University of Leicester

DEPARTMENT OF ECONOMICS

Efficient Aggregation of Panel Qualitative Survey Data

James Mitchell, University of Leicester, UK
Richard J. Smith, Cambridge University, UK
Martin R. Weale, Monetary Policy Committee, Bank of England

Working Paper No. 11/53
December 2011

Efficient Aggregation of Panel Qualitative Survey Data*

James Mitchell*,†, Richard J. Smith‡,† and Martin R. Weale+,†

*Department of Economics, University of Leicester

†National Institute of Economic and Social Research

‡Faculty of Economics, Cambridge University

+Monetary Policy Committee, Bank of England

December 2011.

Abstract

Qualitative business survey data are used widely to provide indicators of economic activity ahead of the publication of official data. Traditional indicators exploit only aggregate survey information, namely the proportions of respondents who report "up" and "down". This paper examines disaggregate or firm-level survey responses. It considers how the responses of the individual firms should be quantified and combined if the aim is to produce an early indication of official output data. Having linked firms' categorical responses to official data using ordered discrete-choice models, the paper proposes a statistically efficient means of combining the disparate estimates of aggregate output growth which can be constructed from the responses of individual firms. An application to firm-level survey data from the Confederation of British Industry shows that the proposed indicator can provide early estimates of output growth more accurately than traditional indicators.

JEL Classification: C35; C53; C80

Keywords: Survey Data; Indicators; Quantification; Forecasting; Forecast Combination

---

*Address for correspondence: James Mitchell, Department of Economics, University of Leicester, University Road, Leicester, LE1 7RH, U.K. Tel: +44 (0) 116 252 2884. E-Mail: jm463@le.ac.uk.

This paper replaces an earlier paper entitled "Aggregate versus Disaggregate Survey-Based Indicators of Economic Activity". The views expressed in this paper are those of the authors and do not represent those of the Monetary Policy Committee at the Bank of England.

1 Introduction

Statisticians and economists are under considerable pressure to produce up-to-date estimates of the state of the economy. In this paper we develop a statistically efficient means of using disaggregate data from qualitative business surveys to produce an indicator of the state of the economy. Such an indicator is valuable because such surveys are generally completed much more rapidly than is the production of official data: they are often available within a few days of the end of the month or quarter to which they relate. These surveys ask *inter alia* whether, after adjusting for normal seasonal movements, output has risen, stayed the same or fallen in recent months. The question thus arises how formally to convert the findings of such surveys into early estimates of movements in economic activity. The traditional approach to this question has been to take the aggregate findings of such surveys, i.e., the proportion of firms reporting that output has risen, stayed the same or fallen, and to relate them to official output data. Approaches suggested have included the probability method [Carlson & Parkin (1975)] and the regression method [Pesaran (1984, 1987)], plus variants of these; see Pesaran & Weale (2006) for a survey. Collectively, we call these approaches "aggregate"; the aggregate datum to which they give rise may then be used on its own or combined with other variables in some form of model such as the factor models produced by Stock & Watson (2002) and Forni et al. (2001).

However, we are concerned with a much more basic 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 quantified and combined if the aim of the survey is to produce an early indication of official output data?[1,2] Indeed there is no intrinsic reason to believe that working with the proportion of firms in each response category is the best basis for linking such surveys to official output data. It may well be that quantification in a manner which allows for a degree of heterogeneity among firms would exploit individual firm information more efficiently than do the traditional approaches and would therefore allow more accurate inferences to be drawn about output movements.

This paper therefore proposes a framework for quantifying and aggregating qualitative survey responses of firms. Individual qualitative responses of each firm are linked

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

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

to the overall official growth rate based on that firm's reporting record. A statistically efficient means of combining the disparate (quantitative) estimates of aggregate output growth is then set out which can be constructed from the responses of individual firms to a qualitative business survey. The principle underlying the approach is similar to those underlying traditional forecast combination, but the qualitative nature of the data obviously raises important new issues. As with traditional forecast combination our approach results in weights which give more emphasis to firms whose answers have a close link to the official data rather than to those whose experiences correspond only weakly or not at all. The resultant estimator is compared with an alternative estimator, considered in Mitchell et al. (2005*b*), which takes a simple average of individual responses across firms. However, it is well-known that simple averaging is not an efficient means of forecast or nowcast combination, irrespective of its performance in real-time applications; see Bates & Granger (1969) and Granger & Ramanathan (1984).

Use of the proposed technique is illustrated in an application to industrial survey data from the Confederation of British Industry (CBI). We find that it explains more of the variation in manufacturing output growth than traditional indicators constructed using "aggregate" data.

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

2 Quantification Across Firms

Consider a survey that asks a sample of N_t manufacturing firms at time t whether their output 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; let N denote the overall number of different firms sampled.

The actual output growth rate y_{it} of firm i at time t is unobserved but the qualitative survey contains data corresponding to whether output has risen, not changed or fallen relative to the previous period. To account for the ordinal nature of the responses and their relationship to the firm-specific growth rate y_{it}, define the indicator variables

$$y_{it}^{j} = 1 \text{ if } \mu_{(j-1)i} < y_{it} \leq \mu_{ji} \text{ and 0 otherwise, } (j = 1, 2, 3), \tag{2.1}$$

corresponding to "down", "same" and "up", respectively, where $\mu_{0i} = -\infty$, μ_{1i}, μ_{2i} and $\mu_{3i} = \infty$ are time invariant firm-specific threshold parameters.

The categorical responses y_{it}^{j}, $(j = 1, 2, 3)$, in the survey are assumed to be related to the output growth rate x_t, as measured quantitatively by the national statistical office, *via* the latent firm-specific growth rate y_{it}, $(i = 1, ..., N_t)$, in the following manner. Here x_t could be the aggregate, i.e., economy-wide, growth rate of output or some published disaggregate such as sectoral output. Importantly, while both y_{it}^j, $(j = 1, 2, 3)$, and x_t refer to time period t, the former is observable, i.e., published, at time t ahead of the x_t data which are published with a lag at time $(t+1)$. Let y_{it}, $(i = 1, ..., N_t)$, depend on x_t according to the linear model

$$y_{it} = \alpha_i + \beta_i x_t + \gamma_i' z_t + \varepsilon_{it}, \tag{2.2}$$

$(t = 1, ..., T)$, where α_i, β_i and γ_i are firm-specific time-invariant coefficients. The firm-invariant vector z_t consists of additional observable exogenous explanatory variables dated t or earlier, e.g., the *aggregate* qualitative survey data $\sum_{i=1}^{N_t} y_{it}^j$, $(j = 1, 2, 3)$, at time t. Their inclusion may accommodate common cross-sectional dependence in firms' categorical responses arising from common shocks or factors, cf. Pesaran (2006), by, for example, capturing those sectoral, cyclical and/or seasonal components in y_{it} not explained by x_t.

2.1 Dependence

Macroeconomic data are widely accepted to exhibit dependence over time. Consequently the error term ε_{it} in (2.2) might also be expected to incorporate some dynamic macroeconomic features. To illustrate suppose that x_t follows the stationary first order dynamic process

$$x_t = \alpha_x + \beta_x x_{t-1} + \gamma_x' z_t + u_t, \tag{2.3}$$

$(t = 1, ..., T)$, with $|\beta_x| < 1$ and u_t an error term. Additional lagged terms in x_t may be included in (2.3) if x_t is thought to be generated by a higher order process. The presence of z_t in (2.3) allows for the possibility of correlation between x_t and z_t, an assumption typically made by aggregate quantification techniques when the proportions of optimistic and pessimistic firms are included in z_t; see below and Appendix C.

If the dependence between ε_{it} and u_t takes the linear form

$$\varepsilon_{it} = \rho_i u_t + \xi_{it}, \tag{2.4}$$

where ρ_i is a firm-specific parameter and ξ_{it} a disturbance term, $(i = 1, ..., N_t)$, then substitution of (2.4) in (2.2) generates the dynamic model

$$\begin{aligned} y_{it} &= \alpha_i + \beta_i x_t + \gamma_i' z_t + \rho_i u_t + \xi_{it} \\ &= \alpha_i^* + \beta_{i0}^* x_t + \beta_{i1}^* x_{t-1} + \beta_{i2}^{*\prime} z_t + \xi_{it}, \end{aligned} \tag{2.5}$$

$(i = 1, ..., N_t)$, where the firm-specific coefficients $\alpha_i^* = \alpha_i - \rho_i \alpha_x$, $\beta_{i0}^* = \beta_i + \rho_i$, $\beta_{i1}^* = -\rho_i \beta_x$ and $\beta_{i2}^* = \gamma_i - \rho_i \gamma_x$.

More precisely, we assume the conditional linear specification $E[y_{it}|\{x_\tau, z_\tau\}_{\tau=1}^t, i] = \alpha_i^* + \beta_{i0}^* x_t + \beta_{i1}^* x_{t-1} + \beta_{i2}^{*\prime} z_t$ where the notation $\{x_\tau, z_\tau\}_{\tau=1}^t, i$ indicates information available to firm i at time t and necessarily includes current and lagged information on x_t. Hence, $E[\xi_{it}|\{x_\tau, z_\tau\}_{\tau=1}^t, i] = 0$ and ξ_{it} is uncorrelated with current and past values of x_t and z_t rendering $\{x_\tau, z_\tau\}_{\tau=1}^t$ predetermined by assumption. The error term ξ_{it} then captures the component of firm-specific output growth y_{it} unanticipated by both firm i and the econometrician at time t given the macroeconomic information on x_τ and z_τ, $(\tau = 1, ..., t)$. We further assume that, conditional on $\{x_\tau, z_\tau\}_{\tau=1}^t$ and i, the error term ξ_{it} is independent of the lagged values of firm-specific growth $\{y_{i\tau}\}_{\tau=1}^{t-1}$ and is normally distributed with common cumulative distribution function (c.d.f.) $F_i(\cdot)$, $(t = 1, ..., T)$.

