A TIME-SERIES STUDY OF THE FORMATION AND PREDICTIVE PERFORMANCE
OF EEC PRODUCTION SURVEY EXPECTATIONS

Dominique M. Hanssens*

and

Piet Vanden Abeele**

August 1984

Wettelijk Depot : D/1984/2376/11

*Graduate School of Management, University of California, Los Angeles,
California 90024, USA

**Department of Applied Economics, Catholic University of Louvain,
3000 Leuven, Belgium.

The authors acknowledge the support of the Commission of European Communities, Brussels, and the Research Center for Managerial Economics and Public Policy, UCLA.

Abstract

This paper provides an empirical analysis of manufacturers' survey responses on short term production expectations. If such surveys provide valid anticipations of production activity in customer industries, they should be valuable information to management in supplier industries.

Two major questions are considered: (1) are survey expectations valid predictors of future production activity (as reported later by the manufacturers and/or as measured by official statistics? and (2) do the expectations make efficient use of information available to the surveyed manufacturers?

The research design involves univariate and multivariate time-series analysis of monthly data for five European countries in three major sectors of the manufacturing industry. The scaling of the survey data is the subject of a special section.

1. INTRODUCTION

There is a long standing tradition in applied economics for the survey based observation of economic agents' attitudes and expectations, in particular those of consumers, businessmen or economic experts. Two kinds of attitudes are registered: judgment questions yield the assessment of current or past level (or change) in some economic variable; expectations yield an assessment of the likely future level (or change) in an economic variable. When expectations refer to outcomes over which the respondent has no control, they are called contingent; if the outcome is (partly) under the respondent's control, we obtain plans or intentional expectations. The objectively measured levels of the economic variable are referred to as the accounts.^a

The inspiration for these surveys is often pragmatic: the results may become available earlier than the official accounts data, thus they can readily be used to forecast economic activity such as consumer buying, manufacturers' investments or interest rates. Whether or not the survey based predictions are useful depends on two factors: (1) the respondents' efficient use of information, i.e. are they rational? and (2) the extent to which the expectations influence the agent's decisions, i.e. are the expectations self-fulfilling?

It is important to know the value of survey expectations because these data are very expensive and time-consuming to collect. For example, the annual cost of gathering production expectations data for the major industries in the Common Market countries is 50 million ECU's. Furthermore, there is a pragmatic alternative for making economic forecasts: econometric or time-series

^aThe terms judgments and accounts are not standard terminology but will be used throughout for ease of exposure.

models on national accounts data are now common and produce statistical forecasts whose reliability is known a priori. Thus the value of survey based expectations data can be challenged on absolute and relative grounds. This is the major objective of this research.

Survey based measures of economic judgments and expectations give rise to a number of research questions. Table 1 presents a simplified and structured overview of these questions in terms of bivariate relations among past and present judgments, expectations and accounts. The table is to be read column-wise, the column entry being the potential dependent and the row entry the potential independent variable. Three possible research interests are mentioned for each cell, depending on whether the focus lies on prediction (P), on the structure of the functional dependence (F) or on the common factor underlying both variables (C).

The table allows us to put our own empirical contribution in perspective. As described below, we will be using survey-based evaluations and expectations of businessmen concerning short term changes in production level. While the data allow the study of aspects of all questions mentioned in the table, the thrust of the analysis will deal with the following major questions:

- (1) how well do expectations forecast subsequent judgments? (cell 4)
- (2) how well do judgments and concurrent realizations correspond? (cell 2)
- (3) how well do expectations forecast subsequent realizations? (cell 11)
- (4) how are production expectations formed? (cells 7,8,9)

The questions in Table 1 can, as usual, be approached empirically by means of (individual) cross-section analysis or by (aggregate) time series analysis. Our study will be of the latter type and will adopt modern time-series analysis methods.

