Collective Sentiment in Qualitative Business Surveys*

Silvia Lui, James Mitchell and Martin Weale
National Institute of Economic and Social Research

March 27, 2009

Abstract

Aggregated qualitative survey data provide timely, but often imperfect, macroeconomic indicators. Exploiting a unique panel dataset for the UK, which contains matched firm-level responses from both qualitative and quantitative surveys, we find that firms' responses are influenced not only by their own current and past output and lagged qualitative responses but also by an indicator of aggregate response. We interpret the latter as showing that firms' answers reflect collective sentiment as well as individual experience. Collective sentiment is shown to have a substantial impact on the summary statistics conventionally used to represent the results of the survey.

JEL Codes C01, C23, C81

Keywords: Qualitative Business Survey, Quantitative Firm Data, Collective Sentiment, Common Factor Test

---

*Address for correspondence: s.lui@niesr.ac.uk; j.mitchell@niesr.ac.uk and m.weale@niesr.ac.uk. We gratefully acknowledge financial support from the ESRC (Award Reference: RES-062-23-0239) and the Bank of England. We thank the CBI and the ONS for their help in facilitating this project, with particular thanks to Rhys Davies (ONS), Robert Gilhooly (ONS), Felix Ritchie (ONS), Richard Welpton (ONS) and Jonathan Wood (CBI).

1 Introduction

Qualitative business survey data are consulted widely as barometers of economic activity. Interest focuses on them because they are more timely than official data, which are typically published at a lag, although the cost of that timeliness is their qualitative nature. Firms are asked a range of questions to which they provide categorical instead of quantitative answers; for example, they are asked whether output has fallen, stayed the same or risen but not by how much it has changed. Traditionally the findings of these surveys are then converted into early quantitative estimates of movements in economic activity by taking the proportion of firms reporting that output has risen, stayed the same or fallen, and relating them to official output data. Most commonly, and simply, this is achieved by relating the proportion of optimists less pessimists from the qualitative survey, the "balance statistic", to the official data; see Pesaran & Weale (2006) for a survey. Such data have been widely used in the construction of macro-economic indicators (e.g., see Hansson et al. (2005), Banerjee et al. (2005), Giannone et al. (2008) and Altissimo et al. (2009)).

It is clear that the balance statistic alone does not always track movements in the official data very well, which in large part explains why such statistics tend to be seen as components of aggregate indicators along with many other variables. However there is no clear or simple answer why that should be the case. One possibility is, of course, that there are aggregation biases. It is by no means obvious that the balance statistic is the most appropriate way of summarising the survey data except in the very special circumstances identified by Anderson (1952).

In this paper we adopt an alternative route by exploring what lies behind the responses of the individual manufacturing firms to a qualitative survey. We match firms' responses in our qualitative survey with the answers given by the same firms to the official quantitative survey used to provide aggregate official statistics on industrial output. This means that we can investigate the influences on qualitative responses of both individual and aggregate variables in order to understand the behavioural factors explaining firms' qualitative responses.

We find that the answers firms provide to the qualitative business survey reflect not only what is happening to themselves but also what other firms believe is happening. What we call firms' collective "sentiment", in other words, is a significant (common) factor explaining movements in qualitative surveys; it captures the tendency for firms, when replying to qualitative surveys, to follow, what some might call, the herd.

Finally, we comment on the implications of the existence of collective sentiment for the manner in which qualitative survey data, typically available only at the aggregated level, are quantified (see Pesaran & Weale (2006) for a review of alternative quantification methods). The precise form these quantification exercises take depends, traditionally, on what is assumed (rather than inferred from data) about how firms' reply to qualitative surveys. Given our access to matched qualitative-quantitative data rather than make an assumption, and then seeing how this works at the macroeconomic level, we suggest that the quantification approach adopted might reflect the way firms reply to qualitative surveys at an individual level.

The plan of the remainder of this paper is as follows. Section 2 sets out our model of firm-level behaviour and describes our characterisation of collective sentiment. Section 3 then describes the matched data-set and reports the estimation results. Section 4 concludes.

