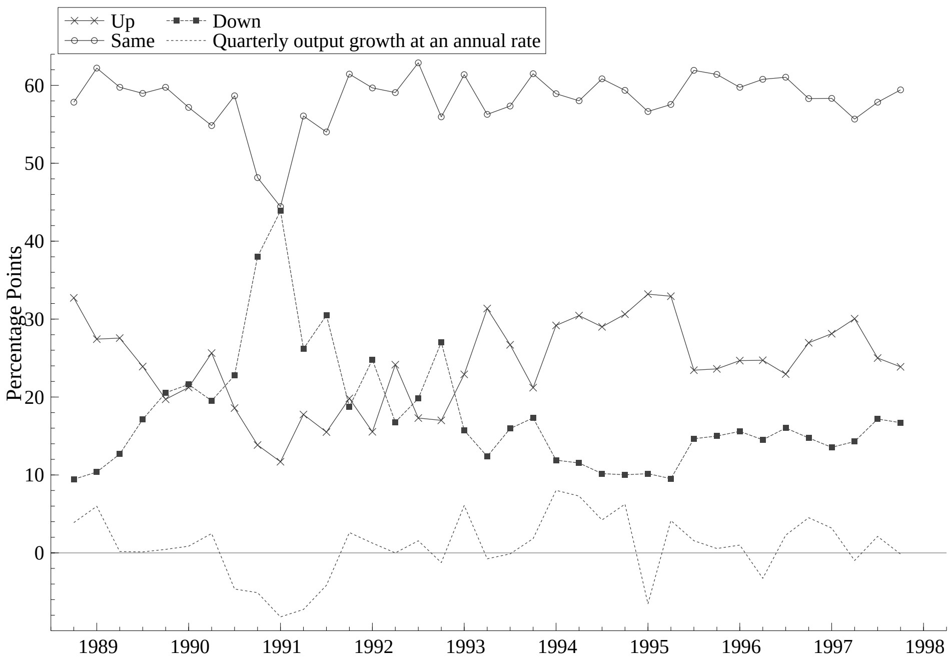

Figure 1: Unweighted percentage of firms expecting output growth to "go up", "stay the same" or "go down", alongside the subsequent outturn for output growth

3.1 Disaggregate indicators

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

The parametric disaggregate indicator $\widehat{x}_t^D$ is based on firm-level estimation. We estimated ordered logit models, (9), for each of the 698 firms. To illustrate the degree of heterogeneity across firms in terms of how they react to changes in the aggregate environment we focus simply on the estimates $\{\widehat{\beta}_i\}$ for these firms. Given that, for identification, the first threshold parameter μ_{1i} is set to zero for all i, comparisons of $\widehat{\beta}_i$ across i are meaningful. It should be added that if we wished also to consider the associated estimated standard errors the estimation of the autoregression (7) at the first-step should not be ignored, otherwise the standard errors will be inconsistent; see Newey and McFadden (1994) and Smith and Blundell (1986, section 3). Since our focus is on constructing forecasts for x_t, it is the magnitude of $\{\widehat{\beta}_i\}$, rather than the precision with which they are estimated, that matters; cf. (18).

In order to give some impression of heterogeneity across firms, Table 1 displays the number of firms that have estimates $\widehat{\beta}_i$ in a specified range.[15]

Table 1: The number of firms that have values for $\widehat{\beta}_i$ in a specified range

$\widehat{\beta}_i$					
$\widehat{\beta}_i \leq -2$	$-2 < \widehat{\beta}_i \leq -1$	$-1 < \widehat{\beta}_i \leq 0$	$0 < \widehat{\beta}_i \leq 1$	$1 < \widehat{\beta}_i \leq 2$	$\widehat{\beta}_i > 2$
19	34	175	242	156	67

Table 1 reveals considerable variation across firms in how their prospective survey responses relate to manufacturing output growth. However, the majority of firms, 67%, exhibit a positive relationship with x_t which is consistent with our prior belief that, in general, we should expect a rise in manufacturing output growth to be anticipated by a rise in expected firm-specific output growth.