2.2 Ordered Discrete Choice Models

The probabilistic foundation for the observation rule (2.1) is given by the conditional probability $P_{jit} = P_i(j|\{x_\tau, z_\tau\}_{\tau=1}^t, i)$ of observing the categorical response $y_{it}^j = 1$ for choice j at time t given the information set $\{x_\tau, z_\tau\}_{\tau=1}^t$ and firm i, i.e.,

$$P_{jit} = F_i(\mu_{ji} - \alpha_i^* - \beta_{i0}^* x_t - \beta_{i1}^* x_{t-1} - \beta_{i2}^{*\prime} z_t) - F_i(\mu_{(j-1)i} - \alpha_i^* - \beta_{i0}^* x_t - \beta_{i1}^* x_{t-1} - \beta_{i2}^{*\prime} z_t),\ (j = 1, 2, 3). \tag{2.6}$$

As discrete choice models are only identified up to scale, including the intercept α_i^* in (2.5) necessitates setting, for example, the first threshold parameter μ_{1i} to zero for identification. Consequently the decision probabilities (2.6) are invariant to multiplying (2.5) by an arbitrary constant, i.e., the parameters in (2.5) are identified only up to the firm-specific time-invariant conditional variance $\sigma_{\xi i}^2 = var[\xi_{it}|\{x_\tau, z_\tau\}_{\tau=1}^t, i]$. In principle, the variance $\sigma_{\xi i}^2$ might be conditionally heteroskedastic also depending on $\{x_\tau, z_\tau\}_{\tau=1}^t$. Like much of the discrete choice literature we normalise $\sigma_{\xi i}^2$ to unity to achieve identification. Under our assumptions, the likelihood function for firm i is

$$L_i = \prod_{t=1}^T P_{1it}^{y_{it}^1} P_{2it}^{y_{it}^2} P_{3it}^{y_{it}^3}. \tag{2.7}$$

Maximum likelihood (ML) based on (2.7) yields consistent and asymptotically efficient estimates $(T \to \infty)$ of α_i^*, β_{i0}^*, β_{i1}^*, β_{i2}^* and μ_{ji} which we denote by $\hat{\alpha}_i^*$, $\hat{\beta}_{i0}^*$, $\hat{\beta}_{i1}^*$, $\hat{\beta}_{i2}^*$ and $\hat{\mu}_{ji}$ respectively.[3] In addition, if the error terms ξ_{it} in (2.5) are independently distributed over firms $(i = 1, ..., N_t)$ conditional on $\{x_\tau, z_\tau\}_{\tau=1}^t$, there is no efficiency loss involved in estimation of the ordered discrete choice models *via* (2.7) firm-by-firm rather than as a

[3]The error terms ξ_{it} in (2.5) may still be serially correlated, $(t = 1, ..., T)$. If, however, ξ_{it} are standard normally distributed conditional on $\{x_\tau, z_\tau\}_{\tau=1}^t$ and i, $(t = 1, ..., T)$, which permits the presence of serial correlation, (2.7) is then a pseudo- or quasi-likelihood function. Note that the coefficient estimates, $\hat{\alpha}_i^*$, $\hat{\beta}_{i0}^*$, $\hat{\beta}_{i1}^*$, β_{i2}^* and $\hat{\mu}_{ji}$, remain consistent but the standard ML asymptotic variance matrix is no longer appropriate and requires adjustment; see, e.g., Robinson (1982). However, we do not require estimator standard errors in the following.

system unless firms are homogeneous in parameters.

An alternative approach to the fixed effects-type approach described above is a random effects-type formulation of (2.5) which incorporates parameter homogeneity across firms and imposes additional conditional independence assumptions. Consequently firms are pooled (across i) and the resultant pooled method would be more efficient when these restrictions hold. Re-express (2.5) as $y_{it} = \alpha^* + \beta_0^* x_t + \beta_1^* x_{t-1} + \beta_2^{*\prime} z_t + \zeta_{it}$, where $\zeta_{it} = (\alpha_i^* - \alpha^*) + (\beta_{i0}^* - \beta_0^*)x_t + (\beta_{i1}^* - \beta_1^*)x_{t-1} + (\beta_{i2}^* - \beta_2^*)' z_t + \xi_{it}$ and $E[\alpha_i^* | \{x_\tau, z_\tau\}_{\tau=1}^t] = \alpha^*$, $E[\beta_{ik}^* | \{x_\tau, z_\tau\}_{\tau=1}^t] = \beta_k^*$, $(k = 0, 1, 2)$.

In general, however, $E[\alpha_i^* | \{x_\tau, z_\tau\}_{\tau=1}^t] = \alpha^*(\{x_\tau, z_\tau\}_{\tau=1}^t) \neq \alpha^*$ and $E[\beta_{ik}^* | \{x_\tau, z_\tau\}_{\tau=1}^t] = \beta_k^*(\{x_\tau, z_\tau\}_{\tau=1}^t) \neq \beta_k^*$, $(k = 0, 1, 2)$. In particular, the firm-level effects α_i^* and β_{ik}^*, $(k = 0, 1, 2)$, are likely to be correlated with the outturn, x_t, rendering a random effects-type panel-data treatment of (2.5) inconsistent through the presence of heterogeneity bias.

The validation of the above assumptions explicit or otherwise is a necessary concomitant in any empirical application. Appendix A details various diagnostic tests used in the application discussed below. Assumptions adopted in (2.5) include linearity, conditional homoskedasticity and that the error term ξ_{it} is standard normally distributed conditional on $\{x_\tau, z_\tau\}_{\tau=1}^t$ and i. In addition, tests for the endogeneity of x_t and dynamic dependence on $\{x_\tau\}_{\tau=1}^{t-1}$ in (2.2) are undertaken together with the assumption of the stationarity of the output growth rate x_t. Cross-sectional independence is examined by the test proposed by Hsiao et al. (2011) adapted for use with nonlinear panel data models.

2.3 Inferring the Official Data: the Proposed Indicator

Given ordered probit models for each firm i, $(i = 1, ..., N_t)$, in either their fixed effects-type or a random effects-type panel data model forms, an estimator for x_t may be inferred from the qualitative survey data. As qualitative survey data are usually published ahead of the official data, this would provide an early quantitative estimate (or nowcast) of x_t. Although we focus below on the former specification where ordered probit models are estimated separately for each firm, we also indicate which alterations need to be made if, for example, a random effects-type panel data model is used instead. Our indicator is designed to address a situation where there is heterogeneity in model parameters across firms, i.e., the coefficients in (2.5) scaled by $\sigma_{\xi i}$. If the data supported the hypothesis of homogeneity in model parameters and justified a fully pooled model, then our approach would be unnecessary because it would be appropriate to give all firms the same weight.

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 indicator requires the density function of x_t conditional on the N_t firms' observed survey responses at time t, $\{j_{it}\}_{i=1}^{N_t}$, and macroeconomic information $\{x_\tau\}_{\tau=1}^{t-1}, \{z_\tau\}_{\tau=1}^t$. Denote this density function as $f(x_t | \{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1})$. Also let $f(x_t | z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1})$ denote the prior conditional density function of x_t given z_t and $\{x_\tau, z_\tau\}_{\tau=1}^{t-1}$, $(t = 1, ..., T)$.

Independence of ξ_{it} across i, $(i = 1, ..., N_t)$, conditional on $\{x_\tau, z_\tau\}_{\tau=1}^{t}$, implies that firms' categorical responses are conditionally independent across firms. Therefore, the joint conditional probability of observing the N_t firms' categorical responses, $\{j_{it}\}_{i=1}^{N_t}$, is the product of their marginal probabilities $P_i(j_{it}|\{x_\tau, z_\tau\}_{\tau=1}^{t}, i)$ (2.6), i.e.,

$$P(\{j_{it}\}_{i=1}^{N_t}|\{x_\tau, z_\tau\}_{\tau=1}^{t}) = \prod_{i=1}^{N_t} P_i(j_{it}|\{x_\tau, z_\tau\}_{\tau=1}^{t}, i). \tag{2.8}$$

Therefore, the joint conditional probability of observing response j across firms i, $(i = 1, ..., N_t)$, given z_t and $\{x_\tau, z_\tau\}_{\tau=1}^{t-1}$, is[4]

$$P(\{j_{it}\}_{i=1}^{N_t}|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}) = \int_{-\infty}^{\infty} \prod_{i=1}^{N_t} P_i(j_{it}|\{x_\tau, z_\tau\}_{\tau=1}^{t}, i) f(x_t|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}) dx_t.$$

Bayes' Theorem states that:

$$f(x_t|\{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}) = \frac{P(\{j_{it}\}_{i=1}^{N_t}|\{x_\tau, z_\tau\}_{\tau=1}^{t}) f(x_t|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1})}{P(\{j_{it}\}_{i=1}^{N_t}|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1})}. \tag{2.9}$$

The proposed indicator D_t is defined as the Bayes estimator (under squared error loss) for x_t given $\{j_{it}\}_{i=1}^{N_t}$, z_t and $\{x_\tau, z_\tau\}_{\tau=1}^{t-1}$, i.e., the mean of the posterior density $f(x_t|\{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1})$,

$$E[x_t|\{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}] = \int_{-\infty}^{\infty} x_t f(x_t|\{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}) dx_t. \tag{2.10}$$

Given $f(x_t|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1})$, and knowledge of the parameters α_i^*, β_{i0}^*, β_{i1}^*, β_{i2}^* and μ_{ji}, $(j = 0, 1, 2, 3)$, $(i = 1, ..., N)$, all of the above integrals may be calculated by numerical evaluation. Estimators $\hat{P}(j_{it}|\{x_\tau, z_\tau\}_{\tau=1}^{t}, i)$ for $P(j_{it}|\{x_\tau, z_\tau\}_{\tau=1}^{t}, i)$ and, thus, $\hat{P}(j_{it}|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}, i)$ for $P(j_{it}|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}, i)$ are given by substitution of the estimators $\hat{\alpha}_i^*$, $\hat{\beta}_{i0}^*$, $\hat{\beta}_{i1}^*$, $\widehat{\beta}_{i2}^*$ and $\hat{\mu}_{ji}$, $(j = 0, 1, 2, 3)$, $(i = 1, ..., N)$, in (2.6). The feasible empirical Bayes estimator

$$D_t = \hat{E}[x_t|\{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}] \tag{2.11}$$

may then be obtained from (2.10) by numerical evaluation. The impact of the use of plug-in (estimated) parameters, instead of priors for these parameters, is expected to be

[4]Let $\theta_i^* = (\alpha_i^*, \beta_{i0}^*, \beta_{i1}^*, \beta_{i2}^{*\prime})'$, $(i = 1, ..., N_t)$. For a random effects-type panel data model where $\theta_i^*|\{x_\tau, z_\tau\}_{\tau=1}^{t} \sim g(\cdot)$,

$$P(\{j_{it}\}_{i=1}^{N_t}|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \prod_{i=1}^{N_t} P_i(j_{it}|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}, \theta_i^*, i) f(x_t|z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}) g(\theta_i^*) dx_t d\theta_i^*.$$

small in circumstances when the likelihood (2.7) dominates these priors; e.g., in large samples and/or when the priors are vague.[5]

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

To illustrate this point suppose that both $\beta^*_{10} = \beta^*_{11} = 0$ and, for simplicity, also set $\beta^*_{12} = 0$. Consequently, firm 1's categorical survey responses offer no information about the official data. For this firm $P_1(j_{1t}|\{x_\tau\}^t_{\tau=1}, 1) = P_1(j_{1t}|1)$. Hence, (2.9) becomes

$$f(x_t|\{j_{it}\}^{N_t}_{i=1}, \{x_\tau\}^{t-1}_{\tau=1}) = \frac{\prod_{i=2}^{N_t} P_i(j_{it}|\{x_\tau\}^t_{\tau=1}, i) f(x_t|\{x_\tau\}^{t-1}_{\tau=1})}{P(\{j_{it}\}^{N_t}_{i=2}|\{x_\tau\}^{t-1}_{\tau=1})}, \tag{2.12}$$

implying firm 1 receives no weight in the indicator D_t.