2 Firm-level sentiment

Qualitative business surveys, such as the Confederation of British Industry's (CBI) Industrial Trends Survey (ITS) in the UK and related surveys in other countries, ask firms many questions, only some of which have a natural counterpart in official surveys - namely surveys conducted by the national statistical office.[1] We focus on the retrospective output question from the ITS which currently takes the form, "Excluding seasonal variations, what has been the trend over the past three months with regard to volume of output?".[2] Firms reply "up", "same" or "down" (and "not applicable"). The ITS is published at the end of the month concerned, and is therefore published ahead of official data on (manufacturing) output explaining its potential value as a timely source of information about the current state of the economy. The ITS can be related to the Office for National Statistics' (ONS') Monthly Production Inquiry (MPI), which asks firms each month for quantitative information on their turnover values. The MPI is a sample-based survey covering all of the UK and is the basis for the monthly Index of Production (IoP).

The ITS asks each firm, i $(i = 1, \cdots, N_t)$ at time t $(t = 1, \cdots, T_i)$, to give qualitative answers to the question about its trend of output (excluding seasonal variations) over

[1] The CBI survey is the basis for the UK data maintained in the European Commission's Euro-wide database of business and consumer surveys.

[2] Until June 2003, firms were asked for their trend over the last four months. The probit models considered below let the data decide the horizon over which the qualitative data relate to the quantitative data.

the past three months. As discussed, the firm can respond either 'up', 'same' or 'down', denoted as j (where $j = 2, 1, 0$, respectively).[3] Assume that there is a continuous latent variable y_{it}^* that triggers firm i's categorical response at time t via the following observation rule:

$$y_{it} = j \text{ if } \mu_j < y_{it}^* \leq \mu_{j+1},\ j = 0, 1, 2, \tag{1}$$

where μ_j's are the unknown thresholds: $\mu_0 = -\infty$, $\mu_j \leq \mu_{j+1}$ and $\mu_3 = \infty$.

2.1 "Collective sentiment" as a common factor

We assume that each firm's sentiment y_{it}^*, as recorded in the qualitative survey, depends on a weighted average of that firm's actual output growth Δx_{it} as measured by the official (quantitative) survey, plus lags of y_{it}^* and a firm-level error ϵ_{it}.The error ϵ_{it} is endowed with a factor structure, where the common factors represent collective sentiment. Collective sentiment may be made up of a number of components ($k = 1, ..., K$). We denote each individual component of collective sentiment in period t as CS_{kt}. The firm-level model is consequently written as:

$$\begin{aligned} y_{it}^* &= \alpha_i(L)\Delta x_{it} + \beta_i(L)y_{it-1}^* + \epsilon_{it} && (2) \\ \epsilon_{it} &= \sum_{k=1}^{K} \lambda_{ik} CS_{kt} + v_{it} && (3) \end{aligned}$$

where Δx_{it} denotes the 1-month logarithmic difference of the volume of turnover for firm i for time (month) t as measured by the MPI, λ_{ik} captures the loading of collective sentiment of type k on firm i, and v_{it}, an idiosyncratic shock, is white noise and uncorrelated across firms. Temporal dependence in y_{it}^* is captured via lagged dummy variables y_{it} (one for $j = 0$ and one for $j = 2$). When a firm does, as instructed by the CBI, base its qualitative reply on its growth over the last three (previously four) months, we should expect lags in (2), beyond the horizon of interest (namely three/four months), to be statistically insignificant.

Therefore, (2) characterises the components of collective sentiment, if present, as a linear combination of common factors which influence all firms' responses to the qualitative survey.

We first test the null hypothesis that collective sentiment plays no role, estimating

[3] Firms can, and a very small number do, also respond 'not applicable'. We ignore these firms below.

equation (2) on the assumption that the null hypothesis of cross-sectional independence is true and then testing whether ϵ_{it} admits a factor structure. We test for cross-sectional dependence (CD) in ϵ_{it} using the CD test developed by Hsiao et al. (2007) for use with nonlinear panel data models. The CD test statistic is based on the average (across firms $i = 1, ..., N$) pair-wise residual correlation coefficient, with the 'residual' measured as the generalised residual from ordered probit estimation of (2). Under the null of cross-sectional independence, the CD statistic tends to a standard normal variate.