2. Literature Survey

Several studies using survey expectations data have been published to date. They generally address one or more of the key research questions listed in Table 1, but they also differ in (a) the economic attitudes and accounts considered (e.g. consumer confidence vs. business investment spending), (b) the nature, size and composition of the sample, and (c) the measurement properties of the data (e.g. interval vs. categorical scales, cross-sectional vs. longitudinal observations). For the purpose of this review, they are classified under forecasting performance vs. attitude formation papers.

2.a. The forecasting performance of economic anticipations data

A distinction is in order between surveys yielding point estimates of the anticipated level of an economic outcome (e.g. how many units do you expect to produce next month?) and those which do not (e.g. do you expect things to improve, get worse or remain unchanged for your company next month?) The former allow a direct (i.e. without fitting) retrospective investigation of their forecasting accuracy. The latter do not, since (1) it is not clear what underlying construct is being tapped and (2) the measurement is not on a scale with properties commensurate with those of candidate-dependent variables.

2.a.1. Point estimate anticipations

Let ${}_tX^*_{t+1}$ be the point forecast obtained as the expectation at t for the level of variable X at $(t+1)$ and let X_{t+1} be the actual level of the forecasted variable. A good forecast should have the properties of unbiasedness and efficiency. Unbiasedness should result in a $(0,1)$ estimate for the parameter vector (α, β) in the regression

$$X_{t+1} = \alpha + \beta {}_tX^*_{t+1} + u_{t+1} \quad (2.1)$$

The efficiency can be assessed in terms of the size of the forecast errors, absolutely or relative to the performance of alternative (e.g. naive) forecasts.

The Livingston data [Ahlers and Lakonishok 1983] have been a favorite testing ground for economic anticipations. They provide a consensus point forecast among a panel of economic experts of the changes in a number of economic series. Ahlers and Lakonishok find that the forecasts are biased over the 1947-1978 period. In terms of equation (2.1), the bias amounts to $\alpha > 0$ (understatement of trend component) and $\beta < 1$ (overstatement of deviations from the trend). The forecasts are not consistently able to outperform naive extrapolation models. The predictive performance of the direct forecasts is known to improve markedly over time, so that the survey forecasts generally outperform naive models in the 1970-1978 period, which can be characterized as rather volatile. The unbiasedness of the Livingston series has been tested in other papers. Brown and Maital [1981] point to weaknesses in OLS-estimates for (2.1) and, after a more appropriate estimation, find many of the series unbiased, thus confirming conclusions of others cited by these authors. Zarnowitz [1979] reports similar findings concerning forecasts of GNP change. Forecasts, whether survey- or model-based tend to beat naive and more complex time series forecasts, and direct survey forecasts perform about equally well as econometric model forecasts. At least for GNP-change forecasts, survey anticipations perform well, absolutely and comparatively. Zarnovitz subscribes the often heard conclusion that no single forecasting method can be shown consistently and generally superior to others.

Next to the direct forecasts of future contingencies, expectations surveys of a more intentional kind are conducted with businessmen concerning investment spending, sales, production, etc. Rippe and Wilkinson [1974] investigate 1 to 4-year anticipations of investment, sales and capacity and

show their predictive performance to be better than for naive extrapolation models or econometric functions. Similar findings were reported by Jorgenson, Hunter and Nadiri [1970] and by Liebling and Russell [1969] for the 50's and 60's. Liebling, Bidwell and Hall [1975] confirm the conclusion for the 73-75 period in a test of direct versus model (Chase, DRI, Wharton) forecasts. A number of other positive references are cited by Rippe, Wilkinson and Morrison [1976]. All these results are fairly encouraging and in line with the early findings by Modigliani and Weingartner [1958]: business survey anticipations (eventually after some bias correction) allow a rather good performance in forecasting. The time series models against which survey based forecasts were tested may be overly simple and thus overrate the comparative performance of survey measures. Joyeux [1980] is one of the few authors studying (business investment) anticipations in a time series methodology framework. She demonstrates Granger causality between accounts and anticipations, but the forecasting performance of the time-series relations in a holdout sample is not consistently superior to that of direct forecasts or of pure survey econometric equations.