The indicator D_t may be contrasted with an alternative indicator $\overline{D}_t$ for economic activity at time t proposed in Mitchell et al. (2005*a*, 2005*b*), which although inefficient does not require the cross-sectional independence of ξ_{it}. Density functions are calculated separately for each firm for x_t conditional on the survey response j_{it}. An average is then taken of these across firms. To be more explicit, again for expositional ease ignoring z_t, the conditional probability of observing response j for firm i is $P_i(j_{it}|\{x_\tau\}^{t-1}_{\tau=1}, i) = \int_{-\infty}^{\infty} P_i(j_{it}|\{x_\tau\}^t_{\tau=1}, i) f(x_t|\{x_\tau\}^{t-1}_{\tau=1}) dx_t$. Bayes' Theorem then states

$$f(x_t|j_{it}, \{x_\tau\}^{t-1}_{\tau=1}, i) = \frac{P_i(j_{it}|\{x_\tau\}^t_{\tau=1}, i) f(x_t|\{x_\tau\}^{t-1}_{\tau=1})}{P_i(j_{it}|\{x_\tau\}^{t-1}_{\tau=1}, i)}. \tag{2.13}$$

For firm i, the Bayes estimator (under squared error loss) for x_t given j_{it}, $\{x_\tau\}^{t-1}_{\tau=1}$ and i is the mean of the posterior density $f(x_t|j_{it}, \{x_\tau\}^{t-1}_{\tau=1}, i)$; *viz.*

$$E[x_t|j_{it}, \{x_\tau\}^{t-1}_{\tau=1}, i] = \int_{-\infty}^{\infty} x_t f(x_t|j_{it}, \{x_\tau\}^{t-1}_{\tau=1}, i) dx_t, \tag{2.14}$$

[5]Random effects can be regarded as a form of Bayesian estimation that uses $g(\cdot)$ as the prior distribution for the firm-level parameters. Note that the resultant estimator of the indicator D_t is still large-T consistent as long as $g(\cdot)$ is constructed from the posterior distribution of the effects. MCMC methods provide a convenient method for the computation of estimators and asymptotically valid confidence intervals. In this sense fixed effects and random effects approaches are rather similar and, thus, neither one may be more robust than the other. We are grateful to the Co-Editor for this point.

which, conditional on $\{x_\tau\}_{\tau=1}^{t-1}$, takes one of three values depending on the observed sample response j_{it} of firm i at time t. If $P_i(j_{it}|\{x_\tau\}_{\tau=1}^{t}, i) = P_i(j_{it}|i)$, i.e., the responses of firm i are unrelated to movements in the official series, the posterior mean estimates (2.14) for each category j will be identical for firm i, i.e., the mean conditional growth rate $E[x_t|\{x_\tau\}_{\tau=1}^{t-1}]$ of the official series, $(t = 1, ..., T)$. In all other cases estimates based on (2.14) will provide some indication about the growth rate of the official series.

By the law of iterated expectations the feasible indicator $\overline{D}_t$ of Mitchell et al. (2005*a*, 2005*b*) is given as

$$\overline{D}_t = \hat{E}[x_t|\{j_{it}\}_{i=1}^{N_t}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}, i] = \sum_{i=1}^{N_t} H_{it}\hat{E}[x_t|j_{it}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}, i]. \qquad (2.15)$$

where H_{it} is the exogenous sampling probability of observing firm i at time t which is thus independent of response j_{it}, z_t and $\{x_\tau, z_\tau\}_{\tau=1}^{t-1}$. If firms $(i = 1, ..., N_t)$ constitute a random sample, then equal weights are appropriate since all firms are equally likely to appear in the sample, i.e., $H_{it} = N_t^{-1}$. However, if firms are drawn according to an exogenous stratified sampling scheme, then H_{it} should reflect the stratum weights. Like D_t, $\overline{D}_t$ is a consistent estimator for the output growth rate x_t. Mitchell et al. (2005*a*, 2005*b*) consider both an equal weighting scheme and one based on firm size.

3 CBI Survey Data

The *Industrial Trends Survey* (ITS) of the CBI, which is conducted on a quarterly basis, gives qualitative opinion from UK manufacturing firms on past and expected trends in output, exports, prices, costs, investment intentions, business confidence and capacity utilisation. Various questions from the survey, typically when aggregated to the "balance of opinion" (namely the proportion of optimists less pessimists), have been the focus of attention by both policy-makers (Ashley et al. 2005) and academics [e.g. see Lee (1994), Driver & Urga (2004) and Pesaran & Weale (2006) for a review]. 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 for deriving timely indicators (or nowcasts) of quarterly output growth x_t (at an annual rate). The number that answer "not applicable" is very small and is 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 $51,225$ responses from the ITS. The sample records the survey responses of, in total, 5422 firms over the period $1988q3$ to $1999q3$. Unfortunately it was not possible to extend the analysis beyond 1999 since in $1999q4$ the CBI moved to a new survey processing platform that involved changing the participant identification numbers making it impossible to match firms pre- and post-December 1999. There are, on average, only 1133 firms in the sample at time t, with 9.4 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 (2.5) without assuming some homogeneity in behaviour across firms. Quantification based on (2.5) requires sufficient time-series observations for a given firm for reliable parameter estimation.

In the application below based on the fixed effects-type formulation of (2.5), we consider twenty observations to be satisfactory. This choice of so-called "cut-off" value is rather arbitrary. In the application below, when examining the performance of the disaggregate indicators D_t and $\overline{D}_t$, we did consider a range of "cut-off" values. In practice the indicators appear to behave rather similarly across quite a wide range of values.[6] If, given i, the error term ξ_{it} in (2.5) is independent of the lagged values of firm-specific growth $\{y_{i\tau}\}_{\tau=1}^{t-1}$ conditional on $\{x_\tau, z_\tau\}_{\tau=1}^{t}$, i, $(t = 1, ..., T)$, observations need not be consecutive. Hence, firms that do not respond to at least twenty surveys are dropped from the sample used to derive the indicator D_t (and $\overline{D}_t$) of output growth. There is a danger that this sample selection could induce bias in the D_t (and $\overline{D}_t$) indicator.

We examined, and subsequently rejected, the possibility of sample selection bias using a comparison of the performance of the aggregate indicators in the "included" and "excluded" samples. In the absence of sample selection, the included sample may be regarded as a random sample from the full-sample and inference from both included and excluded samples should be equivalent apart from sampling error. That is, indicators or statistics derived from both included and excluded samples should not differ significantly. (For more details see also Mitchell et al. (2005a, 2005b).) In any case, notwithstanding the implied theoretical properties of the indicators, their usefulness should primarily be determined by how well they perform in practice relative to the traditionally used quantification techniques employed with aggregate survey data.

The alternative random effects-type approach described above has the advantage of not requiring any firms to be dropped but at the expense of the imposition of parameter homogeneity across firms. The log-likelihood function (2.7), following Butler & Moffitt

[6] Deterioration was more marked for a high rather than a small "cut-off" value; as the number of firms used to compute the disaggregate indicator became very small (< 10) the performance of the indicator also began to deteriorate substantially.

(1982), is revised as

$$\log L = \sum_{i=1}^{N} \log \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left(P_{1i,t}^{y_{i,t}^{1}} P_{2i,t}^{y_{i,t}^{2}} P_{3i,t}^{y_{i,t}^{3}} \right) g(\theta_i^*) d\theta_i^* \tag{3.16}$$

where N is the total number of different firms present over time ($t = 1, ..., T$) and $g(\cdot)$ denotes the conditional density of the parameters $\theta_i^* = (\alpha_i^*, \beta_{i0}^*, \beta_{i1}^*, \beta_{i2}^{*\prime})'$, ($i = 1, ..., N$). As discussed above, while this formulation has the apparent advantage of facilitating construction of the disaggregate indicator using the full panel of firms, however unbalanced this may be, it does rest on the assumptions $E[\alpha_i^* | \{x_\tau, z_\tau\}_{\tau=1}^t] = \alpha^*$ and $E[\beta_{ik}^* | \{x_\tau, z_\tau\}_{\tau=1}^t] = \beta_k^*$, ($k = 0, 1, 2$). Otherwise the random effects-type panel-data estimators will no longer be consistent because of heterogeneity bias.

Over the period $1988q3 - 1999q3$ twenty non-consecutive time series observations are available for 834 manufacturing firms. To give an impression of the nature of the survey responses, Figure 1 plots the percentage of these 834 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.

