THE USE OF QUALITATIVE INFORMATION FOR FORECASTING EXPORTS *

*Fátima Cardoso***

*Cláudia Duarte***

1. INTRODUCTION

The analysis of the evolution of external trade, in particular of exports, is very important in the case of small open economies, such as the Portuguese one. However, the release of external trade data, by *Instituto Nacional de Estatística (INE)*[1] only happens about 10 weeks after the end of the reference quarter. Due to this time lag, short-run exports forecasts play an important role in monitoring economic activity. Nevertheless, forecasting exports is not straightforward. Owing to the volatility of this series, forecasting exports entails a particularly high uncertainty level. Moreover, calculating timely forecasts requires the identification of variables that not only may bring useful information for forecasting exports, but are also timely released.

Survey data can be, in this context, particularly helpful for forecasting exports. According to an increasing literature on the use of qualitative data for forecasting quantitative aggregates, surveys are a very important tool for short-run economic analysis and forecasting. For example, Mourougane and Roma (2002), and Mourougane (2006) find evidence that confidence indicators are useful for forecasting, in the short run, the real rate of change of GDP, of some euro area countries (in the first case), and of Canada (in the second case). More recently, Cuenca e Millaruelo (2006) analyse the relevance of considering a set of qualitative indicators for assessing the behaviour of the main macroeconomic aggregates of the euro area. The use of survey data has drawn more and more attention mainly because these data are released with short lags and are not revised. Furthermore, in the case of questions about agents' expectations for the near future, survey series may be seen as leading indicators.

The aim of this article is to assess the role of qualitative information for forecasting the year-on-year rate of change of goods exports, in nominal terms.[2] In particular, for obtaining short-run exports forecasts, we estimate models in which only survey data are exploited (released by the European Commission) referring to the Manufacturing Industry Survey and to the Economic Sentiment Indicator.

The article is organized as follows. Section 2 describes the data. Section 3 presents the selection of an econometric model for the rate of change of Portuguese exports, in nominal terms, based on qualitative indicators, being its performance evaluated through an out-of-sample forecasting exercise. Finally, section 4 concludes.

* The opinions and findings expressed in this article are those of the authors and do not necessarily coincide with those of the Banco de Portugal. The authors thank António Rua and Paulo Esteves for their suggestions and comments. The usual disclaimer applies.

** Economic Research Department.

(1) Portugal´s National Statistical Institute.

(2) Initially both the nominal and the real rate of change were considered. However, further analysis showed that the results for the real rate of change were significantly worse than those for the nominal rate of change. This evidence suggests that, in the cases considered, the entrepreneurs' assessment, underlying their answers to the survey, is more influenced by nominal measures than by volume measures. Therefore, in this article are only presented the results for the year-on-year rate of change of exports, in nominal terms.

2. DATA

The database used consists of quarterly series covering the period from the first quarter of 1996 to the second quarter of 2006. The purpose of the analysis is to forecast the year-on-year rate of change of goods exports, in nominal terms, released by *INE*. As it can be seen in Chart 1, this variable denotes some volatility.

For forecasting the developments in exports qualitative series are used, both monthly and quarterly, referring to the Manufacturing Industry Survey, released by the European Commission. This enquiry, within the scope of the harmonised European enquiries, aims at collecting information about entrepreneurs' opinion regarding the evolution of their companies' activity. The survey, monthly delivered to a sample of about 1117 companies, consists of 18 questions (8 monthly and 10 quarterly). These questions can refer to *ex-ante* (next three months) or to *ex-post* periods (reference month or last three months). Considering the definition of the answer collecting period, the evaluation of the reference month is based on the first 20 days of that month.

From the set of 18 questions, it was made an initial selection, being only chosen the questions that are directly related to exports. In particular, the analysis encompasses four questions from the quarterly survey and another four questions from the monthly survey.

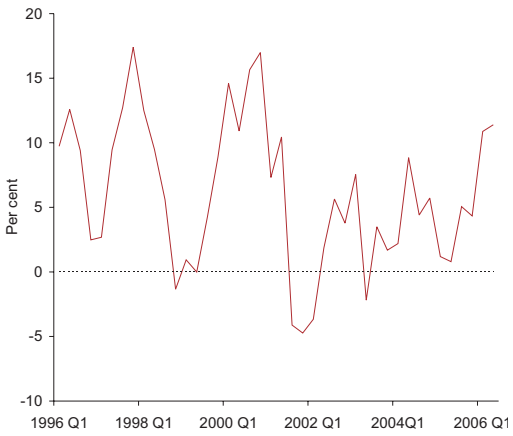
In the case of the quarterly survey, the questions chosen refer to new orders recently received *($nenc_t$)*, to export expectations for the next months *($eexp_t$)* and to the competitive position in the intra- and extra-community markets (pci_t e pce_t). Combining these last two variables, and using as weights the ratio of intra- and extra-community trade to the total external trade, a new variable was built ($comp_t$) which can be interpreted as a global indicator of Portuguese exports competitiveness.