Referring to Table 1, cells 11 and 12, the literature on the predictive value of survey based quantitative anticipations warrants the following conclusions:

1. the forecasts they provide are satisfactory, both in terms of absolute error size and in comparison with simple or more involved time series or econometric models.
2. their absolute (and comparative) predictive accuracy improves over time.
3. some correction for bias in level or regression may be useful.

2.a.2. Indirect anticipations

Economic attitude surveys do not necessarily yield quantitative forecasts in the guise of an anticipation for a well-defined economic outcome. In many

cases the survey data are categorical (e.g. production is expected to decrease, remain unchanged or increase) or are not direct indicators of specific categories of economic behavior (e.g. consumer sentiment data). Such data, eventually after rescaling, can be entered as regressors in econometric equations in order to explain economic phenomena; the fitted equations then also become available for forecasting purposes. Pure survey equations forecast economic outcomes only on the basis of survey data. Mixed equations contain both standard economic regressors and survey data; often, the researcher is interested in the marginal explanatory power or forecasting performance of survey data when added to pure economic regressors.

Adams and Duggal [1984] study the forecasting performance of three versions of the Wharton model, namely the version without, with endogenized and with exogenous (survey based) anticipations. They find that the forecasting performance improves in that order, with the strongest gain for the "hard" anticipations (business investment, housing starts) and with modest gains for "soft" consumer sentiment data. These conclusions were already arrived at earlier by Adams and Klein [1972]. Juster and Wachtel [1972], focusing on consumer spending and saving, conclude that consumer intention and sentiment data do come in significant as regressors next to standard economic variables. The marginal increase in the fit is rather small in most cases, however, while the "non-survey" part of the equations is fairly simple and the equations show substantially autocorrelated errors. The forecasting performance of consumer sentiment data is a much debated issue between proponents and opponents of "soft" (survey) measures. Rather negative evidence is obtained by Shapiro [1972] who demonstrates that consumer sentiment data fail to improve forecasts provided by econometric models of reasonable complexity. The experience with consumer intentions data in the U.S. was largely negative and lead to the termination of related survey efforts [McNeil, 1974]. An optimistic report is

given by Praet and Vuchelen [1983] on EEC consumer survey data. Their positive results may have to be mitigated by the consideration that their work involves extensive fitting. A constant difficulty with survey anticipations is the determination of the optimal lag in forecasting. When the survey question states the lag explicitly, this issue has mostly not been considered empirically. Vuchelen and Praet [1983] find that the prediction lag for consumer sentiment data is often rather short.

Referring to Table 1, the literature on indirect forecasts involving "soft" survey data suggests the following conclusions concerning cells 10, 11 and 12:

1. Pure survey equations will often allow a reasonable forecast of future economic outcomes.
2. The marginal fit or predictive performance of survey data in mixed equations is generally significant but often modest. The results are better for business investment and housing than for consumer spending/saving.
3. The optimal forecasting lag of anticipations data is an issue for empirical investigation.
4. A general assessment of the contribution of survey data, especially of consumer survey results, is rendered difficult by the varying nature of the models and results and by the extensive degree of curve fitting. A common benchmark for the assessment of the results is missing; a time series methodology could provide such a benchmark.

The previous conclusions are derived almost exclusively from aggregate time series tests. Cross-sectional studies of the relations between anticipations, judgments and accounts do not abound. Dunkelberg [1972] finds a significant though modest explanatory power of consumer sentiment data for

subsequent durable goods spending. In the business sector, Nerlove [1983] studies the relationship between short term anticipations (production, demand, prices, etc.) and the subsequent judgments for a German and a French sample of firms (Table 1, cell 4). The marginal distribution of anticipations and of subsequent judgments are found to differ. The conditional distribution of evaluations, given expectations, is temporally stable for one sample, but not for another.