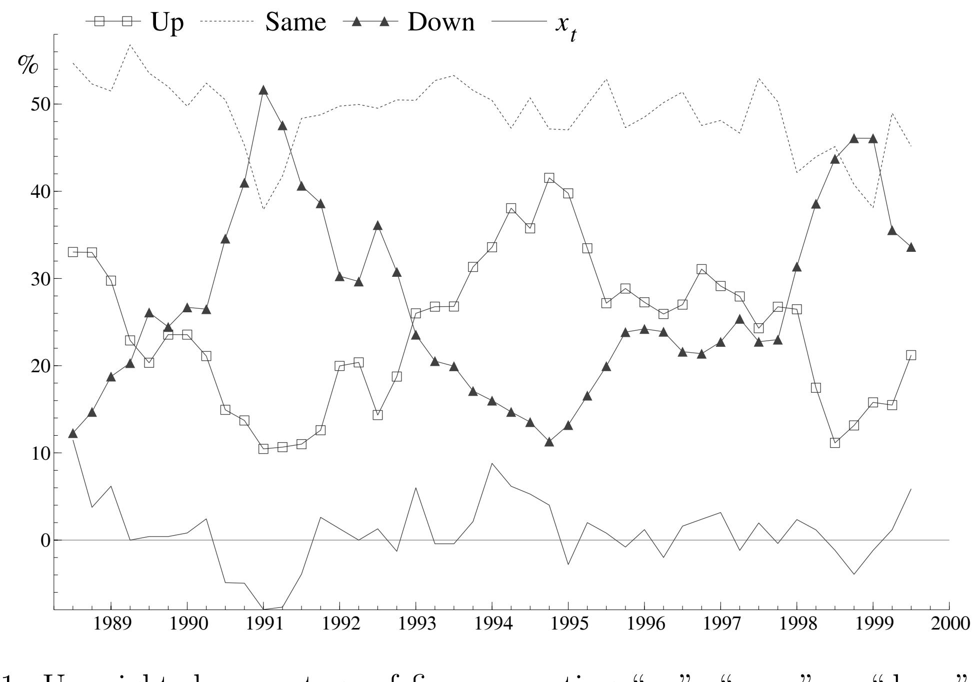

Figure 1: Unweighted percentage of firms reporting "up", "same" or "down" alongside (aggregate) manufacturing output growth x_t

4 Indicators of Sectoral Output Growth

The indicator D_t (2.11) (and $\overline{D}_t$ (2.15)) requires that the relationship (2.5) between the qualitative survey responses and the output growth rate x_t be correctly specified. As detailed in section 2.2 and Appendix A, various specification tests should be conducted in order to establish and validate the preferred nature and form of this statistical relationship. Additional concerns are whether the model is best specified at the sectoral rather than the aggregate level, with, or without, homogeneity restrictions imposed, and which set of additional variables z_t should be included. Consequently the model used as the basis for the indicator D_t (and $\overline{D}_t$) may vary reflecting the statistical properties of the specific datasets employed.

Preliminary estimation of static firm-level models, (2.2), relating the CBI qualitative data solely to aggregate manufacturing output growth, x_t, indicated violation of the independence assumption of ξ_{it} across i required for D_t (but not $\overline{D}_t$); including lags of x_t did not ameliorate this dependence. Furthermore, augmenting the model with the proportions of optimistic and pessimistic firms, i.e., cross-sectional averages of those firms reporting "up" and "down" computed from the CBI survey, as additional variables, z_t, still resulted in a strong rejection of cross-sectional independence.

As a result we considered models specified at the sectoral level. The models now relate firms' categorical responses to the requisite sectoral output growth rate, which we persist in denoting as x_t even though it differs across the sectors. Seven sectoral definitions were considered: (i) Food, Drink and Tobacco; (ii) Chemicals; (iii) Engineering; (iv) Motor Vehicles; (v) Metals; (vi) Textiles and (vii) Other. The additional variables, z_t, were the proportions of optimistic and pessimistic firms.

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

Ordered probit models based on the static formulation (2.2) for each of the 834 firms were estimated at the sectoral level. (An additional 27 firms were dropped because the ML estimation routine failed to converge.) These firm-level models were subjected to the specification tests described in Appendix A and are generally supportive of this specification. Both the test of the statistical significance of x_{t-1} in the firm-level model (2.5), $\rho_i = 0$, $(i = 1, ..., N_t)$, and a score test for misspecification are considered. The score test for misspecification is a joint test for incorrect functional form, based on the omitted variables x_{t-1} and powers of $\widehat{\beta}_i x_t$, conditional heteroskedasticity and the normality of the error terms ξ_{it}; see Machin & Stewart (1990).[7] Table 1 reports the proportion of rejec-

[7] A score test for the conditional independence of the error term ξ_{it} in (2.5) and $\{y_{i\tau}\}_{\tau=1}^{t-1}$ given$\{x_\tau, z_\tau\}_{\tau=1}^{t}$ and i, $(t = 1, ..., T)$, may be based on a test of the significance of lagged generalized residuals; cf. Appendix A. See Gourieroux et al. (1985). Given the highly unbalanced nature of our

tions of the null hypothesis across firms, i, $(i = 1, ..., N_t)$, at a 0.05 significance level. To mitigate the effects of an inflated Type I error when testing across i, $(i = 1, ..., N_t)$, the proportion of rejections using Bonferroni adjusted critical values is also reported.

Table 1: Specification tests for ordered discrete choice models. Proportion of times the specification tests were not rejected and p-values for the CD test

Sector	$\rho_i = 0$		Score		CD
	Individual	Bonferroni	Individual	Bonferroni	p-value
Food, Drink & Tobacco	0.89	1.00	0.87	0.97	0.09
Chemicals	0.98	1.00	0.81	1.00	0.28
Engineering	0.93	1.00	0.89	0.99	0.01
Motor Vehicles	1.00	1.00	0.94	0.97	0.15
Metals	0.93	1.00	0.92	0.99	0.41
Textiles	0.92	1.00	0.89	1.00	0.34
Other	0.87	1.00	0.85	0.99	0.35

A Wald test (fixed N_t) again rejected the null hypothesis $\beta_i = \beta$ for all i with a p-value of 0.00 in each of the seven sectors. Thus, subject to the identifying assumption $\sigma^2_{\xi i} = 1$, firms appear to be heterogeneous in terms of how they react to changes in the sectoral environment. Some firms become more optimistic as the sectoral growth rate x_t increases while others, perhaps because of the nature of the business they run, become more pessimistic; others hardly react to x_t. Table 2 provide some evidence on the dependence of y_{it} on x_t and gives an impression of this heterogeneity, displaying the number of firms that have t-ratios for testing $\beta_i = 0$ in specified ranges; firms are sorted by their industrial sector.[8]

Table 2 reveals considerable variation across firms in how their survey responses relate to sectoral output growth rates. As many firms' qualitative replies are negatively related to sectoral output growth as are positively related. This is a consequence of including the proportion of optimistic and pessimistic firms in z_t. When z_t is excluded, there is a clear preponderance of positive t-ratios for all sectors, although the CD test then rejected cross-sectional independence for each of the seven sectors; this contrasts the CD results, when z_t is included, presented in Table 1. As discussed above, our D_t indicator is designed

panel, however, with limited consecutive observations for a given firm, we maintain this assumption. As discussed in footnote 3 above, our proposed indicator remains consistent using quasi-ML if ξ_{it} is standard normally distributed conditional on $\{x_\tau, z_\tau\}_{\tau=1}^t$ and i, $(t = 1, ..., T)$.

[8]These results are predicated on the assumption that the error term ξ_{it} in (2.5) is conditionally independent of $\{y_{i\tau}\}_{\tau=1}^{t-1}$ given $\{x_\tau, z_\tau\}_{\tau=1}^t$ and i and is standard normally distributed, $(t = 1, ..., T)$. An extension to ordered categorical data as considered here of the test for independence of Pesaran & Timmermann (2009) robust to serial correlation in ξ_{it}, $(t = 1, ..., T)$, would be a useful avenue for future research.

Table 2: t-ratios for $\hat{\beta}_i$: The Number of Firms in Specified Ranges with Firms Sorted by Industrial Sector

Sector	t-ratio (t_i)					
	$t_i \leq -2$	$-2 < t_i \leq -1$	$-1 < t_i \leq 0$	$0 < t_i \leq 1$	$1 < t_i \leq 2$	$t_i > 2$
Food, Drink & Tobacco	3	3	13	17	2	0
Chemicals	2	7	20	19	2	3
Engineering	9	34	92	85	43	4
Motor	1	2	15	10	5	1
Metals	2	13	34	46	11	5
Textiles	4	19	41	42	8	1
Other	8	25	56	72	20	8

precisely to address this heterogeneity across firms.

4.2 Density Function $f(\cdot)$

It remains to specify $f(\cdot|\{x_\tau\}_{\tau=1}^{t-1}, \{z_\tau\}_{\tau=1}^{t})$ for each of the seven sectors. The assumption that x_t is stationary is supported by tests for a unit root in the level series of sectoral output. It is well known that macroeconomic time-series often exhibit structural instabilities or breaks. Thus we should not expect the persistence (or the conditional variance) of these seven series to be time-invariant. However, rather than condition our indicator D_t on a model estimated for x_t over a specific estimation window to reflect their presence, we focus below on illustrating the utility of our indicator in the unconditional case, i.e., when $\beta_x = 0$ and $\gamma_x = 0$. Figure 1, to be discussed in more detail below, suggests that this assumption may not be unreasonable, with considerable volatility displayed in many sectors. We do, though, also consider below the performance of D_t when based on an $AR(1)$ specification for x_t with β_x estimated over the 1988q3-1999q3 sample period. A modified version of the Jarque-Bera test, robust to serial correlation and conditional heteroscedasticity in x_t [see Bai & Ng (2005) and Bontemps & Meddahi (2005)], does not reject the normality of $f(\cdot)$ with p-values of 0.53, 0.21, 0.58, 0.75, 0.27, 0.15 and 0.07 for each of the seven sectors in turn (as listed in Tables 1 and 2).

4.3 Indicator Performance

We first compare the sector-by-sector performance of the indicators D_t and $\overline{D}_t$ against that of four traditional quantification techniques employed on aggregated proportions: the balance statistic [BAL], the probability method of Carlson & Parkin (1975) [CP], the regression approach of Pesaran (1984, 1987) [P] and the reverse-regression approach of Cunningham et al. (1998) [CSW] based on the logistic distribution; Appendix C presents a brief review of these various quantification methods. An assessment of their performance at the aggregate level for manufacturing output growth then follows in Section 4.3.1. The indicators D_t and $\overline{D}_t$ are based on firm-level probit model estimation. Although the homogeneity assumptions are rejected in the sample we compute a random effects-type model indicator D_t [RE] assuming slope homogeneity, $\beta_i = \beta$, $(i = 1, ..., N)$, and evaluate its performance. Finally, as a benchmark, and as a means of assessing the utility of the qualitative survey data, we examine the performance of a pure $AR(1)$ model for x_t.

Table 3 summarises the performance of the indicators for each of the seven sectors. Figure 2 provides a visual impression of the relative performance of the indicators, focusing on BAL as the representative aggregate indicator. It is clearly seen that BAL is too smooth, and unable to pick up the volatility, and for some sectors, business cycle fluctuations, in sectoral output growth. Table 3 reveals that the new indicators provide more accurate early estimates of output growth than all of the traditional indicators employed on the aggregate proportions as well as the $AR(1)$ benchmark, which tends to perform a little worse than the aggregate indicators. Regardless of how the disaggregate indicators are scaled, the higher correlation of the D_t and $\overline{D}_t$ indicators indicates that a stronger signal about the official data may be recovered from them than the aggregate data.