3.1.2 Comparing the Performance of the Aggregate and Disaggregate Forecasts

We compare the performance of the proposed disaggregate forecasts against those of two traditional quantification techniques employed using aggregate proportions: the probability method of Carlson and Parkin (1975) [CP] and the regression approach of Pesaran (1984) [P].[16] See the Appendix for a selective review of these techniques. Consistent with

[15] The maximum likelihood estimation routine employed did not converge for 5 firms; hence results for the parametric disaggregate indicator are based on those 693 firms where there was convergence.

[16] A related approach is the reverse regression approach of Cunningham, Smith and Weale (1998), as modified by Mitchell, Smith and Weale (2002). In contrast to the regression approach official data, rather than survey data, are used as the regressors. Under the assumption that (after revisions) official data

(21) we also considered quantification using the regression approach augmented with a lagged dependent variable to better reflect the dynamic nature of x_t. In fact, results were little different to those based on (A.8) and, therefore, are not presented. The CP indicator can less straightforwardly be adapted to reflect the persistence of x_t.

The CP aggregate forecast is identified up to a scaling parameter [see the Appendix]. Following CP we chose this parameter to ensure that the mean of the quantified series is equal to the mean of the outturn over the sample-period. This does not imply that the forecast is unbiased in the statistical sense. In contrast, the regression forecasts are unbiased since they implicitly estimate the scaling parameter through regression-based methods. Moreover, as indicated above in section 2.2, although the disaggregate indicators have a good correlation with the official data typically we have found that they show much less volatility. This feature has been observed elsewhere with alternative indicators; see, for example, Cunningham (1997) and Mitchell, Smith and Weale (2004). The disaggregate indicators are therefore re-scaled following (20) and (21). To compare fairly between the alternative aggregate (and disaggregate) leading indicators, mean squared error (MSE) criteria therefore are not used to evaluate the performance of these forecasts as MSE is dependent on scale.

To summarise the performance of indicators of manufacturing output growth, we examine their correlation coefficient (corr.) with the outturn for manufacturing output growth over the period $1988q4 - 1997q4$. Correlation informs us about the informational content of the indicator series; when the square of the correlation statistic is strictly positive the indicator series explains some of the (subsequent) variation in manufacturing output growth about its mean.[17] A high value for corr. indicates that a strong signal about the outturn may be recovered from the indicator regardless of how the indicator has been scaled, and whether the 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; see Table 2.[18] However, this may not be true for all possible weighting schemes.

Table 2 summarises the properties of the aggregate and disaggregate forecasts. Table 2 does make clear that the parametric and nonparametric disaggregate forecasts, $\widehat{x}_t^D$ and $\widehat{x}_t^{ND}$, explain more of the variation in output growth one quarter ahead than the aggregate indicators. The disaggregate indicators provide more accurate leading indicators of output growth than traditional aggregate indicators.

---

offer unbiased estimates of the economic variable under consideration this avoids problems caused by measurement error in the data. Results using the reverse regression approach were similar to those using the other two aggregate indicators and are therefore not presented here.

[17] Equivalently, the indicator series has some informational content when, in a linear regression of the outturn on the indicator and an intercept, R^2 is greater than zero.

[18] Therefore, henceforth we confine attention to the unweighted forecasts. In fact, consideration of the weighted forecasts in the out-of-sample analysis in section 3.2, in any case, would be impeded by our lack of knowledge of the weights, based on sales volumes, during the out-of-sample period. Crucially the weights based on sales volumes are time-varying.

Table 2: Aggregate and Disaggregate Forecast Performance: In-sample, $1988q4 - 1997q4$

		Mean	Stand. Dev.	Corr.
Outturn for Manuf. Output Growth		0.807	3.951	
CP	unweighted	0.807	1.333	0.627
	weighted	0.807	1.493	0.625
P	unweighted	0.807	2.628	0.665
	weighted	0.807	2.583	0.654
$\widehat{x}_t^D$	unweighted	0.807	3.580	0.906
	weighted	0.807	3.503	0.887
$\widehat{x}_t^{ND}$	unweighted	0.807	3.559	0.900
	weighted	0.807	3.555	0.899

3.2 Out-of-sample analysis

Having found an improved in-sample fit between the prospective survey responses and official data using the disaggregate rather than aggregate forecasts, this section examines whether the superiority of the disaggregate forecasts extends out-of-sample. To evaluate how accurate the survey-based forecasts of output growth would have been out-of-sample we conduct an experiment designed to mimic "real-time" application of the different quantification approaches. We are nevertheless assessing the performance against near-final rather than initial official data.