If the CD test does suggest cross-sectional dependence, which we identify with collective sentiment, we then seek to identify what influences collective sentiment by exploring whether economy-wide variables serve to remove the dependence. The most obvious candidates to add to (2) as additional explanatory variables are current and lagged values of the balance statistic, Bal_t, defined as the proportion of firms reporting a rise in output less the proportion reporting a fall, and Δx_t, the one-month change in the logarithm of the official manufacturing output index..

There is a question about whether Bal_t and Δx_t should be measured as cross-sectional averages from the matched panel data-set, discussed further in Section 3, or as observed (aggregate) statistics as published by the CBI and ONS. Identification of Bal_t with the former, which amounts to the common correlated effects estimator, is shown by Pesaran (2006) to provide consistent estimators of the factors driving ϵ_{it}, when ϵ_{it} admits a single common factor structure where the factor is unobserved rather than observed. Obviously if there is more than one significant factor the null hypothesis of no cross-sectional dependence will be rejected even when this measure of Bal_t is included in the regression. Empirically we examine both possibilities.

These specifications allow us to explore whether collective sentiment is influenced by one or both of other firms' feelings about their own experience and of actual current and past movements in aggregate output. It should be noted that the CBI is asking firms to report only on their own circumstances and if they are indeed doing as requested neither effect should be present.[4] Indeed there should be no cross-sectional dependence and no role for collective sentiment.

[4]While this is true, the CBI's own evidence obtained by periodic answering practice surveys shows that generally, despite clear instructions via the question wordings themselves, there are a spectrum of considerations and characteristics that seem to influence the participants when answering questions.

3 Matched qualitative and quantitative survey data

We relate the individual responses provided to the CBI with those collected by the ONS on a firm-by-firm basis. To match firms' responses across the two surveys we arranged for the CBI to provide their data set to the ONS in a manner which preserved obligations of confidentiality for both bodies. The ONS then matched the CBI data set to its MPI for the five years 2000-2004, inclusive. This allowed us to match up the response provided by each firm in answering the CBI survey with the response provided by the same firm in providing quantitative information to the ONS in its MPI. The firm-level MPI turnover data are deflated, using the IoP deflator at the 4 digit level, to produce figures in volume terms.

On average (across the five years), there are about 170 firms present in the matched dataset each month, with, in total, across the five years, 807 firms providing at least one contemporaneous response to both the ITS and MPI. This represents an average (in a year) match rate against the ITS (which interviews about 800 firms each month) of about 21% of firms. Overall, since 2584 firms were sampled at least once by the CBI the match rate is 31%.

3.1 Empirical results

Table 1 summarises results from estimation of pooled ($\alpha_i = \alpha$ and $\beta_i = \beta$) dynamic ordered probit models based on (2). The bottom two rows of the table provide results from the CD test.

Starting with a maximum of six lags of each of the explanatory variables, Wald tests are first used to select more parsimonious but good-fitting specifications.[5] These tests indicate that in columns F and G, the most general specifications reported in Table 1, additional lags of the explanatory variables are statistically insignificant.[6] Firms also appear to follow the CBI's instructions and base their qualitative responses on their output movements over the previous three months. We see that the quantitative data are jointly statistically significant. This indicates that there is plainly a relationship between the qualitative and quantitative data and the qualitative data are not simply noise. But the

[5] Seasonal, time, sectoral and size dummies were included in the probit models but found to be statistically insignificant. We also experimented with random effects variants, but the random effect was not significant statistically.

[6] One cannot reject the null hypothesis that excluded variables have a coefficient of zero with p-values of 0.03 and 0.05 in columns F and G. The Bayesian Information Criterion also selects columns F and G over more general alternatives.

quantitative data are insignificant at time t, with lagged values of Δx_{it} explaining firms' qualitative responses. In other words, which remains perfectly consistent with the CBI's question, the firm-level ITS data at time t do not contain a statistically significant signal about growth in the last month, but about growth one month ago relative to two, and two relative to three. There is also a role for what we might call "individual" sentiment, since lagged qualitative data are also playing an important role, with considerable inertia in firms' qualitative responses.

