Quantifying survey data

By Alastair Cunningham of the Bank's Conjunctural Assessment and Projections Division.

In this article[1] *Alastair Cunningham explains how data from economic surveys can be used to complement official statistics. He sets out a simple framework to analyse how firms respond to surveys and outlines the most widely used technique for converting qualitative responses into a quantitative measure. He shows that the results of this technique are often biased, and describes a more rigorous approach. Possible explanations are put forward for why survey data tend to be less volatile than official data. Finally, the use of forward-looking survey data is discussed.*

Introduction

In addition to official data, mainly from the Office for National Statistics (ONS), the Bank receives around 30 regular economic surveys from the private sector. These include the CBI's quarterly *Industrial Trends Survey* and the *Quarterly Economic Survey* produced by the British Chambers of Commerce (BCC). Many private sector surveys offer a qualitative indication of economic conditions whereas the official data (though also usually based on surveys) are quantitative estimates.

The first section of this article explains how surveys can usefully complement official estimates—the context in which we analyse survey data. Before the implications of the survey data can be assessed, we need to convert any qualitative survey responses into quantitative estimates. To do this accurately, we need to understand how the information in the survey is collected and presented. This is explored in the second section of this article. The third section discusses a widely used technique for converting qualitative survey data into quantitative estimates; next, we outline a more rigorous approach before reviewing the issues raised when the official data are more volatile than the private sector survey data. Finally, we discuss the additional factors that we need to consider when interpreting forward-looking surveys of expectations.

How are survey data useful?

Understanding how data from various sources relate to one another is central to economic analysis. We want to know how different variables are related: for example, when analysing the housing market, how do housing completions (data source: Department of the Environment) relate to data on housing sales (such as supplied by the Royal Institute of Chartered Surveyors survey)? And we want to reconcile estimates of a single variable, such as manufacturing output, from different sources (the ONS and private sector surveys).

Research into quantifying survey data is part of a general effort—both in the Bank and by external economists—towards integrating the diverse data available into a systematic analysis of the economy. For example, the National Institute of Economic and Social Research (NIESR) has recently started publishing a monthly indicator of GDP, relating a range of monthly data to total output.

Survey data can help economists to analyse the economy in a number of ways:

(i) Giving early information on the current state of the economy

Official data provide the foundations for economic assessment. UK data are high quality by international standards—of 13 national statistical offices covered in a 1993 survey (published in *The Economist*), the ONS was ranked joint second for the timeliness of publication and the small size of revisions. But there is a lag between the publication of the data and the period to which they refer. And perhaps more importantly, the data are often revised after publication as more information becomes available. Because of the delay before official estimates are finalised, economists may use surveys and other indicators to improve their analysis of the recent past.

(ii) Covering sectors for which official data are less frequent

Not all sectors of the economy are covered equally well by regular and timely official data. For example, the ONS currently produces monthly estimates of manufacturing production, but only a quarterly estimate of output in the (much larger) service sector. Where official data are scarce, other sources of information such as the CBI's *Distributive Trades Survey* become more valuable.

(iii) As an indicator of expectations

Many surveys ask respondents about their expectations, as well as about recent experiences. For most variables,

(1) This article draws heavily on earlier research undertaken in conjunction with Martin Weale (NIESR) and Richard Smith (Bristol University).

surveys are the only source of information on expectations. Information about expectations is useful because economic agents' views of future prospects can affect their current behaviour. For example, if consumers come to expect faster income growth in the future, they may raise expenditure on goods and services today.

Quantifying surveys of expectations is more complex than quantifying backward-looking surveys, because there are rarely any official statistics on expectations to compare the survey data with. Although the bulk of this article focuses on backward-looking surveys, the points also apply to surveys of expectations, and the lack of official data on expectations is discussed in the final section.