The indicator D_t considered in Table 3 assumes an unconditional prior density function $f(\cdot)$ for x_t, i.e., x_t is conditionally independent of z_t and $\{x_\tau, z_\tau\}_{\tau=1}^{t-1}$. When based on the conditional prior density $f(\cdot|x_{t-1})$ using the $AR(1)$ model (2.3) with $\gamma_x = 0$ and with β_x estimated rather than set to zero, the performance of the indicator D_t deteriorates. RMSEs for each of the seven sectors in turn, with the corresponding RMSE estimate from Table 3 in parentheses, are 4.59 (4.12), 5.11 (3.26), 5.47 (2.33), 10.44 (9.90), 4.39 (3.34), 5.36 (3.28) and 4.03 (3.34). These results are consistent with conditional independence between x_t and x_{t-1} indicating that the use of the estimated conditional density rather than the marginal or unconditional density is inefficient.

Table 3 emphasised the importance of basing the indicator D_t on firm-level estimation of (2.2). The indicator D_t [RE] obtained from a random effects formulation performs considerably worse, exhibiting little or no correlation against the outturn, x_t, being explained by β being estimated as zero (to more than three decimal places). Allowing heterogeneous slope coefficients β_i, cf. Table 2, with firm-level ML estimation of the probit model specification, the performance of D_t is much improved. This, of course, is the rationale for our indicator which gives a greater emphasis to those firms whose responses have a close link to the official data than to those whose experiences correspond only weakly or not at all.

The two indicators D_t and $\overline{D}_t$ exhibit a similar correlation against the outturn for manufacturing output growth. However, D_t performs better than $\overline{D}_t$ on the basis of the root mean squared error [RMSE] criterion. Despite the sample mean of $\overline{D}_t$ approximately estimating that of the outcomes x_t correctly, it appears too smooth and thus displays too little volatility as compared with the outturn x_t; see Figure 2. This latter feature has been observed elsewhere for alternative indicators; see, e.g., Cunningham (1997).

Less volatility is also observed because the scale is incorrect which may be explained by consideration of those firms whose responses are poorly correlated with actual output growth. In the extreme case of no correlation, inclusion of these firms reduces the standard deviation of the $\overline{D}_t$ indicator but leaves its correlation with output growth unaffected because in a large time-series, if a firm responds at random the firm-level disaggregate method gives the same score (mean output growth) to all categorical responses, i.e., $E[x_t | j_{it}, z_t, \{x_\tau, z_\tau\}_{\tau=1}^{t-1}, i] = E[x_t]$. For these firms therefore there is no contribution to the variance of $\overline{D}_t$. Excess smoothness of $\overline{D}_t$ may thus be viewed as due to the presence of firms in the sample whose responses contain little or no signal about output growth. However, D_t does not suffer from this problem since, as indicated above, it is designed to give more weight to firms whose answers have a close link to the official data than to those whose experiences correspond only weakly or not at all. Therefore, while D_t has a similar, indeed slightly improved, correlation against x_t, it is not too smooth. D_t better picks up the scale of x_t evidenced by a higher standard deviation and lower RMSE than $\overline{D}_t$.

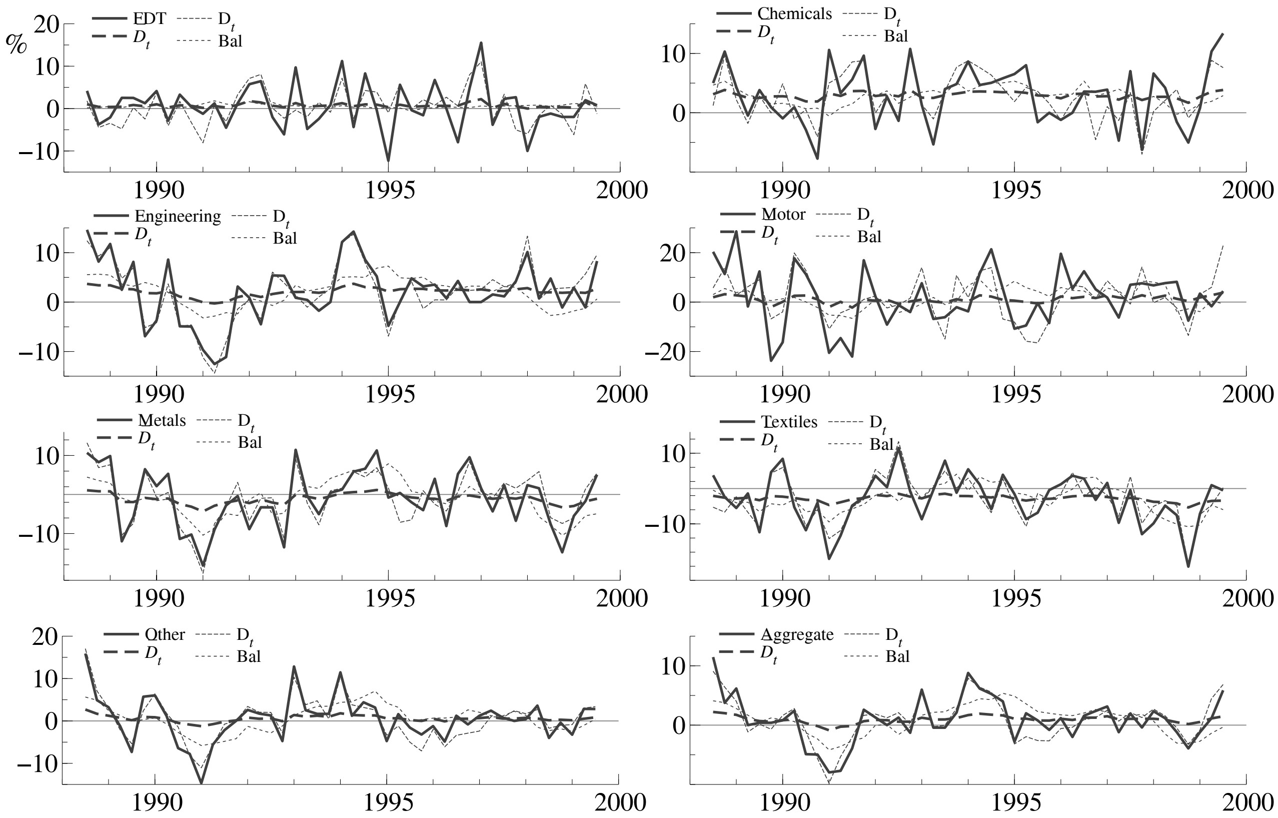

Figure 2: In-Sample Comparison of D_t, $\overline{D}_t$ and BAL against Official Sectoral and Aggregate Output Growth x_t

Table 3: Indicator Performance

	Mean	SD	Corr.	RMSE		Mean	SD	Corr.	RMSE
	Food, Drink & Tobacco					Chemicals			
D_t	0.40	4.26	0.64	4.12	D_t	2.76	3.99	0.76	3.26
$\overline{D}_t$	0.55	0.32	0.60	5.09	$\overline{D}_t$	2.84	0.37	0.76	4.70
D_t [RE]	0.42	0.04	0.01	5.28	D_t [RE]	2.70	0.02	0.10	4.97
CP	0.57	1.54	0.09	5.62	CP	2.86	60.69	0.32	58.59
PES	0.57	0.51	0.10	5.26	PES	2.86	1.54	0.31	4.73
BAL	0.57	0.45	0.08	5.26	BAL	2.86	1.54	0.31	4.73
CSW	0.57	62.11	0.09	61.19	CSW	2.86	17.55	0.29	16.63
AR	0.57	1.20	0.23	5.15	AR	2.86	0.27	0.05	4.96
x_t	0.57	5.34			x_t	2.86	5.03		
NT	308				NT	399			
N	38				N	50			
	Engineering					Motor Vehicles			
D_t	2.14	6.16	0.93	2.33	D_t	1.87	9.20	0.57	9.90
$\overline{D}_t$	2.09	0.36	0.91	5.85	$\overline{D}_t$	1.86	0.83	0.52	11.34
D_t [RE]	1.53	0.11	0.02	6.19	D_t [RE]	1.26	0.16	0.04	11.78
CP	2.08	29.03	0.44	31.93	CP	1.82	42.78	0.36	47.85
PES	2.08	2.87	0.46	5.48	PES	1.82	4.19	0.35	11.01
BAL	2.08	2.87	0.46	5.48	BAL	1.82	4.12	0.35	11.03
CSW	2.08	12.94	0.48	11.20	CSW	1.82	35.99	0.33	33.59
AR	2.08	2.67	0.43	5.58	AR	1.82	2.10	0.18	11.58
x_t	2.08	6.25			x_t	1.82	11.89		
NT	1849				NT	279			
N	260				N	32			
	Metals					Textiles			
D_t	-0.91	6.74	0.89	3.34	D_t	-2.98	6.23	0.88	3.28
$\overline{D}_t$	-1.05	0.456	0.88	6.93	$\overline{D}_t$	-2.99	0.38	0.86	6.65
D_t [RE]	-0.23	0.24	0.43	7.48	D_t [RE]	-2.66	0.10	0.06	6.97
CP	-1.07	4.11	0.63	5.74	CP	-3.00	5.13	0.57	5.85
PES	-1.07	4.67	0.63	5.71	PES	-3.00	4.29	0.61	5.53
BAL	-1.07	4.66	0.63	5.71	BAL	-3.00	3.82	0.54	5.86
CSW	-1.07	11.80	0.63	9.07	CSW	-3.00	12.29	0.57	9.95
AR	-1.07	2.10	0.28	7.04	AR	-3.00	1.98	0.28	6.69
x_t	-1.07	7.42			x_t	-3.00	7.05		
NT	680				NT	643			
N	115				N	121			
	Other								
D_t	0.57	5.13	0.92	2.08					
$\overline{D}_t$	0.55	0.37	0.90	4.96					
D_t [RE]	0.50	0.01	0.15	5.29					
CP	0.54	2.98	0.59	7.40					
PES	0.54	3.15	0.59	4.27					
BAL	0.54	3.09	0.58	4.31					
CSW	0.54	9.17	0.58	7.36					
AR	0.54	1.60	0.30	5.04					
x_t	0.54	5.34							
NT	1251								
N	191								

Notes: N denotes the overall number of different firms, with at least twenty time-series observations, used to compute the indicators. NT denotes the total number of panel-data observations available for these firms.