These results indicate that in evaluating forecasts from anticipations, a correction for bias is needed and this correction may have to be temporally variable. The latter is done precisely by Dramais and Waelbroeck [1984] in an aggregate time series study of business production anticipations and of subsequent judgments (Table 2, cell 4). The last two studies are the only known to us studying judgment data. There is a general scarcity of studies on economic judgment, in contrast with the situation for anticipations. In particular, we are not aware of studies linking judgments to the corresponding accounts (Table 1, cell 2). This situation should cause concern, since business evaluations could serve as components of early leading indicators for the business cycle. One would expect a strong congruence between evaluations and concurrent accounts, barring which, reliance on anticipations would be dubious practice.

2.b. The formation of attitudes

Since economic judgments and anticipations are attitudinal data, psychological theory should yield valuable insight on their formation. Yet, excepting Katona's [1975] speculative discourse on macro-learning, this discipline seems to offer little guidance in the dynamics of aggregate economic attitudinal movements.

It is fair to assume that attitudes with respect to an object of some temporal variability (e.g. attitudes related to the business cycle) are based

on experience and insight and evolve rather smoothly. This intuitively acceptable point of view leads economists to model expectations as extrapolative, based on the past accounts of the series itself, on the past of the attitudes and eventually on the past accounts of other relevant series. The error learning model is a special and popular case of extrapolative expectations. It is found to give a good and parsimonious explanation of expectations by several authors, e.g. Nerlove [1983] for demand, price and production expectations and Carlson and Parkins [1975] for inflation expectations.¹ (Table 1, cells 7,9)

The rational expectations hypothesis looks at the informational basis of expectations in a normative way. In general terms, it states that economic agents form their expectations on the basis of all information available and that they combine this information in an optimal way. Various stringent tests of the rationality of expectations can be performed. Unbiasedness tests were already mentioned above, with mixed but rather encouraging results. From the literature [Ahlers and Lakonishok 1983, Friedman 1980, etc.] it appears that anticipations will generally fail to pass one or several of the tests of unbiasedness, efficiency, consistency or orthogonality of forecast errors. Especially concerning the latter, anticipations data seem to fail to incorporate all relevant information available. That survey based expectations fail the rationality test does not of course imply that they are irrational. Expectations have been shown above to often allow remarkably good forecasts, so that the rationality must be violated only mildly. Rational expectations is a hypothetical concept; "the" rational expectations formation mechanism will never be found. It can only be approximated and survey anticipations

¹In the case of Nerlove, survey expectations are related to judgments. Other studies, such as Carlson and Parkins, relate expectations to accounts.

seem to provide a fair approximation. Finally, all intricacies of testing for rationality seem to be far from resolved at this point [Brown and Maital 1981]. Nerlove [1983] states that the stringent rationality tests may be unreasonable and proposes the weaker criterion that expectations be shown to present no pattern of systematic error. He adds that when economic outcomes are not undergoing structural change, one will generally find a univariate time series relationship between a variable and its own past; then quasi-rational expectations (i.e. containing no pattern of systematic error) can generally also be generated as a time series forecast from the past realizations of the outcome concerned.

The literature offers some studies of expectations in a univariate or multivariate time series perspective. Joyeux [1980] regresses business investment anticipations on their own past, but does not study their formation deeper. Dramaïs and Waelbroeck [1984] apply an ARMA model to an evaluation series and find its forecasting RMSE generally larger than for econometric forecasts on the basis of past expectations. Praet [1984] regresses consumer sentiment indicators on their prior one or two lagged values and finds R^2 values in the order of .90; objective economic indicators can improve the fit only marginally in view of the already excellent fit. In a time series perspective, his results show that the consumer sentiment series can be explained by simple ARMA models. All of these results confirm that anticipations evolve rather smoothly and can be well fitted from their own immediate past. (Table 1, cells 7 and 9)

A number of other studies have investigated the formation of economic attitudes outside of a strictly time series framework [Shapiro 1972, Adams and Klein 1972, Foster and Wachter, 1972, Nerlove 1983, Hymans 1970]. These regression models usually incorporate the first lag of the survey attitude as well as a number of economic indicators as explanatory variables. The results

of the later studies as well as those mentioned earlier on the formation of economic attitudes (Table 1, Cells 3, 4, 5, 7, 8, 9) seem to warrant the following conclusions.