Regarding the monthly survey, the questions analysed refer to the recent production trend *($prod_t$)*, to total and export order books ($cenc_t$ e $cexp_t$) and to production expectations for the next months *($eprod_t$)*.

The qualitative series are presented as balances, defined as the difference between positive and negative answers, and are seasonally adjusted. In addition to the aggregated totals, the variables from

Chart 1

YEAR-ON-YEAR RATE OF CHANGE OF EXPORTS, IN NOMINAL TERMS



Source: *INE* (International Trade Statistics).

the Manufacturing Industry Survey are also available disaggregated by economic activities (NACE).[3] The aggregation of the series for each question reflects the weights of each branch (measured by the turnover) in the total. Since, in this case, the aim is modelling and forecasting the year-on-year rate of change of exports, instead of only using the original overall series, it was also considered a set of series resulting from aggregating the disaggregates by economic activity, using as weights the ratios of the exports of each branch to the total exports (henceforth denoted by the suffix *p*).

Furthermore, adding to the variables from the Manufacturing Industry Survey, it was built a new variable, based on the Economic Sentiment Indicators of the several countries of the European Union (EU) *(pise$_t$)*. This variable, which can be seen as a proxy of the external demand of Portuguese exports, is a weighted average of the Economic Sentiment Indicators of the several Members of the EU, whose weights reflect the structure of Portuguese exports disaggregated by destination country.

In the context of a quarterly database, monthly series have to be transformed into quarterly series. In this case, the quarterly figures were obtained by averaging the monthly ones, considering more than one way of distributing the months among the quarters. Traditionally, the first quarter refers to the months of January, February and March, and so on so forth. Moreover, we also considered lags of one month (henceforth denoted by the suffix *d1*), corresponding the first quarter to December, January and February, and of two months (henceforth denoted by the suffix *d2*), with the first quarter equal to the average of November, December and January. This procedure allows testing the existence of several kinds of lags (in months) in the relation between the qualitative variables and exports.

3. MODELLING

3.1. First approach: cross correlations analysis

Considering the size of the database, in a first approach, bivariate correlations between the variation of exports and potential explanatory variables were analysed. From the results attained, it is possible to conclude that the correlations calculated from the weighted series, whose weights reflect the disaggregation by NACE of Portuguese exports, are, in general, higher than the correlations from the original data. Taking into account this evidence (subsequently confirmed by robustness checks during model estimation) it was chosen to work with the weighted series, whose weights are based on the Portuguese exports structure.

Analysing the correlations calculated from these series, it is possible to identify three different situations: in the case of export expectations for the next months *(eexpp$_t$)* the highest correlation is registered in a lagged period (t-2); for the variables referring to new orders recently received *(nencp$_t$)*, to export order books, contemporaneous and lagged one month (*cexpp$_t$* e *cexpd1p$_t$*) and to the proxy of external demand (*pise$_t$*) the highest correlation values are registered in the contemporaneous period; in all the other cases the highest value is most commonly registered in the period t+1 (see Table 1). This evidence suggests that the variables *nencp$_t$*, *eexpp$_t$*, *cexpp$_t$*, cexpd1pt e *pise$_t$* may be seen as coincident indicators (and lead indicator, in the case of *eexpp$_t$*) of the rate of change of exports. Nevertheless, even in the cases where the highest correlation is registered in periods after t, the value of the correlations in t and in lagged periods is still reasonably high.

In addition, the analysis of the table suggests that using the series from the monthly survey lagged one and two months does not seem to significantly improve correlations. In fact, the use of lags of less than

(3) NACE – Statistical classification of economic activities in the European Community. Note that, at least at the disaggregation level considered in this article (two-digit), there is a direct correspondence between this classification and the current Portuguese classification of economic activities (CAE).