4.3.1 Aggregate Performance

We have focused on a sectoral analysis because, as noted early in section 4, estimation of the model without distinguishing between sectors led to cross-sectional dependence in the model residuals ξ_{it}, $(i = 1, ..., N_t)$, $(t = 1, ..., T)$. However, users of our indicator are likely to be at least as interested in estimates for aggregate manufacturing growth than as its fit for the component sectors.

Aggregate growth rates are computed by National Accounts statisticians as the weighted average of the sectoral growth rates, the weights being the sectoral shares of value-added observed in the previous year. Accordingly, we use the shares of a given sector in total manufacturing value-added in the previous year to aggregate our proposed sectoral indicators, D_t, $\overline{D}_t$ and D_t [RE] for the current year. We compare their performance to those of the aggregate indicators, CP_t, PES_t, BAL_t, CSW_t and AR_t, calculated directly from the aggregate survey data for manufacturing and/or the index of output for the manufacturing sector as a whole. We present in Table 4 evaluation statistics comparable to those for the sectoral indicators seen in Table 3. The bottom right panel of Figure 2 also provides a visual impression of the relative performance of the indicators, again focusing on BAL as the representative aggregate indicator.

Table 4: Aggregate Indicator Performance

	Mean	SD	Corr.	RMSE
D_t	0.98	3.66	0.92	1.75
$\overline{D}_t$	0.95	0.61	0.88	3.69
D_t [RE]	0.84	0.04	-0.31	4.23
CP	0.93	5.02	0.64	7.97
PES	0.93	2.50	0.65	2.90
BAL	0.93	2.47	0.64	2.93
CSW	0.93	5.95	0.65	4.48
AR	0.93	1.86	0.48	3.34
x_t	0.93	3.85		

These results in Table 4 confirm the good sectoral performance of our indicator D_t. Not surprisingly, an approach that works well with sectoral data also displays a good performance when the sectoral indicators are aggregated to produce an indicator of aggregate output growth.

4.4 Confidence Intervals

The above results and discussion have been concerned solely with point estimates for sectoral and aggregate growth rates.

It is also informative to provide interval estimates for output growth. We briefly describe one possible simulation scheme to do so. Let $\hat{\theta}_i$ denote the ML estimator for θ_i where $\theta_i = (\{\mu_{ji}\}_{j=0}^{3}, \alpha_i, \beta_i, \gamma_i')'$, $(i = 1, ..., N)$. Given $\hat{\theta}_i$, $(i = 1, ..., N)$, and $\{x_\tau, z_\tau\}_{\tau=1}^{T}$, and a T-vector random draw from the standard normal distribution, since $\varepsilon_{it} \sim N(0,1)$, generate the indicators j_{it}^r, $(i = 1, ..., N_t)$, $(t = 1, ..., T)$. Calculate ML estimators $\hat{\theta}_i^r$, $(i = 1, ..., N_t)$, and thus feasible indicators $\hat{D}_t^r$, $(t = 1, ..., T)$. Repeat this sequence R times. The empirical distribution function of $\hat{D}_t^r - \hat{D}_t$, $(r = 1, ..., R)$, then provides an asymptotically valid approximation to the distribution of $\hat{D}_t - x_t$, $(t = 1, ..., T)$. Appendix B provides an analytical expression for the asymptotic variance of $\hat{D}_t - D_t$ that could be used to give some indication of how estimation error is likely to affect the estimator $\hat{D}_t$ relative to the infeasible indicator D_t.

Our results indicate no statistically significant cross-sectional dependence between error terms for firms in any one sector and those of any other sector. Recall that there is no evidence of significant cross-sectional dependence of error terms within sectors. Hence, the variance of the aggregate indicator may be estimated by the sum of appropriately weighted estimated variances of the sectoral indicators, the weights being the squares of the aforementioned value-added shares. Its square root is then used to calculate 90% confidence intervals for the aggregate indicator.

Figure 3 implements the above simulation scheme with the number of replications R set at 500 and plots the feasible indicator $\hat{D}_t$ for each sector and that for the aggregate indicator, together with approximate 90% confidence intervals.

Figure 3 clearly indicates the uncertainty concerning both sectoral and manufacturing output growth in the indicator $\hat{D}_t$. It also provides some idea of the influence of the contribution of estimation error in $\hat{D}_t$ over and above the sampling error in D_t. The confidence bands displayed in Figure 3 often indicate uncertainty about the sign of sectoral output growth, cf. Food, Drink & Tobacco, that of aggregate manufacturing output growth.

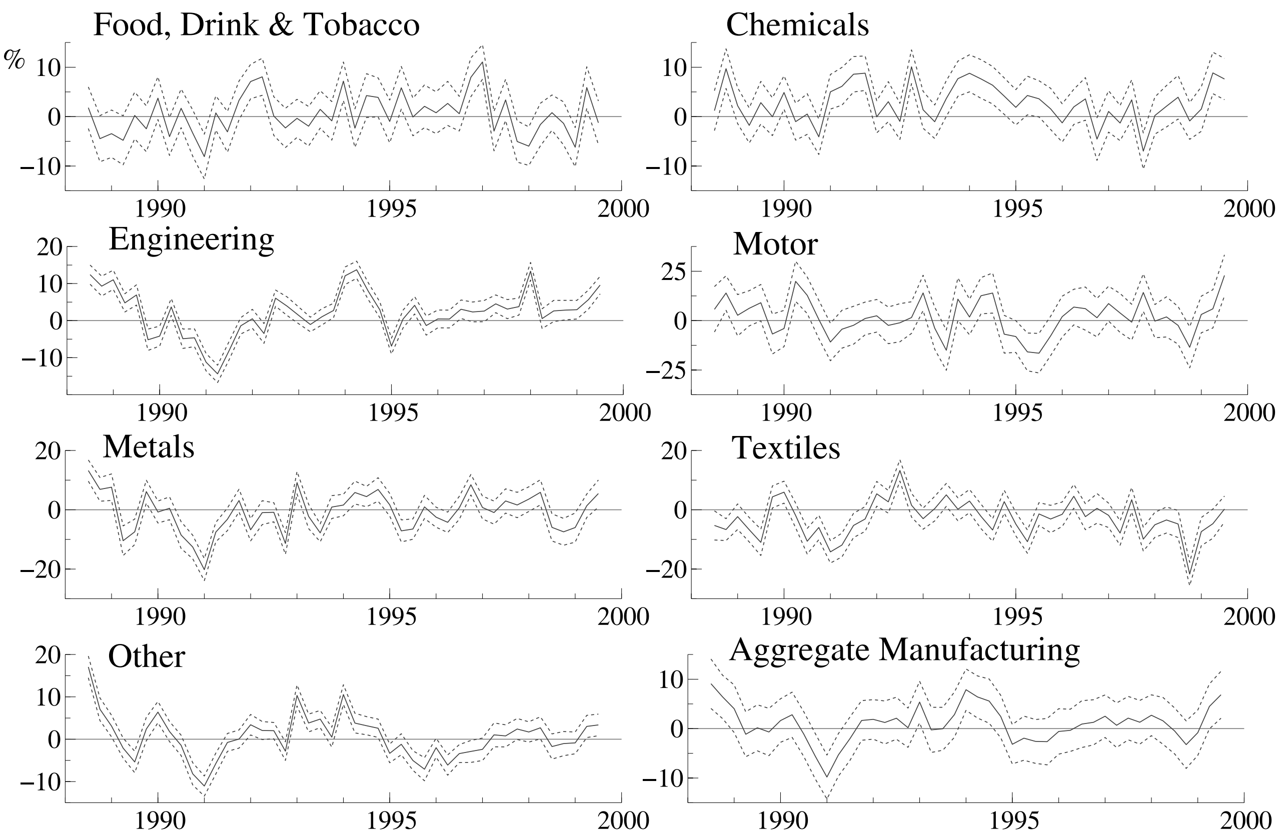

Figure 3: Confidence intervals around $\widehat{D}_t$ at the sectoral and aggregate level

5 Concluding Comments

This paper develops an efficient means of extracting a quantitative signal about the business environment from qualitative survey data. The approach is statistically coherent, being derived from an application of Bayes' Theorem to a statistical model for individual qualitative responses to the survey. Unlike methods based on aggregate data it takes account of the relative informational content of each individual survey response. From a practical perspective an improved means of extracting the underlying signal from qualitative categorical data ahead of the publication of official data should mean that economic policy setting can be undertaken with more confidence. The method developed is applicable to other qualitative surveys. In addition our approach could be adapted to address questions on expected future output growth.

In an in-sample application to survey data from the CBI, the proposed indicator outperformed traditional indicators in terms of anticipating movements to sectoral output growth. This satisfactory performance is also mirrored in that of the indicator of aggregate manufacturing output growth constructed from these sectoral indicators. Out of sample testing is possible only when a panel data set with a longer time-series dimension becomes available; this research will be undertaken in future work, since the time dimension of the quarterly CBI observations available since 2000, when the processing platform at the CBI changed, will soon exceed that used in this study. But the satisfactory performance of the indicator set out above, and that under the stated assumptions it offers an efficient means of aggregating qualitative survey data, suggests this should be a worthwhile exercise. Of course, in practice, because of estimation error, found to be significant here, and structural instabilities, it may be the case that the equal-weighted indicator $\overline{D}_t$ outperforms the weighted indicator D_t. Similar outturns have been found when combining quantitative (point) forecasts; see Timmermann (2006) for a recent survey.

Qualitative survey data are often collected by non-government bodies, by, e.g., the CBI in the UK and the Conference Board in the US, and are generally publicly available only in aggregate form. Perhaps the importance of the associated microeconomic-level survey data demonstrated in this paper may facilitate an improvement in the availability of such data.

Acknowledgements

We wish to thank J. Morrison and J. Wood of the CBI for their advice about the survey data. We are also grateful to the Co-Editor Manuel Arellano and three anonymous referees for their helpful comments. Earlier versions of this paper and its predecessor have been presented at the Universities of Cambridge, Exeter and Melbourne, the Office for National Statistics, H.M. Treasury, the Reserve Bank of New Zealand, the 2006 Royal Economic Society Conference, University of Nottingham, and the 2008 "Time-Series and Panel Modelling" Conference in honour of M.H. Pesaran at Goethe-Universität Frankfurt. Financial support from the ESRC (RES-000-22-1009) and a 2002 Leverhulme Major Research Fellowship is also gratefully acknowledged.

