QUANTIFYING DATA FROM QUALITATIVE SURVEYS*

The external demand for goods and services

Rui Esteves**

1. INTRODUCTION

In Portugal, the results of qualitative surveys on economic activity are disclosed on a monthly or quarterly basis. These results are indicators of economic agents’ expectations and behaviours, regarding various aspects of the developments in the economy in several activity sectors, and therefore are particularly relevant in short-term economic analysis.

The qualitative surveys on economic activity have many characteristics that make indispensable their utilisation in assessing the economy: their results are disclosed prior to quantitative statistics and, unlike these, the data are not subject to revisions in following releases; these surveys cover some sectors for which quantitative statistics are unavailable, or are infrequently released; they provide the only source of information on economic agents’ expectations, which allows to foresee the behaviour of the economy in the short-run.

In these surveys, industrials of companies in the sample indicate whether a given aspect of the company’s activity exhibited an increase, a decrease or relative stability in a certain period(1). The results are then presented in the form of shares of each of the three types of answers possible, and hence do not allow for an immediate interpretation of the evolution of a particular aspect for the surveyed sector as a whole. Frequently, the balance of respondents (i.e., the proportion of “increase” answers, minus the proportion of “decrease” answers) is calculated to condense the information of the surveys. However, since this statistic is an aggregation, it represents a loss of information (the share of “stable” answers is lost), and it builds on a very restrictive assumption: that an “increase” answer is worth as much as a “decrease” answer in relative terms.

The combination of qualitative information from these surveys with quantitative information on a common aspect of economic activity allows to understand the relationship between the proportions of the various answers and the state of economic activity. The methods of quantifying qualitative information meet these objectives. Many methodologies have been suggested in this sense. This article focuses on one technique that, despite being simple, yields widely acknowledged results. This technique builds upon the assumption that the real change in a surveyed variable is a weighted average of the average changes within each of the three types of answer.

The delay in the publication of quantitative statistics attributes great relevance to the quantification methods, as these allow to obtain estimates for the future behaviour of quantitative indicators.

This article analyses the relationship between qualitative and quantitative information, so as to estimate changes in quantitative statistics and to

---

* The opinions of this paper represent the views of the author, they are not necessarily those of the Banco de Portugal.
** Economic Research Department.
(1) The possible answers do not follow necessarily this scheme, since they vary according to the question. Answers can be either good / satisfactory / deficient, above normal / normal / below normal, rise / stabilisation / reduction, increase / stable / decrease, etc.
Questions can also regard expectations about a future period.
interpret more correctly the results of qualitative surveys, using a quantification method. To attain this, we use data on the external demand for Portuguese goods and services.

The remainder of this article is structured as follows: section 2 presents the qualitative and quantitative data used in the quantification process and the way these data were obtained and treated. A brief description of the quantification methodology used in this research is presented in section 3. The fourth section presents and comments the main results obtained through the application of this methodology. Section 5 concludes.

2. DATA

The indicators analysed in the article refer to the external demand for Portuguese goods and services on a quarterly basis. The analysis is based upon two kinds of information: qualitative and quantitative data.

Qualitative responses used in this research refer to the external order book, and are drawn from the Monthly Manufacturing Industry Survey (ICIT) of the Instituto Nacional de Estatística. Every month the entrepreneur in each company is asked the following question: “Do you consider that, given the season, your order book (or demand) from abroad is presently: above normal/nor-mal/below normal?”(2).

The Monthly Manufacturing Industry Survey is available in the current shape since June 1994. Data in the previous shape is available for the period running from January 1987 to March 1997. After analysing the proportions of the three answers in the period of overlap of both survey shapes, the correlation between both was calculated. This proved to be high, so the break adjustment was made by correcting the level of the answers in the previous shape. The break adjustment was made in the month for which the twelve following months recorded the highest correlation coefficient between the results in both surveys. The values in the previous survey had their levels corrected by the average difference between both surveys in the 12-month period scoring maximum correlation.