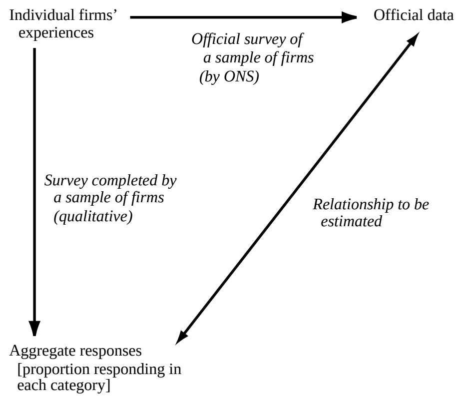
Economists may also be interested in how good an indication of the state of the economy the various surveys give, and indeed whether the official estimates are the best indicator. That may be the focus of future Bank research, but is not discussed in this article. Here we are concerned with how best to match survey and official data, regardless of their relative performance as indicators. We do not aim to model the official data in any behavioural sense. Instead, we wish to transform the qualitative survey data into a quantitative estimate that is (on average) consistent with the official estimates, once they have been finalised.

To make the fullest use of survey data as a complement to official data, we need to turn qualitative survey responses into quantitative estimates. There is a wide range of possible techniques that economists can use to relate survey data to official estimates. To choose between them, we need to understand how the information in surveys is collected and presented.

A simple framework for analysing surveys

Chart 1 represents the relationships between the official and survey-based estimates of an economic variable. The official data are the (weighted) average experiences of a sample of firms. Surveys are also based on questions about the experiences of a sample of firms, but these questions are usually qualitative and published in aggregated form as the proportion of firms answering in each of a series of categories—typically 'rise', 'fall', or 'the same'.

Chart 1
Relationships in surveys



We can use a simple framework to reveal any implicit assumptions that we make when analysing survey data. This should help us to choose between the various possible techniques we can use to interpret the survey. The framework has two parts, corresponding broadly to the first links in the triangle in Chart 1:

- Sampling: how do the experiences of individual firms covered by the official (ONS) sample differ from those of the firms completing the private sector survey?

- An 'observation rule': how do the responses given by the individual firms completing the survey relate to their experiences?

The answers to these two questions can inform our approach to the final leg of the triangle: relating the aggregate survey responses to official estimates of the variable.

All the numerous techniques that economists use to relate survey data to official estimates rest on assumptions about the relationships embodied in the survey, in particular the nature of the observation rule. Our simple framework can be used to judge between these different approaches.

Sampling errors

It is easiest to describe the framework in terms of a specific variable, such as output. Each firm's output growth can be divided into two parts: the economy-wide average plus some firm (or industry)-specific influence. These specific influences must average zero across all firms and industries, so the economy-wide average will approximately equal the average growth rate experienced by individual firms.(1)

But the official and survey data are both usually based on samples of the firms in the economy. So the average experience recorded will equal the economy-wide average *plus* a random sampling error. Because the official data and the survey data are usually based on different samples, they are subject to different sampling errors.

Attempts to match the data will be impeded if there is a pattern—a systematic variation—in the differences in sampling error. This will only occur if firms sampled in the official data experience consistently different conditions from those in the survey data. So for example, some commentators have been concerned that the CBI's *Industrial Trends Survey* may be biased towards exporters. If this is so, and exporters' experiences are not thought to be representative of the economy as a whole, then the user of the survey must make allowances for this.

(1) For the purposes of this paper we ignore any differences due to 'aggregation bias'.

An observation rule for individual firms

The next (and more complex) step is to understand how firms report their individual experiences. In most surveys, firms are not asked to report their output directly, but to state, for example, whether output has 'risen', 'fallen' or 'stayed the same'. To understand how firms' experiences relate to their responses, we need to define the range of output growth that firms regard as falling into each category. In the usual case, where there are only three categories, this merely involves defining the range of outcomes that firms regard as 'the same'.

At first glance, the answer seems obvious—firms should only report 'the same' if output growth is exactly zero. But the probability of output growth being *exactly* zero is very small under most circumstances, so firms should rarely give this response. But we observe, for example, that since 1972, an average of 48% of respondents to the CBI's *Industrial Trends* survey have reported unchanged output volumes in any one quarter.

If firms are reporting 'the same' when output has changed, there must be some range of output variation that they regard as essentially unchanged. This range, which is termed the 'indifference band', underlies the information offered by the survey.

The CBI periodically investigates the answering practices of its respondents. The results of the most recent enquiry were published in 1990 (Junankar 1990) and suggested that the indifference band could be significant. When asked 'what range of movement would you regard as falling within the reply 'the same'?', only 11% responded 'up to 1%' and over a quarter responded 'up to 4%–8%'.