1. Survey attitudes can be explained to a significant extent by prior or by concomitant economic indicators. These indicators are often those known or expected to explain the accounts corresponding to the attitudes. The forecasting lag on the indicators is often very short. These findings can be taken as evidence for rapid updating in attitude formation.
2. Survey attitudes can usually be well explained by their own past.
3. One has reason to doubt that the variance in economic attitudes not explained by the own past or by economic indicators is useful for prediction purposes.

2.c. Conclusion of the literature survey and research objectives

Our assessment of the literature differs on whether a pragmatic or a theoretic stance is adopted. Pragmatically, the conclusion is that survey based attitude data allows good or even improved forecasts and that the survey attitudes themselves can be forecasted. On the more fundamental side, it is not clear that, or why, survey data are superior to alternatives. While the latter issue will be clarified as the result of further investigations on new data, advances in knowledge will also hinge on considerations of method. In particular, we feel that there is more need for the following types of research:

1. Cross-section studies at the level of the individual respondent vs. aggregate time series studies. Attitudes are individual variables. Their formation and decision impact takes place in the individual before it shows up in the aggregate. Aggregate time series analyses can obscure the individual nature of the processes for a variety of reasons.

2. Time-series studies as opposed to econometric fitting. The latter explodes one to the dangers inherent in spurious correlation, in curve-fitting and in general, to the dangers of an overly generous attitude towards causality. Most of the time series dealt with here are strongly autocorrelated, so that the modern time series methodology is indicated. A restrictive stance towards causality, as embodied in the Granger causality concept is recommended.
3. Construct validation studies in the psychometric tradition. When working with survey responses one should guard against naive faith in one's measures. The multiple survey and other variables should be subjected to convergent/discriminant validity tests in order to learn what it is that one has actually measured.

Our empirical research investigates the formation and predictive value of short term business attitudes obtained from surveys. These short-term attitudes have received relatively little scientific treatment up to now, so that a first contribution lies in expanding our insight in the value of these particular attitudes. A second contribution lies in the application of a research approach which meets the last two needs and especially the second need above. A third contribution lies in expanding our knowledge of the value of economic attitudes in general and of the appropriate research approach in particular.

A final contribution can be found in the multi-sector/multi-nation character of our results, which allows empirical generalization.

3. Econometric and Time Series Methodology

Since the popularization of time-series theory by Box and Jenkins, a controversy has arisen over the use of "classical" econometric vs. time-series

methods in empirical economics. Briefly speaking, the "pure" time-series analysts use statistical methodology for the specification and the estimation of multivariate models, whereas "pure" econometricians employ these techniques only after a suitable model has been specified on a priori grounds.

In our opinion - and definitely within the present research context - an intelligent combination of the two is preferable to either extreme. First, it is possible to specify the following simple economic model of production activity:

current production = $f(\text{previous orders, previous expectations})$

i.e. production (P) is predicted by known demand (orders O) and subjectively estimated future demand (expectations E). Second, it is difficult to specify the dynamics of this model on a priori grounds alone; furthermore, the indivisibilities in industrial production make the data series fairly smooth over time, so that current levels may be strongly related to previous ones. Thus we need time-series methods to operationalize the economically plausible model on monthly data. A similar argument holds for expectation formation: if expectations are related to actual production, then they too should be smooth over time.