Although in the question on external demand responses should be given regarding the season, the series of the three types of answers exhibited a seasonal pattern, though not a marked one. Therefore, the qualitative answers were seasonally adjusted. This adjustment was carried out using the TRAMO and SEATS programs, assuming an additive breakdown.

Since the information of the qualitative surveys used in this study is released on a monthly basis, these data were aggregated to quantify quarterly qualitative indicators. Several aggregation procedures were tested to study if the relationship between the responses to the survey and the figures for external demand is a contemporary, led or lagged one. The quarterly qualitative indicators were built as 3-month averages of the proportions of each answer. These averages comprise not only the three months in the quarter, but also the three months beginning one to three months later (or sooner) to the beginning of the quarter — corresponding to one to three-month lagged (or led) variables(3).

We chose the year-on-year rates of change of external demand for goods and services in volume terms to relate with the proportions of the answers to the ICIT(4). The estimates for this indicator were built upon geometric means of the year-on-year real rates of change of the imports of seven countries(5), weighted by the relative share of each country in Portuguese exports. The values used are based on statistics drawn from Datastream.

This research covers the period running from the first quarter of 1988 and the last quarter of 1996. The period under scrutiny is limited as the ICIT was first published in 1987. Since the first observations of surveys are usually irregular — due to a period of adaptation of the respondents and

---

(2) Question of the Monthly Manufacturing Industry Survey version started June 1994.

(3) For instance, relating the quantitative variable to the 1-month lagged aggregate answers means that the last variables correspond to the aggregation of data for March to May, June to August, September to November and December to February.

(4) In addition to the year-on-year rates of change of external demand, also the year-on-year rates of change of Portuguese exports of goods and services in volume were used. The results obtained with this reference value were clearly worse, exhibiting a very low correlation with the ICIT answers.

(5) The countries here considered were: Germany, France, Spain, Italy, United Kingdom, United States and Japan.
the surveying institution — data from the first year of the survey were disregarded.

3. METHODOLOGY

This section outlines the methodology used in obtaining quantitative indicators from qualitative responses (6).

The method — hereafter referred to as MPR (method of proportions of responses) assumes that the expected value of the reference quantitative variable equals the sum of the mean of that variable, conditioned to each kind of response, times the probability of each type of response. It therefore is based upon the following identity:

\[
E(g_t) = H(g_t | A_t) Pr(A_t) + H(g_t | E_t) Pr(E_t) + H(g_t | D_t) Pr(D_t)
\]

(1)

where \(g_t\) is the reference quantitative variable, \(A\), \(E\) and \(D\) stand for the possible answers of the respondents (increase, stabilisation and decrease, respectively).

Assuming that the probability of respondents answering each type of response is given by the proportions of the answers to the qualitative surveys, \(a_t, e_t,\) and \(d_t,\) and that \(g_t = E(g_t) + \varepsilon_t\) where \(\varepsilon_t\) is random residual verifying the usual proprieties (i.e., \(E(\varepsilon_t) = 0, E(\varepsilon_t^2) = \sigma^2\varepsilon\) and \(E(\varepsilon_t, \varepsilon_k) = 0\) if \(t \neq k\)) and if \(\theta_{ij}\) stands for the expected values of the reference variable, conditioned to each type of answer, yields the following equation:

\[
g_t = \theta_{1j} a_t + \theta_{2j} e_t + \theta_{3j} d_t + \varepsilon_t
\]

(2)

Many are the hypothesis that can be made on parameters \(\theta_{ij}\) in estimating equation (2). One of these hypothesis is that the conditioned averages are constant in time. This being the case, (2) can be estimated through the ordinary least squares. Other hypothesis are such that the expected values of the reference variable, conditioned to the answer types, are allowed to change over time — e.g., are a function of other variables.