One reason for the existence of an indifference band may be that firms are uncertain about what has happened and how to report it. The CBI's investigation of answering practices suggests one potential source of uncertainty: the timing of the period used to assess changes in output. Respondents to the quarterly survey are asked about the trend 'over the past four months'. The CBI's investigation found that this was interpreted differently by different respondents. Around one half of the respondents compared the latest four-month period with the previous four months. But significant minorities compared the start with the end of the four-month period (21%); or compared with experiences a year earlier (9%); or even used a combination of the three (16%).

If a firm is uncertain about its experiences, then it is likely to regard a small change as essentially the same. The firm will only record a rise if it is sufficiently certain that the change is significant. If this reasoning is correct, then we might expect the indifference bands to be widest for questions about experiences of which the firm is most uncertain. In that case, indifference bands should be wider for surveys of expectations than for surveys of experiences, because of the additional uncertainty faced when taking a view about the future.

If a firm becomes less uncertain about its experiences then we might expect the indifference band to become narrower over time. So for example, uncertainty may have risen as markets have become increasingly global. Alternatively, the introduction of computerised stock control may have reduced uncertainty. Because we have no prior view about how (and if) uncertainty has shifted over time, the approaches discussed in this article all make the simplifying assumption that the bands are constant through time.

In the remainder of the article, we discuss techniques that may be used to quantify survey data in the light of our discussion. The methods set out are illustrated by an application to the CBI's *Industrial Trends Survey*. But the purpose of this article is to set out the techniques rather than the results of these specific regressions. The points made should apply to any qualitative surveys. Details of the regression results underpinning the charts in this article are set out as an appendix.

A common approach: the 'balance' statistic

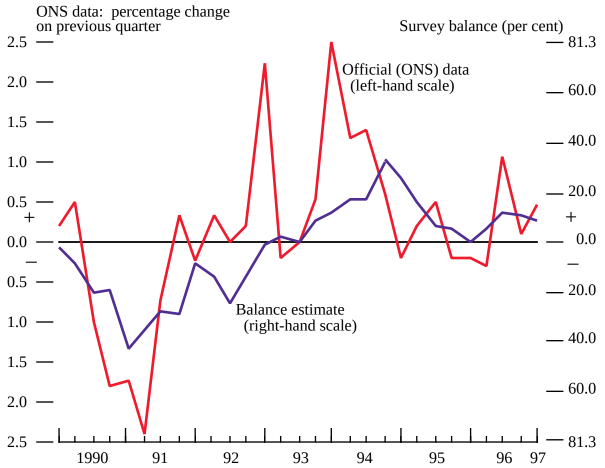
One of the most commonly used representations of survey data is the 'balance' statistic: the difference between the proportion of firms reporting a rise and those reporting a fall. Because it is a single figure, the balance statistic is often used to summarise the information in a survey; with a positive balance being associated with output growth and a negative balance associated with falling output.

The balance statistic is frequently used (informally) to quantify the extent of any growth or shrinkage. Here the balance may be plotted alongside the official estimates as in Chart 2, with a balance of zero associated with zero growth. This implicitly regresses the balance against the data, as in equation **(1)**:

$$Data_t = \beta \, Balance_t + \varepsilon_t \qquad \textbf{(1)}$$

The summary of the survey given by the balance statistic assumes (implicitly) that the average increase in output

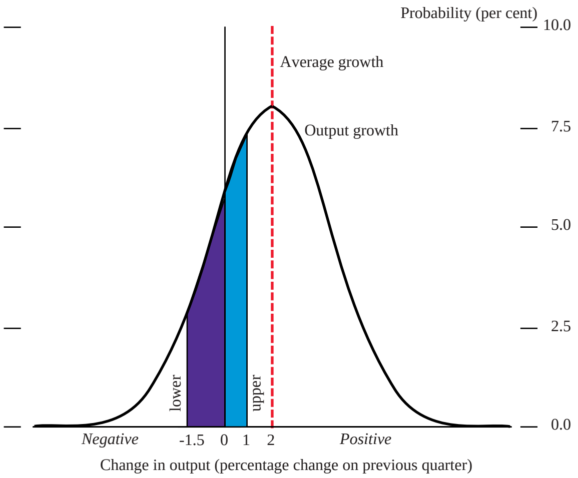
Chart 2
Example of balance estimate: applied to manufacturing output



reported as a 'rise' has the same magnitude as the average reported 'fall'. This will only be true if the range of outcomes that the firm perceives to be insignificantly below zero is the same size as the range perceived to be insignificantly above zero. In other words, the indifference band is symmetric around zero growth. But this will only normally occur if growth averages zero—a property violated by many macroeconomic time series.(1) So the balance statistic will generally be biased.