Table 1

CORRELATION COEFFICIENTS

	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4
Manufacturing Industry Survey									
Quarterly survey									
New orders recently received *(nencp$_t$)*	-0.042	0.178	0.428	0.624	0.678	0.631	0.533	0.293	0.123
Export expectations for the next months *(eexpp$_t$)*	0.258	0.400	0.552	0.513	0.540	0.431	0.246	0.143	-0.094
Competitive position in the intra-community market *(pci$_t$)*	-0.122	-0.041	0.000	0.410	0.526	0.679	0.695	0.390	0.126
Competitive position in the extra-community market *(pce$_t$)*	0.161	0.195	0.201	0.238	0.497	0.577	0.555	0.317	0.134
Global indicator of competitiveness *(comp$_t$)*	-0.068	0.009	0.049	0.400	0.541	0.686	0.692	0.389	0.131
Monthly survey									
Recent production trend *(prod$_t$)*	0.032	0.239	0.429	0.517	0.591	0.593	0.469	0.448	0.227
Total order books *(cenc$_t$)*	-0.033	0.168	0.368	0.563	0.601	0.616	0.505	0.363	0.262
Export order books *(cexp$_t$)*	-0.216	0.132	0.447	0.683	0.773	0.666	0.446	0.107	-0.085
Production expectations for the next months *(eprod$_t$)*	0.204	0.301	0.475	0.463	0.522	0.563	0.529	0.516	0.314
Lagged one month									
Recent production trend *(prod$_t$)*	-0.042	0.206	0.327	0.508	0.556	0.598	0.512	0.429	0.333
Total order books *(cenc$_t$)*	-0.071	0.103	0.297	0.515	0.594	0.631	0.541	0.403	0.298
Export order books *(cexp$_t$)*	-0.249	0.017	0.349	0.632	0.756	0.728	0.540	0.216	-0.005
Production expectations for the next months *(eprod$_t$)*	0.139	0.311	0.394	0.464	0.489	0.544	0.593	0.489	0.417
Lagged two months									
Recent production trend *(prod$_t$)*	-0.064	0.105	0.274	0.484	0.548	0.612	0.539	0.477	0.388
Total order books *(cenc$_t$)*	-0.102	0.032	0.225	0.445	0.580	0.618	0.592	0.460	0.331
Export order books *(cexpt)*	-0.271	-0.110	0.234	0.555	0.716	0.750	0.617	0.312	0.040
Production expectations for the next months *(eprod$_t$)*	0.114	0.254	0.330	0.467	0.437	0.539	0.565	0.533	0.511
Economic Sentiment Indicators									
Proxy of the external demand *(pise$_t$)*	-0.171	0.046	0.325	0.600	0.737	0.704	0.479	0.279	0.062
Lagged one month *(pised1$_t$)*	-0.226	-0.030	0.226	0.519	0.707	0.739	0.559	0.351	0.123
Lagged two months *(pised2$_t$)*	-0.274	-0.107	0.134	0.431	0.661	0.754	0.639	0.411	0.200

Note: All correlations are calculated for the weighted series, whose weights reflect the Portuguese exports structure. The values refer to the correlation between the rate of change of exports in the period t and the qualitative indicators in period $t + i$ $\left(i \in \{-4, \ldots, 4\}\right)$. The shaded area denotes the highest correlation coefficient.

a quarter (one and two months) does not, in general, increase the correlation with the variation of exports, for lagged and contemporaneous periods.

3.2. Model estimation

The aim of this section is to choose the model specification, namely the variables to include, which allows attaining the best results in terms of forecasting the variation of exports. This choice takes into account the principle of parsimony, but, even so, we try to cover an information set as wide as possible.

To represent the rate of change of exports, *Autoregressive Distributed Lags* (ADL) models are considered, with the following form:

$$\theta(L)\, y_t = \sum_{i=1}^{k} \delta_i(L)\, x_{i,t} + \varepsilon_t \qquad k = 1,\ldots,n$$

in which y_t is the year-on-year rate of change of Portuguese exports, in nominal terms, $x_{i,t}$ are the qualitative indicators, $\theta(L)$ and $\delta_i(L)$ are lag polynomials and ε_t is the residual variable ($\bar{\varepsilon} = 0$ and standard deviation $= \sigma^2$). For model identification purposes it was used a sample from 1997Q1 to 2004Q2.