Two hypothesis on the behaviour of parameters \(\theta_{ij}\) were tested in this research. A first approach admits these parameters are constant over the course of the period under scrutiny. The second approach considers that the conditioned averages follow a random walk, that is:

\[
g_{t+1} = \theta_{ij} a_t + \theta_{2j} e_t + \theta_{3j} d_t + \varepsilon_t
\]

(3)

\[
\theta_{ij} = \theta_{ij-1} + \gamma_{ij}
\]

where \(\varepsilon_t\) and \(\gamma_{ij}\) are gaussian white noises.

Using the Kalman filter to estimate (3) we can obtain estimates for both the conditioned averages \(\theta_{ij}\) and \(g_{ij}\), based on the results of the monthly surveys.

4. RESULTS

The choice of the reference variable to quantify the results of the qualitative surveys is an extremely important step in the application of the above described methods. If good results are to be achieved, the quantitative variable used must refer to the same aspect economic agents are asked about, or should be strongly correlated to it. Hence, in most cases quantification of the survey results cannot be done due to the lack of reference variables, or sometimes due to errors of statistical measurement — specially in the reference quantitative variables, since these are based on complex sampling procedures (7).

In this research, the reference variable chosen to quantify the results of the question on external demand of the ICIT (8) was the time-series of the year-on-year rates of change of external demand for goods and services (9).

As mentioned in section 2, different aggregations of the monthly qualitative information were attempted, including monthly leads and lags in the process. In several models, the best results

(6) Another technique, based upon developments of the methodology initially proposed in Carlson and Parkin (1975), was also used. However, since the latter is more complex and does not give better results, it is not presented here.

(7) The estimates for the quantitative variables usually present greater errors than the qualitative answer proportions, since the former are obtained from more complex sampling and aggregation procedures.

(8) In the previous section, when presenting the methods generically, the three possible response choices considered were “increased”, “remained stable” and “decreased”. In this section we consider the admissible answers comprised in the ICIT: “above normal”, “normal” and “below normal”.

(9) Other variables could have been used. However, the chosen variable seemed the most adequate among the available data.
were obtained using the proportions of aggregate responses with a 1-month lag, which suggests that the external order books are basically met in the month following to their production. All results presented below were obtained by using response proportions aggregated in this fashion.

To estimate the relationship between the year-on-year rate of change of external demand and the proportions of the ICIT answers, we started by admitting that average changes within each type of answer (the \( \theta \) coefficients in (2)) are constant through time. The results show that a very strong relationship exists between the answers to the “external order book” of the ICIT and external demand directed towards the Portuguese market. The estimated values for the coefficients resulting from this estimation were 24.5 per cent, 10.1 per cent and -15.4 per cent for the proportions of the “above normal”, “normal” and “below normal” answers, respectively.

The hypothesis of having the reference variable’s average stable over time is at start quite restrictive, since no evidence suggests that such stability exists. The trend exhibited by the residuals of the adjustment may indicate a change in the average rates recorded in each group.

To quantify the qualitative results without imposing the constraint of having constant coefficients over time, these were allowed to follow a random walk. To attain this, (3) was estimated using the Kalman filter.

According to the estimation of (3), the average year-on-year rates of change of the “above normal”, “normal” and “below normal” answers in the last quarter of 1996 were 27.7, 7.7 and -13.7 per cent respectively, while the averages for the estimation period as a whole were 26.6, 9.3 and -14 . Chart 1 presents the values adjusted using the Kalman filter, compared with the values for the year-on-year rate of change of external demand. As expected, the adjusted and the effective values are quite similar, since the flexibility granted to the parameters allowed a great adaptation to the behaviour of the reference series.

Chart 2 exhibits the average year-on-year rates of change of each type of answer, estimated according to the Kalman filter. Note that, despite some variation, the estimated rates never diverge significantly from their mean.

Since one of the major advantages of the quantification methods consists of providing first estimates for quantitative variables yet to be disclosed, based on the results of monthly qualitative surveys, we analysed the one-step forecasts resulting from using the MPR with fixed and dynamic parameters, for the period running from the first quarter of 1996 and the third quarter of 1997\(^{10}\).