A Appendix A: Specification Tests

A test of $\rho_i = 0$, $(i = 1, ..., N_t)$, or the exclusion of the error term u_t in (2.5) jointly tests for the absence of dynamics and the weak exogeneity of x_t in (2.2). A simple two-step test of $\rho_i = 0$ may be formulated similarly to the procedures described in Smith & Blundell (1986) and Newey (1987). Firstly, (2.3) is estimated by least squares (LS) which yields the consistent estimates $(T \to \infty)$, $\hat{\alpha}_x$, $\hat{\beta}_x$ and $\hat{\gamma}_x$ and the LS residual $\hat{u}_t = x_t - \hat{\alpha}_x - \hat{\beta}_x x_{t-1} - \hat{\gamma}_x' z_t$, $(t = 1, ..., T)$. Secondly, the augmented model (2.5) is estimated by ordered Probit as in section 2 after substitution of $\hat{u}_t$ for u_t. Finally, the hypothesis $\rho_i = 0$ is then assessed by a standard ordered Probit t-test based on the resultant estimate of ρ_i. Failure to reject $\rho_i = 0$ supports the use of (2.2) while its rejection implies that the official data should be inferred using the augmented conditional model (2.5); see section 2.3 above.

Score or Lagrange multiplier tests for the implicit assumptions of linearity, conditional homoskedasticity and that the error term ε_{it} is normally distributed appropriate for the use of ordered Probit are employed to ascertain the empirical validity of (2.5); see, e.g., Chesher & Irish (1987) and Machin & Stewart (1990).

The cross-sectional independence of ε_{it}, $(t = 1, ..., T)$, can be tested using the test proposed by Hsiao et al. (2011) adapted for use with nonlinear panel data models; *viz.*

$$CD = \sqrt{\frac{2}{N(N-1)}} \sum\nolimits_{i=1}^{N-1} \sum\nolimits_{k=i+1}^{N} \sqrt{T_{ik}} \hat{r}_{ik}$$

where T_{ik} is the number of time-series observations when qualitative survey responses are available for both firms i and k and $\hat{r}_{ik}$ is the pair-wise sample correlation coefficient between the estimated generalised residuals $E[y_{it} - \alpha_i^* - \beta_{i0}^* x_t - \beta_{i1}^* x_{t-1} - \beta_{i2}^{*\prime} z_t | y_{it}^j, \{x_\tau, z_\tau\}_{\tau=1}^t, i]$ and $E[y_{kt} - \alpha_k^* - \beta_{k0}^* x_t - \beta_{k1}^* x_{t-1} - \beta_{k2}^{*\prime} z_t | y_{kt}^j, \{x_\tau, z_\tau\}_{\tau=1}^t, k]$ obtained from the ordered Probit models for firms i and k, see Gourieroux et al. (1987), where $E[\cdot | y_{it}^j, \{x_\tau, z_\tau\}_{\tau=1}^t, i]$

denotes the conditional expectation operator under the null hypothesis of cross-sectional independence. Under cross-sectional independence, $CD \xrightarrow{d} N(0,1)$; cf. Hsiao et al. (2011).

B Appendix B: Estimation Error

For simplicity this exposition ignores the presence of dynamics and additional variables z_t.

Write the conditional probability $P(j_{it}|\{x_\tau\}_{\tau=1}^t, i)$ of j_{it} given $\{x_\tau\}_{\tau=1}^t$ and i as $P(j_{it}|\{x_\tau\}_{\tau=1}^t; \theta_i)$ where θ_i summarises the unknown parameters for the ith firm, $(i = 1, ..., N_t)$. Correspondingly, the estimator $\hat{P}(j_{it}|\{x_\tau\}_{\tau=1}^t, i)$ for $P(j_{it}|\{x_\tau\}_{\tau=1}^t, i)$ is written as $P(j_{it}|\{x_\tau\}_{\tau=1}^t, \hat{\theta}_i)$ where $\hat{\theta}_i$ is the ML estimator for θ_i, $(i = 1, ..., N_t)$. Thus, the estimator $\hat{P}(\{j_{it}\}_{i=1}^{N_t}|\{x_\tau\}_{\tau=1}^{t-1}; \{\theta_i\}_{i=1}^{N_t})$ for $P(\{j_{it}\}_{i=1}^{N_t}|\{x_\tau\}_{\tau=1}^{t-1}; \{\theta_i\}_{i=1}^{N_t})$ is written as

$$P(\{j_{it}\}_{i=1}^{N_t}|\{x_\tau\}_{\tau=1}^{t-1}; \{\hat{\theta}_i\}_{i=1}^{N_t}) = \int_{-\infty}^{\infty} \prod_{i=1}^{N_t} P(j_{it}|\{x_\tau\}_{\tau=1}^{t}; \hat{\theta}_i) f(x_t|\{x_\tau\}_{\tau=1}^{t-1}) dx_t.$$

For ease of exposition we drop the macroeconomic conditioning information $\{x_\tau\}_{\tau=1}^{t-1}$ and the index t, cf. section 2.3, and provide an analysis for scalar θ_i, $(i = 1, ..., N)$, which may straightforwardly, but at the expense of more complex notation, be extended to the vector case. Let $\hat{\theta}$ denote the ML estimator of θ where θ collects together θ_i, $(i = 1, ..., N)$. The large sample distribution of the ML estimator is given by $T^{1/2}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \mathcal{I}^{-1})$ where $\mathcal{I}$ denotes the (asymptotic) information matrix. Let $\hat{\mathcal{I}}$ denote a consistent estimator for the information matrix $\mathcal{I}$ and $\hat{\imath}^{ij}$ the (i,j)th element of the inverse of the estimated information matrix $(\hat{\mathcal{I}})^{-1}$. The feasible indicator $\hat{D}$ is then defined as

$$\begin{aligned}
\hat{D} &= E[x|\{j_i\}_{i=1}^N; \{\hat{\theta}_i\}_{i=1}^N] \\
&= \int_{-\infty}^{\infty} x f(x|\{j_i\}_{i=1}^N; \{\hat{\theta}_i\}_{i=1}^N) dx \\
&= \int_{-\infty}^{\infty} x \frac{P(\{j_i\}_{i=1}^N|\{\hat{\theta}_i\}_{i=1}^N) f(x)}{P(\{j_i\}_{i=1}^N|\{\hat{\theta}_i\}_{i=1}^N)} dx \\
&= \int_{-\infty}^{\infty} x \frac{\prod_{i=1}^{N_t} P(j_i|\hat{\theta}_i)}{\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|\hat{\theta}_i) f(x) dx} f(x) dx.
\end{aligned}$$

with the infeasible index

$$D = \int_{-\infty}^{\infty} x \frac{\prod_{i=1}^{N} P(j_i|x;\theta_i)}{\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x) dx} f(x) dx.$$

Let $P_{\theta_i}(j_i|x;\theta_i) = \partial P(j_i|x;\theta_i)/\partial\theta_i$, $(i = 1, ..., N)$. A Taylor expansion of $\hat{D}$ about θ_i, $(i = 1, ..., N)$, yields

$$\begin{aligned}
\hat{D} - D \quad = \quad & \frac{1}{(\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx)^2} \\
& \times \sum_{i=1}^{N} \Big[\Big(\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx \Big) \Big(\int_{-\infty}^{\infty} x \prod_{k=1,k\neq i}^{N} P(j_k|x;\theta_k) P_{\theta_i}(j_i|x;\theta_i) f(x)dx \Big) \\
& - \Big(\int_{-\infty}^{\infty} x \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx \Big) \Big(\int_{-\infty}^{\infty} \prod_{k=1,k\neq i}^{N} P(j_k|x;\theta_k) P_{\theta_i}(j_i|x;\theta_i) f(x)dx \Big) \Big] \\
& \times (\hat{\theta}_i - \theta_i) \\
& + O_p(\max_{i=1,\ldots,N} \left\| \hat{\theta}_i - \theta_i \right\|^2).
\end{aligned}$$

An estimator for the variance of $\hat{D} - D$ is given by substituting $\hat{\theta}_i$ for θ_i, $(i = 1, ..., N)$, in

$$\begin{aligned}
& \frac{T^{-1}}{(\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx)^4} \\
& \times \sum_{i=1}^{N} \sum_{j=1}^{N} \Big[\Big(\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx \Big) \Big(\int_{-\infty}^{\infty} x \prod_{k=1,k\neq i}^{N} P(j_k|x;\theta_k) P_{\theta_i}(j_i|x;\theta_i) f(x)dx \Big) \\
& - \Big(\int_{-\infty}^{\infty} x \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx \Big) \Big(\int_{-\infty}^{\infty} \prod_{k=1,k\neq i}^{N} P(j_k|x;\theta_k) P_{\theta_i}(j_i|x;\theta_i) f(x)dx \Big) \Big] \\
& \times \Big[\Big(\int_{-\infty}^{\infty} \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx \Big) \Big(\int_{-\infty}^{\infty} x \prod_{k=1,k\neq j}^{N} P(j_k|x;\theta_k) P_{\theta_i}(j_j|x;\theta_j) f(x)dx \Big) \\
& - \Big(\int_{-\infty}^{\infty} x \prod_{i=1}^{N} P(j_i|x;\theta_i) f(x)dx \Big) \Big(\int_{-\infty}^{\infty} \prod_{k=1,k\neq j}^{N} P(j_k|x;\theta_k) P_{\theta_i}(j_j|x;\theta_j) f(x)dx \Big) \Big]^{-1} \hat{\imath}^{ij}.
\end{aligned}$$

C Appendix C: Aggregate Quantification Techniques

This appendix reviews four alternative quantification methods: the balance statistic and the probability approach of Carlson & Parkin (1975); the regression approach of Pesaran (1984, 1987); the reverse-regression approach of Cunningham et al. (1998) and Mitchell et al. (2002). Although motivated in different ways, these approaches are shown to share a common foundation. Our discussion compares the latter two methods to the probability approach and draws on Pesaran (1987) and Mitchell et al. (2002). For alternative reviews and extensions of the probability and regression approaches, see Pesaran & Weale (2006).

Let U_t and D_t denote the proportion of firms that report an output rise and fall.

C.1 The Balance Statistic and the Probability Approach

The "balance statistic" $U_t - D_t$ [Anderson (1952)], up to scale, 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.Theil (1952) provides a motivation for this approach based on the probability approach.