This strong conclusion follows from our interpretation of the indifference band as a confidence interval covering growth rates that the reporting firms perceive to be insignificantly different from zero. As the band is a confidence interval, the probability of growth being positive and insignificantly different from zero *must* equal the probability of growth being negative but insignificantly so. Chart 3 shows why, given this property of confidence intervals, the indifference band must be asymmetric when average growth is not zero.

Chart 3
Asymmetric indifference band



The chart plots a normal (bell-shaped) probability distribution of the firm's output growth over time. The average growth rate is 2% per quarter. In this case, the probability of output growth being between 0% and +1% must be greater than the probability of growth between 0% and -1%, because the positive range is closer to the average. So if we want to equalise the probabilities of growth falling within the positive and negative ranges, the negative range must be larger than the positive range (say from 0% to -1.5% as in Chart 3). Because we think of the indifference band as a confidence interval, we do want the probabilities to be equal, so the band must be asymmetric.

In line with previous research (see for example Pesaran 1984), this approach suggests that estimates based on the balance statistic will often be biased. In addition, the bias may vary over time. Any bias should fall as the proportion of firms reporting 'same' increases, because the bias derives from average reported 'rises' being different from average reported 'falls'. As the proportion reporting 'same' increases, there are fewer 'rises' and 'falls' to generate bias.

Analysis when the balance is biased

Despite the shortcomings of balance statistics, economists often choose to use them. Indeed, some surveys are only published as a balance statistic (for example the BCC's *Quarterly Economic Survey*), and so the user cannot distinguish between rises and falls. Even when survey data are published by category, the ease of presenting balance statistics makes them useful for *ad hoc* analysis. Although balance statistics may be biased, we can still gain valuable insights from them, especially if we can predict what any bias is likely to be and so correct (or allow) for it. This may be possible: our framework suggests some properties that any bias is likely to have.

Analysis of the CBI's *Industrial Trends Survey* suggests that the bias may be (fairly) significant. For example, a balance-based estimate of quarterly manufacturing output derived using equation **(1)** was, on average, 0.4 percentage points lower than an alternative using the 'best-practice' estimation technique described later. And since the 'best-practice' technique provides unbiased estimates, the 0.4 percentage points difference may be considered as bias in the balance estimate.

We can correct for average bias fairly trivially by including a constant in the balance regression, as in equation **(2)**:

$$Data_t = \alpha + \beta \, Balance_t + \varepsilon_t \qquad \textbf{(2)}$$

Because of the constant, a balance of zero may not be associated with zero growth. Any *ad hoc* assessment of survey balance statistics should allow for this.

Although we can correct for average bias by including a constant term, balance statistics may still distort the results, if bias varies over time. Our empirical work confirms that bias varies, though often not to an extent that would significantly affect our view of the trend in that variable. For example, in the application to manufacturing output, the variance of the bias was equivalent to just 3% of the total variance of output, as estimated using the balance statistic. But for some other variables, such as manufacturing export volumes, we found a greater variance, which could affect our conclusions.

Economic inference will be improved if we can predict when bias is likely to vary most. Again, the simple framework suggests an answer. The more uncertain firms are about their experiences and how to record them, the greater the variation of bias. This follows because any asymmetry is likely to become more marked as the width of the bands increases. And band width increases with uncertainty. If we think that uncertainty is related to the volatility of the economic variable, we may be able to 'predict' when the bias will vary most.

(1) Technically, the balance statistic may be unbiased despite a non-zero average growth rate, if output growth is not distributed normally across time.

Our analysis can be improved by using known properties of the bias in the balance statistic. But we could improve our analysis further by avoiding the bias altogether.