Now turn directly to column A, the most restricted specification reported, which presents results from the benchmark model, which sets $\lambda_{ik} = 0$. The CD test rejects the null hypothesis of cross-sectional independence at a 95% significance level: "sentiment", of the sort seen at the macroeconomic level in Figure 1, arises because of collective, as well as individual, sentiment.

The other columns in Table 1 can then be interpreted as different attempts to identify collective sentiment and thereby capture the cross-sectional dependence found in column A. They involve adding Bal_t and/or Δx_t, both as published and as cross-sectional averages of those firms present in column A (denoted *match*), to the benchmark model.

Columns B and C indicate that the balance statistic alone, irrespective of the sample from which it is computed, does appear to be the single common factor which characterises collective sentiment (i.e., $K = 1$); in columns B and C the p-value of the CD test rises well above 0.05 and one can no longer reject the null hypothesis of cross-sectional independence at 95%. In turn, columns D to G show that including Δx_t, irrespective of how it is measured, does not alone capture the cross-sectional dependence. Moreover, Δx_t is not statistically significant in the probit models.

3.1.1 Implications of the existence of collective sentiment

Traditional aggregate quantification approaches, such as the probability approach of Carlson & Parkin (1975) and the related regression approach of Pesaran (1987), in the absence of firm-level categorical information, typically focus on the sample (size N_t) proportion of firms that reported an output fall and rise, respectively. In terms of the firm-level model (2), when $\lambda_{ik} = 0$, these proportions are given as:

$$D_t \xrightarrow{p} P(y_{it}^* \leq \mu_1) \quad (4)$$

$$U_t \xrightarrow{p} P(y_{it}^* \geq \mu_2) \quad (5)$$

The presence of collective sentiment can be interpreted as meaning that these thresholds are, in fact, time-varying, such that:

$$D_t \xrightarrow{p} P(y_{it}^* \leq \mu_1 - \lambda_1 Bal_t) \quad (6)$$
$$U_t \xrightarrow{p} P(y_{it}^* \geq \mu_2 - \lambda_1 Bal_t) \quad (7)$$

Similarly, aggregate quantification methods have been developed that allow for a time-varying threshold; e.g., see Smith & McAleer (1995) for a review and extensions. Most commonly, the thresholds are either assumed to follow a random walk or assumed to depend on the macroeconomic growth rate Δx_t and possibly its lagged values. Since there is a correlation of 0.6 between Δx_t and Bal_t this finding could be interpreted as a proxy for a situation where the thresholds in fact depend on the aggregate response. However, it is only possible to provide micro-economic foundations, in terms of our firm-level model (2)-(3), for the use of aggregate quantification methods with time-varying thresholds under specific conditions. Traditionally, these aggregate quantification approaches assume (i) $\Delta x_{it} = \Delta x_t + u_{it}$, where u_{it} is an i.i.d. mean zero random variable; (ii) $y_{it}^* = \Delta x_{it} + u_{it}$ and (iii) the thresholds μ_j are fixed and common across firms and time. Smith & McAleer (1995) review and propose extensions to this approach, which let the μ_j's vary over time. If $y_{it}^* = \Delta x_{it} + \lambda_1 Bal_t + u_{it}$ then it is as though the thresholds μ_j become $(\mu_j - \lambda_1 Bal_t)$, thereby generating a particular form of time-variation in the thresholds. Moreover, if y_{it}^* also depends on lagged values of Δx_{it} then the thresholds are also affected by lags of Δx_{it}. But our model, as estimated in Table 1, shows that y_{it}^* is affected by lags of y_{it}^*. It is therefore not possible to establish a formal analytical link between these quantification methods and our own micro-based findings. However, heuristically our findings might be taken to offer some support for use of aggregate quantification approaches with time-variation in the thresholds.