Due to the interest in obtaining timely forecasts, during model specification the qualitative variables were only introduced contemporaneously and with lags *(n = 4)*. Furthermore, adding to economic interpretation, model selection was based on the Schwarz criterion. Based on this criterion, the model chosen,[4] as follows,[5]

$$\underset{}{\hat{y}_t} = \underset{[-1.64]}{-25.34} \underset{[-5.00]}{-0.44 y_{t-4}} \underset{[1.92]}{+0.19 eexpp_{t-2}} \underset{[2.24]}{+0.22 cexpp_t} \underset{[2.44]}{+0.34 pise_t} \qquad (1)$$

Adjusted R^2 = 0.79

Dependent variable standard deviation = 6.33

Residuals standard deviation = 2.88

has as regressors: the dependent variable (rate of change of exports) lagged four periods (y_{t-4}), which enables the model to capture base effects; from the quarterly survey, export expectations for the next months, lagged two periods *($eexpp_{t-2}$)*; from the monthly survey, export order books *($cexpp_t$)*; and, lastly, the proxy of external demand of Portuguese exports *($pise_t$)*. There is no evidence of autocorrelation until the order 4 in this model.

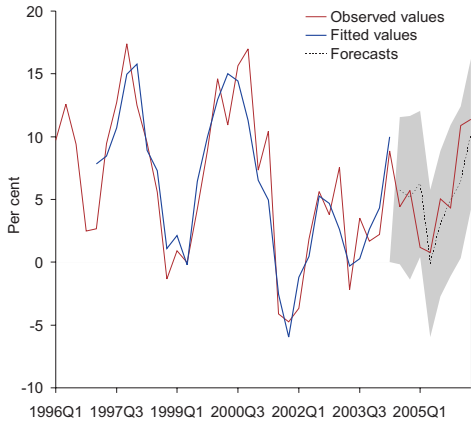
Considering the sample that was not used in model identification (2004Q3 to 2006Q2) it was simulated an out-of-sample forecasting exercise. The results of this exercise were evaluated according to the root mean squared forecast error (RMSFE), which, in this case, is 2.60 (see Chart 2).

(4) It is worth noting that other models, with similar results, were obtained. However, the model chosen presented the best fit in-sample and the best forecasting performance.

(5) The t-ratios are presented in square brackets.

Chart 2

YEAR-ON-YEAR RATE OF CHANGE OF EXPORTS, IN NOMINAL TERMS



Note: The shaded area corresponds to a 95% confidence interval.

Comparing the model chosen with a univariate benchmark,[6] one concludes that the model with qualitative variables has a better performance, both in-sample and out-of-sample, than the univariate model, which presents a residuals standard deviation equal to 4.59 and a RMSFE of 3.21.

In order to assess the robustness of the model to the sample used it was carried out a sensitivity analysis. According to the results of this analysis, the model appears to be robust to changes in the sample dimension. In particular, when estimating the selected model using the full sample (1997Q1 to 2006Q2) the coefficients remain stable and statistically significant.

Sample: 1997Q1 – 2006Q2

$$\hat{y}_t = \underset{[-1.66]}{-23.55} \underset{[-5.20]}{-0.44 y_{t-4}} + \underset{[2.41]}{0.17 eexpp_{t-2}} + \underset{[2.44]}{0.24 cexpp_t} + \underset{[2.51]}{0.33 pise_t} \qquad (2)$$

Adjusted R^2 = 0.77

Residuals standard deviation = 2.81

3.3. More timely estimates: indicators with incomplete quarterly information

The selected model, described in the previous section, presents the best forecasting performance, and gives one estimate to the rate of change of exports about 70 days before the release of the official data.

(6) The univariate benchmark has the following specification:

$$\hat{y}_t = \underset{[3.07]}{4.50} + \underset{[4.59]}{0.62 y_{t-1}} \underset{[-2.39]}{-0.32 y_{t-4}}$$

with t-ratios presented in square brackets.

In an attempt to obtain an estimate even earlier, alternative models were estimated in which, starting from the previously selected model (see (1)), the contemporaneous regressors are replaced with equivalent series that result from only considering information about a fraction of the quarter. Obviously, this is possible because, in the model chosen, the contemporaneous variables are from the monthly survey. Therefore, for both these variables (export order books *(cexpp$_t$)* and proxy of external demand *(pise$_t$)*), alternative series were calculated, where the quarterly observations are replaced by the data referring to the first month (henceforth denoted by the suffix *m1*) or to the average of the first two months of the respective quarter (henceforth denoted by the suffix *m2*).