More rigorous estimation

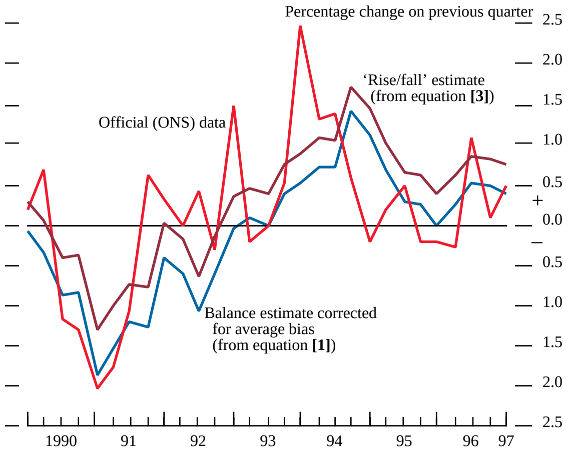
A slightly more complex approach (associated with Pesaran 1984 and 1987) uses the information contained in both the 'rise' and 'fall' proportions. In this approach, the official data is regressed against the proportions reporting in each category:

$$Data_t = \alpha + \beta_r RISE_t + \beta_f FALL_t + \varepsilon_t \qquad \textbf{(3)}$$

This approach does not impose symmetry. Indeed, it can be used to test for symmetry, which requires the coefficients attached to the 'rise' and 'fall' responses to be equal and opposite. Our research has rejected symmetry for the majority of survey questions that we have tested.

Chart 4 plots an estimate of manufactured output derived by applying this technique to the CBI's *Industrial Trends Survey* and compares it with the simple balance estimate from equation **(1)**. There are clear observational differences between the series—in particular, the balance estimate suggests flat output (zero growth) in 1996 Q1 while the 'rise/fall' estimate shows continued (albeit slowing) growth. The differences between estimates from the 'corrected' balance in equation **(2)** and the 'rise/fall' estimate are much smaller. But using equation **(2)** does not remove all of the bias from the balance statistic.

Chart 4
Example of 'rise/fall' estimate: applied to manufacturing output(a)



(a) Estimated for the period 1981 to 1994.

As predicted, the 'fall' term in equation **(3)** had a negative coefficient (so that greater proportions reporting falls were associated with lower output growth). The 'rise' coefficient was positive. We can reject symmetry because the 'fall' coefficient was significantly larger than the 'rise' coefficient (-0.07 compared with +0.01). The larger size of the 'fall' coefficient accords with our framework, because average output growth is greater than zero (as in the example in Chart 3).

Although equation **(3)** is an improvement on the simple balance model, all three equations share a problem. When a relationship is estimated using regression analysis, the explanatory variables appear on the right-hand side, and the dependent variable is on the left-hand side. The error term should be correlated with the left-hand side dependent variable but not the explanatory variables. Equations **(1)** to **(3)** all use survey responses as an explanatory variable, with the official data as the dependent variable. This may be the wrong way round. Intuitively, the survey response is being transformed to predict (or model) the official data. But in order to get unbiased and efficient estimates of the relationships between the survey and official data the survey data should be the dependent variable. This is because of the assumption that (after any revisions) the official data give an unbiased indication of the state of the economy, while survey data may contain measurement errors. In that case, the survey data should be on the left-hand side of any regression.

Of course, as noted earlier, most official data *are* subject to measurement error, since they are based on the experiences of a sample of firms. And it is possible that the official data may be biased—in other words that any measurement errors do not average zero. For example, the official data may pick up new firms with a lag and those firms' experiences may differ from the economy-wide average. In some cases private sector surveys may be less prone to such error.

Economists may wish to test whether the survey data give a better indication than the official data of the 'true' state of the economy. If they do, then survey data may *substitute* for official data. This possibility raises a number of interesting issues—in particular, how to quantify the survey data in the absence of a base against which to match it. But this article focuses on the best way to quantify survey data when it is used to *complement* official data.

When an equation is mis-specified by reversing explanatory and dependent variables, the results are not biased but the efficiency of the estimation process is reduced. In other words, any confidence intervals will be wider than they would be under efficient estimation.