To gain a quantitative impression of how collective sentiment affects our panel of firms' qualitative responses, via the observation rule, (1), we compute the proportions $\frac{1}{N_t}\sum_{i=1}^{N_t} \Pr(y_{it} = j)$ for $j = 0$ (the pessimists) and $j = 2$ (the optimists) from the estimated probit model seen in Table 1. The balance statistic $(U_t - D_t)$ is then plotted both using the probit estimates seen in column B of Table 1, which accommodates collective sentiment, and assuming collective sentiment plays no role.

Figure 1, which compares these proportions, reveals that collective sentiment clearly affects firms' responses to the qualitative survey leading to a more depressed impression of the state of the economy in the recession of the early 2000s and a more optimistic view in the subsequent period of sustained growth. Thus the effect of collective sentiment is, as might be expected, largely to amplify movements in the balance statistic. Comparison of the two balance statistics with the bottom panel of Figure 1, which plots the 12 month growth rate of aggregate IoP, reveals there to be less informational content in the qualitative survey when we ignore the effects of collective sentiment - the correlation coefficient against the aggregate growth rate drops from 0.60 to 0.47.

4 Conclusions

Exploiting a unique panel dataset we seek to understand what lies behind firms' responses to qualitative business surveys. We find that firms, when replying to qualitative surveys, react not just to the hard (quantitative) facts but to the "herd". Sentiment, in other words, has a common collective component, as well as individual components. This feedback does explain the tendency for the balance statistic to under/over react to macroeconomic events. It also implies that the thresholds which trigger firms' categorical responses, and which are central to aggregate quantification methods like the probability approach, are time-varying and a function of the balance statistic.

Table 1: Panel ordered probit results explaining firms' qualitative responses, with robust t-values in parentheses

y_{it}	A	B	C	D	E	F	G
Explan variab							
Δx_{it}	0.013	0.008	0.013	0.010	0.009	0.008	0.005
	(0.59)	(0.39)	(0.62)	(0.42)	(0.39)	(0.32)	(0.24)
Δx_{it-1}	0.091	0.091	0.088	0.091	0.091	0.091	0.091
	(3.88)	(3.9)	(3.77)	(3.88)	(3.88)	(3.9)	(3.9)
Δx_{it-2}	0.124	0.122	0.123	0.125	0.125	0.122	0.122
	(5.62)	(5.5)	(5.55)	(5.63)	(5.64)	(5.5)	(5.52)
y^u_{it-1}	0.878	0.873	0.877	0.879	0.880	0.873	0.874
	(16.42)	(16.24)	(16.33)	(16.43)	(16.43)	(16.25)	(16.25)
y^d_{it-1}	-0.833	-0.822	-0.821	-0.833	-0.834	-0.822	-0.823
	(-15.73)	(-15.56)	(-15.5)	(-15.73)	(-15.76)	(-15.56)	(-15.58)
y^u_{it-2}	0.493	0.497	0.493	0.492	0.492	0.497	0.496
	(9.39)	(9.45)	(9.38)	(9.37)	(9.36)	(9.44)	(9.43)
y^d_{it-2}	-0.393	-0.383	-0.384	-0.393	-0.392	-0.383	-0.382
	(-7.67)	(-7.48)	(-7.48)	(-7.67)	(-7.65)	(-7.48)	(-7.46)
Bal^{match}_t	-	0.010	-	-	-	0.010	0.010
		(6.23)				(6.22)	(6.18)
Bal_t	-	-	0.010	-	-	-	-
			(5.43)				
Δx^{match}_t	-	-	-	0.0138	-	0.004	-
				(0.29)		(0.08)	
Δx_t	-	-	-	-	2.171	-	1.719
					(1.31)		(1.03)
μ_1	-0.750	-0.779	-0.765	-0.750	-0.751	-0.779	-0.780
	(-23.72)	(-24.23)	(-23.99)	(-23.72)	(-23.72)	(-24.24)	(-24.24)
μ_2	0.840	0.821	0.832	0.839	0.839	0.821	0.821
	(26.80)	(25.88)	(26.38)	(26.76)	(26.75)	(-25.86)	(25.85)
$\sum_{i=1}^N T_i$	4079	4079	4079	4079	4079	4079	4079
N	381	381	381	381	381	381	381
LL	-3430.5	-3411.6	-3416.3	-3430.4	-3429.7	-3411.6	-3411.1
BIC	6935	6906	6915	6943	6942	6914	6913
CD test stat	1.6651	0.2124	0.8722	1.6564	1.7273	0.2094	0.1435
$p-value$ of CD test	0.0479	0.4159	0.1915	0.0488	0.0421	0.4171	0.4429