\(^{10}\)The year-on-year real rates of change of external demand for 1997 are subject to revisions.
The MPR with dynamic parameters yielded better forecasts, presenting a 1.75 p.p. square root of the mean quadratic prediction error, while in the fixed-parameter case this statistic was of 2.1 p.p.. As shown in chart 3, the forecasts obtained through the simple MPR are clearly biased, unlike the MPR estimated through the Kalman filter. However, the dynamic MPR always “misses” the sign of the change of the year-on-year rate of change of external demand — i.e., every time an acceleration took place, the predictor anticipated a slowdown, and vice-versa. Nevertheless, we can conclude that the MPR estimated using the Kalman filter provides better predictors of external demand.

The application of quantification methods allows not only to obtain quantitative indicators from responses to qualitative surveys, but also to infer about the behaviour of the respondents in relation to the questions that are asked. When applying the MPR with dynamic parameters to the quantification of external demand, the estimated average rates of change within each group of answers were 27.7 per cent, 7.7 per cent and -13.7 per cent in the last quarter of 1996, for the respondents answering to “above normal”, “normal” and “below normal”, respectively. On average, the respondents indicating “normal” recorded a quite positive change, while the changes of “above normal” answers are greater than those of “below normal” answers, in absolute value. This situation is understandable, since this is a variable which almost always exhibits positive rates of change. In these cases, the respondents tend to indicate the situation is normal when in fact what is being recorded is a growth. However, the average rate of variation of “normal” answers is greater than the year-on-year average rate of change of external demand (5.6 per cent). This means that individuals report a stable behaviour not only due to the “historical” growth of external demand, but also due to a certain pessimism.

These results evidence the main drawbacks of the balance of respondents as an indicator for economic agents’ perception of the behaviour of their activities. The balance of respondents disregard the proportion of “normal” responses that, as we have seen, indicate a growth in external demand. An analysis based upon balances of respondents builds on the assumption that all extreme answers should be weighted equally, which is also untrue since on average the rates associated to “above normal” answers are greater (in absolute value) than those associated to “below normal” answers.

The estimates for the parameters show that equal changes in the proportions of both extreme responses — yielding no impact on the balance — indicate that on average the change respondents face is not null. Considering the estimates of each parameter for the last quarter of 1994 (25.16 per cent, 9.28 per cent and -15.97 per cent), if the proportions of both extreme answers rose 1 per cent (hence reducing “normal” answers by 2 per cent), the year-on-year rate of change of the external demand for goods and services would decrease by about 0.1 p.p.. If in turn 1 per cent of respondents change their answer from “below normal” to “normal”, the year-on-year rate of change of the reference value records a greater revision (0.25 p.p.) than if this same proportion of respondents change their answer from “above normal” to “normal” (-0.16 p.p.) (11).

(11) Changes in the proportions of the responses that leave the balance unchanged do not always imply changes in the estimates for the reference variable. Only when $\theta = (\theta_1 + \theta_2)/2$, or a similar relationship holds do identical changes in the proportion of extreme answers yield null or negligible changes in the reference variable.
5. CONCLUSIONS

This article presents a method of estimation of the change in external demand using qualitative surveys on economic activity. The main conclusions are the following:

— the results of the qualitative surveys on external demand fit well the real year-on-year rate of change of external demand for goods and services;
— the answers on external demand of the ICIT provide an advanced indicator of the year-on-year rate of change of quarterly external demand, since the better results in quantification were achieved when quarterly aggregating the results of the 1-month lagged survey;
— the method based on the direct utilisation of the proportion of answers, with the parameters following a random walk, proved better than the constant parameter method as regards one of the major functions of quantification methods — short-term prediction of the reference quantitative variable. This is an important function, as quantitative variables are usually disclosed with considerable delay from the periods to which they refer;
— the balance of respondents has some drawbacks as an indicator of the behaviour of external demand. Responses indicating a “normal” order book indicate on average an increase. “Over normal” responses represent a change which is in absolute value greater than “below normal” responses. These facts reveal that economic agents are somewhat pessimistic in the appraisal of their external order book.

REFERENCES