An efficient estimator

Recent Bank research undertaken with Martin Weale (NIESR) and Richard Smith (Bristol University) has derived an approach in which the survey responses are treated as dependent variables. We set up two equations—one for the 'rise' and one for the 'fall' proportions. In each case, we regress the survey responses on the official data.[1]

(1) The rise and fall proportions must be within a range of 0% to 100%. To avoid violation of this range, the survey variables are subjected to a logistic transformation prior to estimation.

$$RISE_t = \alpha_r + \beta_r Data_t + \varepsilon r_t$$
$$FALL_t = \alpha_f + \beta_f Data_t + \varepsilon f_t \quad \textbf{(4)}$$

Once this system has been estimated, we can rearrange the equations to generate two transformations of the survey responses. Both give quantitative estimates of the economic variable, which we use to produce a single weighted average. The weights are chosen to minimise the variance of any errors in the estimate. The final estimate is simply a transformation of the survey data. It is not part of a behavioural model of the official data. Nor is the estimation an attempt to maximise the fit of an equation 'explaining' the official data.

In practice, the differences between the best-practice estimates and those derived using equations **(2)** and **(3)** are very small. But the best-practice approach should be preferred for any rigorous analysis: it does not require any further data; it is not much more complex to use; and it is potentially more efficient.

Problems encountered when survey data are smoother than official data

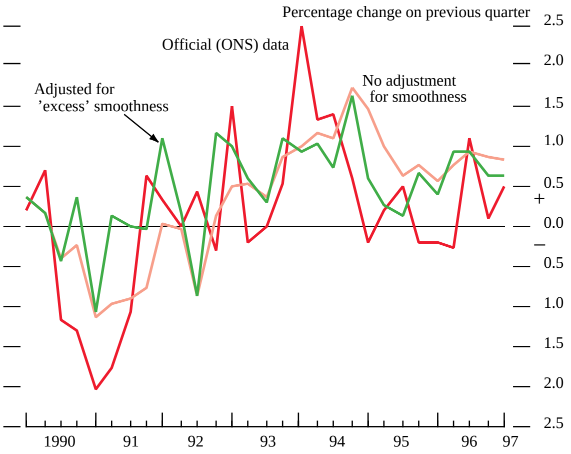
The quantitative estimates derived from surveys are often smoother than the official data they complement. As a result, we tend to find patterns in the residuals of the regressions used to match the survey data to the official data. These patterns, termed 'serial correlation', are a common problem in economic analysis. They may bias our estimates if the serial correlation is caused by omitted variables: perhaps the survey fails to pick up all the shocks to the economy and so is excessively smooth. In that case, we may add variables to 'absorb' the serial correlation.(1) If we think that the official data are more volatile than the (true) economic variable, despite being accurate on average, then there may not be an omitted variable problem and our estimates may not be biased.

Chart 5 compares our best-practice estimates of manufacturing output with and without an adjustment for serial correlation. The analyst needs to decide whether the surveys are too smooth, or whether the official ONS data are just more volatile than the 'true' data. This depends on what caused the relative smoothness. We put forward three possible explanations here.

Seasonal adjustment problems

Seasonal patterns and trends are a widespread problem in macroeconomic analysis, making it difficult to compare growth rates in different months or quarters. The ONS seasonally adjusts much of its data, and most surveys also ask respondents to allow for seasonal variations. But seasonal adjustment is complex, and neither the ONS nor survey respondents are likely to adjust perfectly. These problems may cause serial correlation when we try to match data from different sources:

Chart 5
Approaches to smoother survey data: applied to manufacturing output(a)



(a) Estimated for the period 1981 to 1994.

- Occasionally, the ONS adjustments may leave seasonal 'spikes' in the data. For example, the ONS seasonal adjustment of earnings data has been complicated by large and variable bonuses paid in the first quarter of each year.

- Survey respondents' attempts to adjust for seasonal variation may smooth the data more (or indeed less) than the ONS does. This is because any adjustments tend to be subjective. The CBI's investigation of answering practices found that only 26% of those who made adjustments did so 'by an established quantitative procedure'.

As a preliminary test of whether seasonal adjustment problems caused the serial correlation found in our empirical work, we re-ran our regressions after seasonally adjusting both the ONS and the survey data. Serial correlation was still present, suggesting other causes.

Firms infrequently update their responses

Surveys ask firms a number of questions, each of which may take some thought, and even research, to answer correctly. Firms may not be prepared to bear the cost of this research every time the survey is circulated. Or they may choose not to change their responses until the experience has changed significantly (perhaps reflecting their own uncertainty). If responses are only changed infrequently, then the survey estimates will be relatively smooth.

We have not devised a statistical test of this hypothesis. But the CBI's investigation of answering practices gave no indication of this problem (though it did not ask respondents explicitly about how often they reviewed their position).