The probability method of quantification assumes that the response of firm i concerning x_t is derived from a subjective probability density function for x_t, $f_i(\cdot|i)$, which may differ in form across firms and is conditional on information available to firm i at time t; the dependence of $f_i(\cdot|i)$ on t is suppressed in the discussion. Denote the mean of $f_i(\cdot|i)$ by $x_{it} = \int x f_i(x|i)dx$.

The responses of firm i are classified as follows: "up" is observed 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 $a_{it}, b_{it} > 0$.

Assume that firms are independent and that $f_i(\cdot|i)$ is the same and known for all firms, i.e., $f_i(\cdot|i) = f(\cdot|i)$. Consequently, $x_{it} = \int x f(x|i)dx$ can be regarded as an independent draw from an aggregate density $f(x) = \int f(x|i)F(di)$, where $F(\cdot)$ denotes the distribution function of firms i; the density $f(\cdot)$ is conditional on aggregate information available to all firms at time t, the dependence on which is again suppressed. Assume $f(\cdot)$ has mean x_t.

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

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

where $F_t(\cdot)$ is the cumulative distribution function obtained from $f(\cdot)$ where, now, we indicate explicitly the dependence on time t. 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 $F_t(\cdot)$.

C.1.1 Carlson and Parkin's Method

Carlson & Parkin (1975) assumes that $f(\cdot)$ is a normal density function with mean x_t and variance σ_t; alternative densities are considered in, e.g., Batchelor (1981) and Mitchell (2002). From (C.1), the estimator for x_t is given as the solution to the equations

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

where $\Phi(\cdot)$ is the $N(0,1)$ c.d.f. Solving (C.2)

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

and thus

$$\hat{x}_t = \lambda\left(\frac{\Phi^{-1}(1-U_t) + \Phi^{-1}(D_t)}{\Phi^{-1}(1-U_t) - \Phi^{-1}(D_t)}\right), \tag{C.3}$$

where $\Phi^{-1}(\cdot)$ is the inverse function. The scale parameter λ remains to be determined. Carlson & Parkin (1975) invoke unbiasedness over the sample period, $(t = 1, ..., T)$, i.e.,

$$\hat{\lambda} = \left(\sum\nolimits_{t=1}^{T} x_t\right) / \sum\nolimits_{t=1}^{T}\left(\frac{\Phi^{-1}(1-U_t) + \Phi^{-1}(D_t)}{\Phi^{-1}(1-U_t) - \Phi^{-1}(D_t)}\right). \tag{C.4}$$

C.2 The Regression Approach

Suppose that aggregate output x_t is a weighted average of the sample of firms' outputs x_{it}, $(i = 1, ..., N_t)$, *viz.*

$$x_t = \sum\nolimits_{i=1}^{N_t} w_i x_{it}, \tag{C.5}$$

Categorising firms according to whether they reported an "up" (+) or a "down" (−), (C.5) can be rewritten as

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

where the unobserved $x_{it}^+ = x_{it}$ if "up" and 0 otherwise, likewise, $x_{it}^- = x_{it}$ if "down" and 0 otherwise with w_i^+ and w_i^- the associated weights. Anderson (1952) assumes that, up to a mean zero disturbance ξ_{it}, $x_{it}^+ = \alpha$ and $x_{it}^- = -\beta$, $\alpha, \beta > 0$, giving

$$\begin{aligned} x_t &= \alpha\sum\nolimits_{i=1}^{N_t} w_i^+ - \beta\sum\nolimits_{i=1}^{N_t} w_i^- + \xi_t && \text{(C.6)} \\ &= \alpha U_t - \beta D_t + \xi_t, && \text{(C.7)} \end{aligned}$$

where $\xi_t = \sum_{i=1}^{N_t} w_i \xi_{it}$ and U_t and D_t now denote the respective (weighted) proportions of firms reporting an output rise and fall. 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 (C.7) 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 (1984, 1987) extends this basic model to allow for an asymmetric relationship between x_t and x_{it}.

C.2.1 Relating the Regression Approach to the Probability Approach

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

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

with corresponding cumulative distribution function

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

From (C.1),

$$U_t = \frac{q + \hat{x}_t - \lambda}{2q}, D_t = \frac{q - \hat{x}_t - \lambda}{2q}, \tag{C.8}$$

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

$$\hat{x}_t = q(U_t - D_t), \tag{C.9}$$

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

A generalisation of (C.9) is obtained by relaxing the assumption that the "no change" interval is symmetric; that is, replace $(-\lambda, \lambda)$ by $(-a, b)$. Hence, (C.8) becomes

$$U_t = \frac{q + \hat{x}_t - b}{2q}, D_t = \frac{q - \hat{x}_t - a}{2q}.$$

with the estimator for x_t as

$$\hat{x}_t = \alpha U_t - \beta D_t,$$

which is equivalent to the estimator for x_t in (C.7) based on U_t and D_t for the single time period t, where the two scaling parameters are defined as

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

C.3 The Reverse-Regression Approach

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

Let the unobserved firm-specific output growth rate y_{it} be related to x_t through the linear representation

$$y_{it} = x_t + \eta_{it} + \varepsilon_{it}. \tag{C.10}$$

which may be expressed in terms of (2.2) by defining $\eta_{it} = \alpha_i + (\beta_i - 1)x_t, (i = 1, ..., N_t, t = 1, ..., T)$. In (C.10), η_{it} is the difference between y_{it} and x_t anticipated by firm i while ε_{it} is an unanticipated component, i.e., $E[y_{it}|i] = x_{it} = x_t + \eta_{it}$.

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

$$y_{it}^j = 1 \text{ if } c_{j-1} < y_{it} \leq c_j \text{ and 0 otherwise, } j = 1, 2, 3, \tag{C.11}$$

where $c_0 = -\infty$ and $c_3 = \infty$ 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. From (C.10), the observation rule (C.11) becomes

$$y_{it}^j = 1 \text{ if } c_{j-1} - x_t < \eta_{it} + \epsilon_{it} \leq c_j - x_t \text{ and 0 otherwise.} \tag{C.12}$$

A probabilistic foundation may be given to (C.12) by letting the scaled error terms $\{\sigma(\eta_{it} + \epsilon_{it})\}$, $\sigma > 0$, $i = 1, ..., N_t$, possess a common and known cumulative distribution function $F(\cdot)$ which is parameter free and assumed time-invariant. Then,

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

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

C.3.1 Motivating the Regression Formulation

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

$$P_t^1 = D_t = F(\mu_1 - \sigma x_t) + \xi_t^1, P_t^3 = U_t = F(-(\mu_2 - \sigma x_t)) + \xi_t^3. \tag{C.13}$$

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

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

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 $F(\cdot)$ is strictly monotonic, the non-linear regressions (C.13) 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

$$F^{-1}(D_t) = \mu_1 - \sigma x_t + u_t^1, F^{-1}(U_t) = -\mu_2 + \sigma x_t + u_t^3, \tag{C.14}$$

where $u_t^1 = (f_t^1)^{-1}\xi_t^1 + o_p(N_t{}^{-1})$, $u_t^3 = (f_t^3)^{-1}\xi_{t,3} + o_p(N_t{}^{-1})$ with $f_t^1 = f(\mu_1 - \sigma x_t)$, $f_t^3 = f(-(\mu_2 - \sigma x_t))$ and the density function $f(z) = dF(z)/dz$.

Since x_t is observed, feasible and asymptotically efficient estimation of (C.14) is achieved by generalised least squares (or minimum chi-squared) estimation given the structure of the variance matrix of u_t^1 and u_t^3.

C.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 (C.14) and let

$$\hat{x}_t^1 = \frac{\hat{\mu}_1 - F^{-1}(D_t)}{\hat{\sigma}}, \hat{x}_t^3 = \frac{\hat{\mu}_2 + F^{-1}(U_t)}{\hat{\sigma}}. \tag{C.15}$$

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 (C.15).

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

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

$$1 - U_t = F(\frac{b - \hat{x}_t}{\hat{\sigma}_t}), D_t = F(\frac{-a - \hat{x}_t}{\hat{\sigma}_t}).$$

From the symmetry of $F(\cdot)$,

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

Hence,

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

Therefore, in comparison with (C.14), $\mu_1 = -a/\sigma_t$, $\mu_2 = b/\sigma_t$ and $\sigma = 1/\sigma_t$.

References

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

Ashley, J., Driver, R., Hayes, S. & Jeffery, C. (2005), 'Dealing with data uncertainty', *Bank of England Quarterly Bulletin* **Spring**, 23–29.

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

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

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

Bontemps, C. & Meddahi, N. (2005), 'Testing normality: a gmm approach', *Journal of Econometrics* **124**(1), 149–186.

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

Butler, J. & Moffitt, R. (1982), 'A computationally efficient quadrature procedure for the one-factor multinomial probit model', *Econometrica* **50**, 761–764.

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

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

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

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

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

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

Gourieroux, C., Monfort, A., Renault, E. & Trognon, A. (1987), 'Generalised residuals', *Journal of Econometrics* **34**(1-2), 5–32.

Gourieroux, C., Monfort, A. & Trognon, A. (1985), 'A general approach to serial correlation', *Econometric Theory* **1**(03), 315–340.

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

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

Hsiao, C., Pesaran, M. H. & Pick, A. (2011), 'Diagnostic tests of cross section independence for nonlinear panel data models', *Oxford Bulletin of Economics and Statistics* . Forthcoming.

Lee, K. (1994), 'Formation of Price and Cost Inflation Expectations in British Manufacturing: a Multisectoral Anlaysis', *Economic Journal* **104**, 372–386.

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

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

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

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

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

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

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

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

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

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

Pesaran, M. H. (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica* **74**(4), 967–1012.

Pesaran, M. H. & Timmermann, A. (2009), 'Testing dependence among serially correlated multicategory variables', *Journal of the American Statistical Association* **104**(485), 325–337.

Pesaran, M. H. & Weale, M. R. (2006), Survey Expectations, *in* G. Elliott, C. W. J. Granger & A. Timmermann, eds, 'Handbook of Economic Forecasting Volume 1', North-Holland, pp. 715–776.

Robinson, P. M. (1982), 'On the asymptotic properties of estimators of models containing limited dependent variables', *Econometrica* **50**(1), 27–41.

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

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

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

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

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

Timmermann, A. (2006), Forecast combinations, *in* G. Elliott, C. W. J. Granger & A. Timmermann, eds, 'Handbook of Economic Forecasting Volume 1', North-Holland, pp. 135–196.