The out-of-sample analysis is conducted using the prospective survey responses over the 8 periods, $1997q4 - 1999q3$, so that forecasts for output growth are obtained for $1998q1 - 1999q4$. Just as with the in-sample analysis conducted above, on an out-of-sample basis we relate the survey data published at time t (but assumed to refer to $(t+1)$) to official data for $(t+1)$. However, out-of-sample we need to reflect the fact that the official data for output growth are published with a lag. Indeed, in previous work we found that the retrospective survey responses (published at time t and referring to t) can be exploited to obtain useful 'early' estimates of these official data, given that the survey data are published ahead of the official data; see Mitchell, Smith and Weale (2004).

The analysis is performed by conducting the following recursive experiments. When using the prospective survey responses published in $1997q4$ to forecast output growth in $1998q1$ since the official data for $1997q4$ are assumed not yet published, the in-sample estimates, used as the basis for the out-of-sample forecasts, relate the prospective survey data published in $1988q3$ to $1997q2$ to official data for $1988q4$ to $1997q3$. Section 2.3 details how the parametric disaggregate indicator is made operational out-of-sample. A similar delay is used in the application of the aggregate methods. Then we forecast output growth for $1998q2$ using the prospective data published in $1998q1$, given the in-sample estimates based on relating the prospective survey data from $1988q3$ to $1997q3$ to official data from $1988q4$ to $1997q4$. This recursive process is carried on until survey data published in $1999q3$ are used to forecast output growth in $1999q4$, given in-sample estimates based on relating the survey data published in $1988q3$ to $1999q1$ to official data for $1988q4$ to

1999q2. Both the aggregate and disaggregate out-of-sample estimates are re-scaled by recursively regressing their in-sample counterparts against the outturn for output growth following (20) and (21); we denote these forecasts $\widehat{\widehat{x}}_t^D$ and $\widehat{\widehat{x}}_t^{ND}$. In this way no *ex post* information about output growth is used when quantifying the survey data in "real-time".

As is traditional when evaluating forecasts, the performance of the aggregate and disaggregate indicators is evaluated in terms of their root MSE against the outturn. The results of this recursive exercise are summarised in Table 3. Table 3 also contrasts the performance of the aggregate and disaggregate survey based forecasts with the forecasts from a benchmark time-series model, AR. As discussed in section 2.3 to avoid conditioning on x_{t-1} when forecasting x_t, as data are not published in time, these AR forecasts are constructed as two-quarter ahead linear projections based on the following autoregressive model with lag order p in the growth rate of output growth:

$$x_t = \phi_1 + \sum\nolimits_{l=2}^{p} \phi_l x_{t-l} + \tau_t \tag{25}$$

where ϕ_l $(l = 1, ..., p)$ are parameters, estimated recursively using data up to $t-2$, and τ_t is a mean zero disturbance. The lag order p was selected recursively by the Bayesian information criterion with $2 \leq p \leq 6$.

The first conclusion from Table 3 is that the nonparametric disaggregate indicator produces more accurate forecasts than the aggregate indicators.[19] The nonparametric disaggregate survey based forecasts also beat those of AR, the benchmark time-series model. However, the parametric disaggregate indicator does not deliver as accurate forecasts as its nonparametric cousin. Its forecasts are only marginally more accurate than the aggregate indicators, and only then when new firms are allowed to enter the sample during the out-of-sample period. More work is required to consider whether the improved in-sample performance of the parametric disaggregate indicator, relative to the aggregate indicator, can be translated into better performance out-of-sample. The manner in which this indicator is re-scaled should be central to this.

We should conclude on a cautionary note. Of course, out-of-sample analysis, particularly with small samples, is always sensitive to the period chosen. This is particularly so in this application where output growth is far less volatile out-of-sample than in the in-sample period; output growth has a standard deviation of 3.9% over the period 1988q4 to 1997q4, while it has a standard deviation of 2.6% from 1998q1 $-$ 1999q4. Experimentation with a 16 rather than 8 quarter out-of-sample period did, however, deliver similar results to those in Table 3.