Timing issues

Respondents to a survey may not record changes in output over the same period as the official data, even when the series purport to cover similar periods:

(1) We can absorb serial correlation by adding lagged dependent variables or estimating the correlation pattern directly.

- ONS data are recorded using a rigid set of rules. These define the period for which a change is recorded, for example comparing output on two days or averages for a quarter. Occasionally these rules may cause the data to be lumpy because output is recorded in discrete chunks. For example, in the Balance of Payments, exports and imports are only recorded on delivery, with no account made for progress payments. This may explain why trade data are relatively volatile.

- Survey respondents may smooth their responses by comparing recent experience with that a year earlier, even if asked about the trend during the past four months.

We have not devised a test of either ONS lumpiness or survey smoothing. But the preliminary evidence against other potential causes of serial correlation makes this explanation a likely focus for further analysis.

Surveys of expectations

This article has focused on questions about firms' past experiences. The same intuition and arguments can be applied to forward-looking questions. But there is a further problem to address: namely that there are rarely any quantitative estimates of expectations against which to match the survey data. There are two possible solutions:

(a) Assume that parameters estimated for the backward-looking responses apply to the expectations questions

Some surveys ask questions about both experiences and expectations. In this case, it is possible to estimate parameters for the backward-looking questions and apply them to the forward-looking survey responses.

But this procedure violates one aspect of the intuitive reasoning developed in our earlier discussion of how surveys work. The assumption that the parameters are the same is equivalent to assuming that the average indifference bands underlying the survey are the same in both the backward and forward-looking questions. But firms may attach a greater margin of error to their expectations of the (uncertain) future than to their perceptions of the past. In that case, the framework outlined earlier suggests that the indifference band will be wider in the expectations responses. Then the parameters of the forward-looking estimate should be larger. So imposing the backward-looking parameters will induce error—the estimate will have too little variance.

(b) Estimate parameters using the official data to model expectations

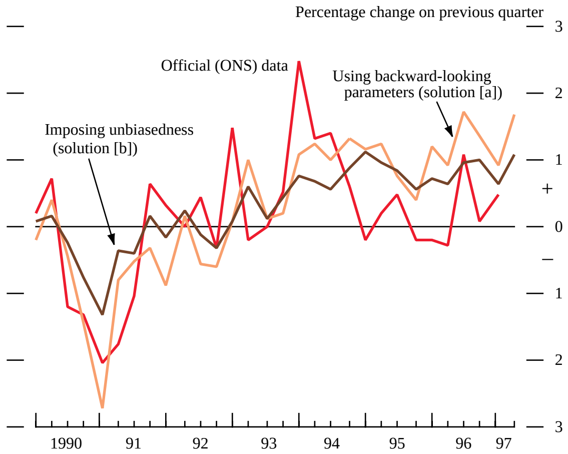
The alternative technique does not use information on respondents' past experiences. Instead, it makes an assumption about how expectations are formed, to derive 'expectations' against which to quantify the survey.

Individual firms are assumed to form expectations that are on average correct. In that case, expectations can be proxied by the official data on actual growth.

This method does not make any assumptions about the indifference bands in the forward-looking responses relative to those in the backward-looking responses. But the vital assumption that firms' expectations are on average correct has not been tested.

Chart 6 compares two estimates of expectations of quarterly output growth (derived using the two approaches) relative to the ONS estimate of growth in the same quarter. There are clear differences, which may cause concern, particularly if we wish to use the derived expectations for statistical analysis.

Chart 6
Approaches to expectations estimation: applied to manufacturing output(a)



(a) Estimated for the period 1981 to 1994.

Neither approach is entirely satisfactory. The first violates the intuition of our framework and the second makes an untested imposition. Perhaps the safest approach is to use both where possible. If they suggest similar results, then those results are at least robust.

Summary

To make the best use of the qualitative survey data to complement quantitative official estimates, survey data need to be converted as accurately as possible into quantitative estimates. This article has set out a simple framework to analyse the assumptions made in different techniques for making this conversion. Using this analysis, it has argued that the widely used balance statistics tend to give biased estimates and that, even when corrected for bias, the equations generally used are mis-specified and so reduce the efficiency of the estimation process. Re-specifying the equations results in a more efficient estimator, which should be preferred for economic analysis. Areas for further research include more work on serial correlation and on assessing how best to analyse survey data on firms' expectations.