Consequently, it is possible to obtain a first estimate of the quarterly variation of exports from a model with a specification similar to (1) but in which the variables referring to the export book orders *(cexpp$_t$)* and to the proxy of external demand of the Portuguese exports *(pise$_t$)* are replaced by the corresponding variables with information on the first month of each quarter only (*cexpm1p$_t$* e *pisem1$_t$*).

$$\underset{[-2.15]}{\hat{y}_t = -26.86} \underset{[-5.00]}{-0.44 y_{t-4}} + \underset{[2.18]}{0.18 eexpp_{t-2}} + \underset{[2.39]}{0.23 cexpm1p_t} + \underset{[3.12]}{0.36 pisem1_t} \quad (3)$$

In the same way, it can be obtained a second estimate using the variables with information on first two months of each quarter (*cexpm2p$_t$* e *pisem2$_t$*).

$$\underset{[-1.64]}{\hat{y}_t = -23.74} \underset{[-5.11]}{-0.44 y_{t-4}} + \underset{[2.21]}{0.18 eexpp_{t-2}} + \underset{[2.26]}{0.25 cexpm2p_t} + \underset{[2.51]}{0.33 pisem2_t} \quad (4)$$

Considering as a third estimate the one obtained from the model initially chosen, that is, from the model in which all indicators refer to complete quarters, Table 2 summarises the estimation and forecasting results over the three stages:

As expected, the results obtained, both for the RMSFE and for the accuracy on predicting direction changes, worsen as less information concerning each quarter is used. Nevertheless, by incurring this slight loss of quality it is possible to obtain more timely estimates. Furthermore, even considering the model with information only for the first month of the quarter, its results are still better than the ones of the univariate model.

The results obtained suggest that it is possible to have more timely forecasts for the variation of exports, without its quality being significantly affected, by using incomplete quarterly information. Moreover, including these variables in the model allows taking into account the monthly data as soon as they are released, without having to wait for gathering information for the whole quarter.

Table 2

MODEL COMPARISON

Models	RMSFE	Matching of the variation sign[a] (%)	Forecast timeliness[b] (days)
One month information	2.96	70.3	130
Two months information	2.85	73.0	100
Complete quarter information	2.60	73.0	70

Notes: **(a)** The matching of the variation sign measures the number of times that the sign of the variation of the series of interest (which is the rate of change of exports) is equal to the one of the variation of the fitted values and forecasts of the model (in percentage). **(b)** Forecast timeliness refers to the period of time elapsed between the calculation of the forecast and the release of exports data.

4. FINAL REMARKS

In this article it is assessed the importance of using qualitative data, from opinion surveys, for forecasting exports in the short run. In particular, a model for the quarterly year-on-year rate of change of exports is estimated, considering qualitative series from the Manufacturing Industry Survey, referring to Portugal, and the Economic Sentiment Indicators, referring to the EU countries, both released by the European Commission.

Based on the model selected, some alternative models are also estimated in an attempt to obtain estimates as early as possible, *vis-à-vis* the release of the data. The models presented produce estimates for the evolution of exports between 70 and 130 days before the release of the exports figures.

For the purpose of obtaining timely forecasts for the variation of exports, one concludes that the qualitative information is particularly relevant for this matter and that it provides a leading indication on exports' behaviour. Furthermore, the results obtained point to the stability of the model in relation to the sample used. It should be noted, however, that forecasting exports entails some difficulties due to the volatility of the series.

The favourable results that were achieved, in this case, can be seen as an incentive to further research on the use of this methodology for forecasting other quantitative aggregates.

BIBLIOGRAPHY

Artís, M. *et al*. (2003), *Forecasting models currently applied to indicators computed on the basis of surveys results*, European Commission.

European Central Bank (2001), "The information content of composite indicators of the euro area business cycle", *Monthly Bulletin*, November.

Britton, E. e Cutler, J. (1999), "The Bank's use of survey data", *Quarterly Bulletin*, May, Bank of England.

European Commission (2006), *Using factor models to construct composite indicators from BCS data – a comparison with European Commission confidence indicators*

Cuenca, J. e Millaruelo, A. (2006), "*Las encuestas de opinión y el análisis de coyuntura de la actividad real de la UEM*", *Boletin Económico*, January, Banco de España.

Hansson, J. et al. (2003), "Business Survey Data: Do They Help in Forecasting the Macro Economy?", The National Institute of Economic Research, *Working Paper* nº 84.

Mourogane, A. e Roma, M. (2002), "Can confidence indicators be useful to predict short term real GDP growth?", *Working paper* nº 133, European Central Bank.

Mourougane, A. (2006), "Forecasting monthly GDP for Canada", Economics Department *Working paper* nº515, OCDE.

Rua, A. (2002), "Composite indicators for the Euro Area economic activity", *Economic Bulletin*, September, Banco de Portugal.

Santero, T. e Westerlund, N. (1996), "Confidence indicators and their relationship to changes in economic activity", Economics Department *Working paper* nº170, OCDE.

Santos, R. (2003), "The use of qualitative data for short term analysis", *Economic Bulletin*, September, Banco de Portugal.