4 Concluding Comments

Using a panel of firm level survey responses obtained from the CBI disaggregate forecasts for output growth are derived using ordered discrete choice models relating firms'

[19] Our out-of-sample period is too small for sensible use of formal statistical tests to test whether this improvement in forecast accuracy is significant statistically; see Ashley (2003).

Table 3: Aggregate and Disaggregate Forecast Performance: Out-of-sample, $1998q1 - 1999q4$

	Root MSE
CP	3.763
P	3.906
$\widehat{\widehat{x}}_t^D$	3.732
$\widehat{\widehat{x}}_t^{ND}$	2.117
AR	2.653

Notes. The results for the disaggregate forecasts presented in Table 3 allow "new" firms to enter the sample during the out-of-sample period. If we restrict attention in the out-of-sample analysis to those 693 firms present in the in-sample period similar results are obtained; the parametric and nonparametric disaggregate forecasts now have root MSE, respectively, of 3.850 and 2.298, rather than 3.732 and 2.117.

prospective categorical survey responses to a quantitative measure of economic activity. There was considerable heterogeneity across firms in how their responses relate to the measure of economic activity. The disaggregate forecasts outperformed traditional aggregate forecasts in terms of anticipating movements in manufacturing output growth on an in-sample basis. Out-of-sample the evidence was mixed with only the nonparametric disaggregate indicator providing more accurate forecasts than those of the aggregate indicators. Nevertheless, confirming the results of Mitchell, Smith and Weale (2004), our results indicate that better signals of economic behaviour can be derived from analysis of the panel data set of individual responses underlying the aggregate responses. In future work we aim to enhance further the appropriate means of extracting this signal.

References

[1] Albert, J.H. and S. Chib (1993), "Bayesian analysis of binary and polychotomous response data", *Journal of the American Statistical Association*, 88, 669-679.

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

[3] Ashley, R. (2003), "Statistically significant forecasting improvements: how much out-of-sample data is likely necessary?", *International Journal of Forecasting*, 19, 229-239.

[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] Bouton, F. and H. Erkel-Rousse (2002), "Conjonctures sectorielles et prévision à court terme de l'activité: l'apport de l'enquête de conjoncture dans les services", *Économie et Statistique,* 359-360, 35-68.

[8] Branch, W.A. (2004), "The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations", *Economic Journal*, 114, 592-621.

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

[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] Driver, C. and G. Urga (2004), "Transforming qualitative survey data: performance comparisons for the UK", *Oxford Bulletin of Economics and Statistics*, 66, 71-90.

[13] Entorf, H. (1993), "Constructing leading indicators from non-balanced sectoral business survey series", *International Journal of Forecasting*, 9, 211-225.

[14] Gregoir, S. and F. Lenglart (2000), "Measuring the probability of a business cycle turning point by using a multivariate qualitative hidden Markov model", *Journal of Forecasting*, 19, 81-102.

[15] Hild, F. (2002), "Une lecture enrichie des réponses aux enquêtes de conjoncture", *Économie et Statistique*, 359-360, 13-33.

[16] Horvath, B., M. Nerlove and D. Willson (1992). "A reinterpretation of direct tests of forecast rationality using business survey data". In *Business Cycle Analysis by Means of Economic Surveys: Part I* (eds. Oppenländer, K.H. and Poser, G.). Avebury, Aldershot.

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

[18] Madsen, J.B. (1993), "The predictive value of production expectations in manufacturing industry", *Journal of Forecasting*, 12, 273-289.

[19] McIntosh, J., F. Schiantarelli and W. Low (1989), "A qualitative response analysis of UK firms' employment and output decisions", *Journal of Applied Econometrics*, 4, 251-264.

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

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

[22] Mitchell, J., R.J. Smith and M.R. Weale (2004), "Aggregate versus disaggregate survey-based indicators of economic activity". Revision of National Institute of Economic and Social Research Discussion Paper No. 194.

[23] Nerlove, M. (1983). "Expectations, plans, and realizations in theory and practice". *Econometrica*, 51, 1251-1279.