Notes: Δx_{it} is the rolling 1-month growth rate of the quantitative data; y^u_{it-1} and y^d_{it-1} are dummy variables indicating a firm's previous qualitative response; Bal_t is the balance statistic computed from the qualitative survey and Δx_t is the 1-month growth rate of the macroeconomic data. CD is the test statistic of Hsiao et al. (2007).

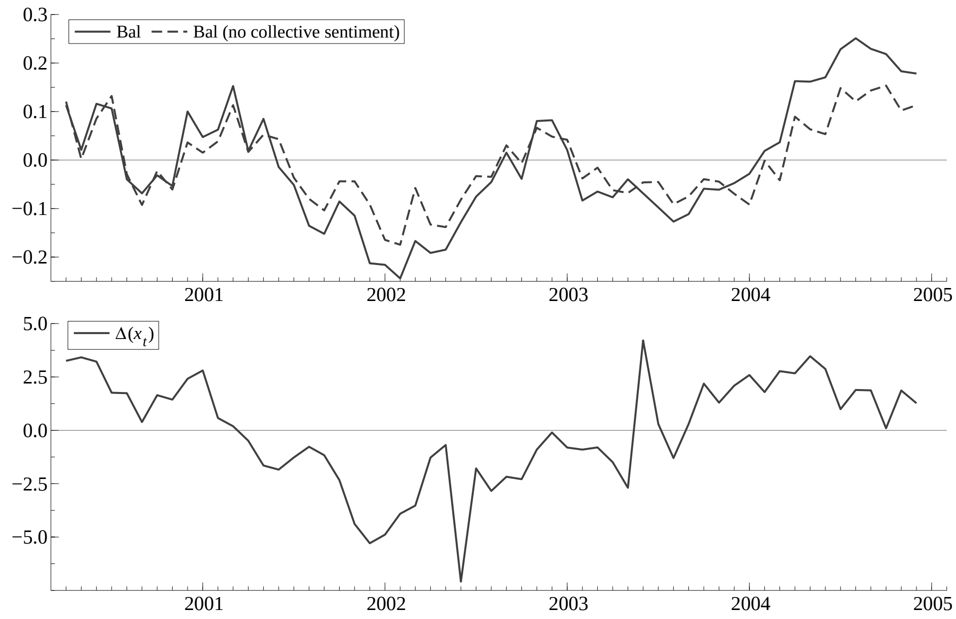

Figure 1: The role of collective sentiment

References

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M. & Veronese, G. (2009), 'New eurocoin: Tracking economic growth in real time', *Review of Economics and Statistics* . Forthcoming. Available as CEPR Discussion Paper 5633.

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.

Banerjee, A., Marcellino, M. & Masten, I. (2005), 'Leading indicators for euro-area inflation and gdp growth', *Oxford Bulletin of Economics and Statistics* **67**(s1), 785–813.

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

Giannone, D., Reichlin, L. & Small, D. (2008), 'Nowcasting: The real-time informational content of macroeconomic data', *Journal of Monetary Economics* **55**(4), 665–676.

Hansson, J., Jansson, P. & Lof, M. (2005), 'Business survey data: Do they help in forecasting gdp growth?', *International Journal of Forecasting* **21**(2), 377–389.

Hsiao, C., Pesaran, M. H. & Pick, A. (2007), Diagnostic tests of cross section independence for nonlinear panel data models. Cambridge University Working Paper.

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. & 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.

Smith, J. & McAleer, M. (1995), 'Alternative procedures for converting qualitative response data to quantitative expectations: an application to Australian manufacturing', *Journal of Applied Econometrics* **10**, 165–185.