Appendix

Estimation results

This appendix gives a brief description of the regression results underpinning the charts used.

All the regression techniques were applied to question 8 of the CBI's *Industrial Trends Survey*, which asks about changes in output. Since the survey covers manufacturers, the qualitative responses are matched to ONS estimates of manufacturing output. The variable *Data* is quarterly growth of manufacturing output.

The equations were estimated for the period 1981 Q1 to 1994 Q4. Later observations were omitted from the estimation because they may still be prone to revision.

Balance estimate—equation ***(1)***

$Data_t = 0.043\ Balance_t + \varepsilon_t$

R^2 0.37 — LM(2) serial correlation: 0.12

S.E. 1.04 — White heteroskedasticity: 0.62

Balance estimate corrected for average bias—equation ***(2)***

$Data_t = 0.38 + 0.04\ Balance_t + \varepsilon_t$

R^2 0.37 — LM(2) serial correlation: 1.58

S.E. 0.98 — White heteroskedasticity: 0.62

Separate rise and fall proportions—equation ***(3)***

$Data_t = 0.16 + 0.045 RISE_t - 0.035 FALL_t + \varepsilon_t$

R^2 0.37 — LM(2) serial correlation: 1.65

S.E. 0.99 — White heteroskedasticity: 0.32

'Efficient' estimator—equation ***(4)***

The 'rise' and 'fall' proportions have been given a logistic transform.(1) The transformed variables are denoted by *LRISE* and *LFALL*.

$LRISE_t = -1.26 + 0.24\ DATA_t + \varepsilon_t$

R^2 0.32 — LM(2) serial correlation: 9.87

S.E. 0.43 — White heteroskedasticity: 0.08

$LFALL_t = -1.26 - 0.27\ DATA_t + \varepsilon_t$

R^2 0.30 — LM(2) serial correlation: 16.3

S.E. 0.49 — White heteroskedasticity: 0.89

Note the serial correlation in the 'rise' and 'fall' equations.

'Efficient' estimator adjusted for serial correlation

$$LRISE_t = -1.02 + 0.11\ DATA_t + 0.43 LRISE_{t-1} - 0.24 LFALL_{t-1} + \varepsilon_t$$

R^2 0.75 — LM(2) serial correlation: 2.45

S.E. 0.27 — White heteroskedasticity: 0.41

$$LFALL_t = -0.83 - 0.08\ DATA_t - 0.24 LRISE_{t-1} + 0.59 LFALL_{t-1} + \varepsilon_t$$

R^2 0.82 — LM(2) serial correlation: 0.58

S.E. 0.26 — White heteroskedasticity: 0.61

Expectations estimator

Note that approach (a)'s system (page 298) was estimated for the backward-looking question, and so is identical to the 'efficient' estimator with no adjustment for serial correlation. The logistic transforms of expected rises and expected falls are denoted *LERISE* and *LEFALL*.

$LERISE_t = -1.25 + 0.14\ DATAt_{t+1} + \varepsilon_t$

R^2 0.24 — LM(2) serial correlation: 5.15

S.E. 0.32 — White heteroskedasticity: 0.48

$LEFALL_t = -1.66 - 0.25\ DATAt_{t+1} + \varepsilon_t$

R^2 0.30 — LM(2) serial correlation: 11.9

S.E. 0.49 — White heteroskedasticity: 0.81

(1) Under a logistic transform, LRISE = $ln\left[\frac{rise}{100 - rise}\right]$.

References

‘Good Statistics Guide’ (11 September 1993), *The Economist*, page 83.

Cunningham, A W F, Smith, R S, and Weale, M R (1995), ‘Measurement Errors and Data Estimation: the Quantification of Survey Data’, paper presented to Department of Applied Economics 50th Anniversary Conference, Cambridge.

Junankar, S (1990), ‘How Do Companies Respond To The CBI’s Industrial Trends Survey?’ in *Economic Situation Report*, Confederation of British Industry.

Pesaran, M H (1984), ‘Expectations Formations and Macro-econometric Modelling’ in Malgrange, P and Muet, P A (eds.) , *Contemporary Macroeconomic Modelling*, Basil Blackwell, Oxford.

Pesaran, M H (1987), *The Limits to Rational Expectations*, Blackwell.