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

[25] Newey, W.K. and D.L. McFadden (1994), "Large sample estimation and hypothesis testing". In *Handbook of Econometrics Volume IV* (eds. Engle, R.F. and D.L. McFadden), Elsevier.

[26] Parigi, G. and G. Schlitzer (1995), "Quarterly forecasts of the Italian business cycle by means of monthly economic indicators", *Journal of Forecasting*, 14, 117-141.

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

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

[29] Pesaran, M.H. and M.R. Weale (2004), "Survey expectations", Unpublished paper, NIESR.

[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] 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.

[33] 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 in period $t-1$ whether output growth in period t, x_t, is expected to "fall", "stay the same" or "rise" relative to the previous period. Since the proportion of respondents who reply "fall", "stay the same" and "rise" sum to unity the survey contains two pieces of independent information at time t.[20] Let U_t and D_t denote the proportions of firms that respectively expect output growth to rise and fall in period t.

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 movement of x_t can be derived given certain assumptions. Here we briefly outline two alternative methods of quantification [for more details and references to the literature see Mitchell, Smith and Weale (2004)]:

- the probability approach of Carlson and Parkin (1975) [CP];
- the regression approach of Pesaran (1984, 1987) [P].

A.1 The Probability Approach

The probability method of quantification assumes that firm i's expectation of economy-wide manufacturing output growth x_t is derived from a subjective probability density function for x_t, $f_i(x_t|\Omega^i_{t-1})$, conditional on the information set Ω^i_{t-1} available to the firm, that is, $x^*_{it} = E(x_t|\Omega^i_{t-1})$. The responses of firm i are classified as follows.

- "up" if $x^*_{it} \geq b_{it}$;
- "down" if $x^*_{it} \leq -a_{it}$;
- "same" if $-a_{it} < x^*_{it} < b_{it}$,

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

Assuming that firms are independent and that the structure of $f_i(x_t|\Omega^i_{t-1})$ is the same for all firms then x^*_{it} can be regarded as an independent draw from an aggregate density $f(x_t|\Omega_{t-1})$ where Ω_{t-1} is the information set available to all firms. The density $f(x_t|\Omega_{t-1})$ is assumed to have 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

$$D_t \xrightarrow{p} P(x_t \leq -\lambda \mid \Omega_{t-1}) = F_t(-\lambda), \tag{A.1}$$

$$U_t \xrightarrow{p} P(x_t \geq \lambda \mid \Omega_{t-1}) = 1 - F_t(\lambda), \tag{A.2}$$

[20] The number of firms answering "not applicable" tends to be very small and is ignored here.

where $F_t(\cdot)$ is the cumulative distribution function obtained from $f(x_t|\Omega_{t-1})$.

The traditional [CP] approach assumes that $f(\cdot)$ is a normal density function with mean x_t^* and standard deviation σ_t^*; alternative densities $f(\cdot)$ may be also considered, see, for example, Batchelor (1981) and Mitchell (2002).

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

$$D_t = \Phi(\frac{-\lambda - x_t^*}{\sigma_t^*}), \tag{A.3}$$

$$1 - U_t = \Phi(\frac{\lambda - x_t^*}{\sigma_t^*}), \tag{A.4}$$

where $\Phi(\cdot)$ is the standard normal c.d.f.. Using (A.3) and (A.4) to solve for x_t^* and σ_t^*,

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

where $\Phi^{-1}(\cdot)$ denotes the inverse standard normal c.d.f.. Thus,

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

which leaves only λ undetermined. In the literature λ has been calculated in various ways. [CP] assume unbiasedness over the sample period, $t = 1, ..., T$; that is, λ is estimated as

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

For alternative approaches, see *inter alia* Batchelor (1981,1982), Pesaran (1984), and Wren-Lewis (1985). Since λ is constant over time, its role is merely to scale x_t^*.

A.2 The Regression Approach

The regression approach, at its simplest, relates output growth x_t to the aggregate proportions U_t and D_t as follows

$$x_t = \alpha U_t - \beta D_t. \tag{A.8}$$

The unknown parameters α and β can be estimated *via* regression of x_t on U_t and D_t. The fitted values from this estimated regression then provide the quantified prospective 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.