The tests on the various models are based on the concept known as Granger causality (Granger 1969). To summarize this concept, X is said to Granger cause Y with respect to an information set containing X and Y, if the error in forecasting Y from its own past and the past of X, is lower than the prediction error when only the past of Y is used. This definition applies well to expectations modeling: that current production is partially predictable from previous expectations seems intuitively obvious. However, the critical issue is whether or not this predictive quality holds over and above simple time series extrapolation of past production data. In other words: do the survey respondents know and reveal something about the future that cannot be captured

by statistical analysis of production data? Since the collection of survey data is costly and time-consuming, it can be justified only if survey expectations Granger cause industrial production.

The successful execution of a Granger causal analysis depends on specifying an adequate information set, and using statistical methods to disentangle the intrastructure (i.e. past of Y) and the interstructure (past of X) in the data. Our information set includes perceived production and orders, the expectations (i.e. the survey data), and the production accounts data. This is a fairly complete information set given that accounting data on orders are not available.

Statistical methodology for separating intrastructure and interstructure was developed mainly in the seventies. One popular method correlates the residuals of univariate Box-Jenkins analyses on the series of interest at various lags (Haugh 1976, Pierce 1977). This method is very efficient for removing intrastructure, but is restricted to pairwise interstructure analysis, which may be dangerous. A second method collects the Box-Jenkins residuals in a multiple regression equation, called a "dynamic shock model" (Haugh and Box 1977). This is a comprehensive approach, which is used infrequently because of practical problems: the lag structures may be distorted, the findings may be sensitive to the choice of a Box-Jenkins prewhitening model, and the operation is cumbersome and time-consuming. Lastly, one may use traditional least squares estimation in multiple regression models containing the past of Y and X (e.g. Sims 1972, Granger 1971). For example, for the information set $\{Y, X, \}$:

$$Y_t = d + \sum_{k=1}^{\infty} a_k Y_{t-k} + \sum_{\ell=1}^{\infty} b_{\ell} X_{t-\ell} + u_t,$$

X would Granger cause Y if $\{b_{\ell}\}$ are statistically significant. If maximum lags K and L can be chosen without causing truncation bias, and if a sufficient

number of degrees of freedom are available, then these tests can be executed fairly safely. On the matter of testing the significance of $\{b_{\ell}\}$, some Monte Carlo simulation by Geweke, Meese and Dent (1983) points to the superiority of Wald's Chi-square test over other likelihood ratio tests.

We will use the last method of Granger causality testing. Since this technique involves multiple regression analysis, the results must be interpreted with caution so as to avoid unstable or spurious findings. Therefore the following precautions are taken in the analysis:

1. the models are estimated on stationary data, possibly after applying a stationarity-inducing transformation. In addition, possible collinearity among the lags is reduced by prefiltering the data with a filter which removes the highest common autoregressive order (e.g. Liu and Hanssens 1982). The state-of-the art in collinearity and outlier diagnostics is used to monitor the stability of the results (Belsley, Kuh and Welsch 1980);

2. the cumulative effect of any regressor, e.g. the sum of $\{b_{\ell}\}$ must have a plausible sign;

3. the in-sample results are supplemented by a 24-month forecasting test, using 1982 and 1983 data. Comparisons are made on the basis of the one-step mean squared forecast error.

4. Data and Scaling Issues

The empirical analysis is based on data from the European Community Business Survey. This mail survey is carried out by national institutions in the respective EEC member countries on a monthly basis. The EEC harmonizes the questionnaires between the member countries so as to insure the cross-national homogeneity of the measurement instrument and of the results. Each month a sample of some 20,000 enterprises is surveyed.

Five countries with sufficiently long time series for the survey data and the national accounts data were selected for analysis. The sample includes France (from June 1968), Germany (from January 1969), Belgium (from January 1970), Italy (from January 1971) and Holland (from December 1972). In all cases the last month reported was December 1983, and a common 24-month forecasting sample (1982-1983) was set aside.

In addition to investigating five different countries, the study also considers three different sectors of the economy: the consumer, investment and intermediate goods sectors. This choice was based on the fact that the accelerator effect is likely to make the investment goods sector more volatile than the intermediate sector, which in turn should be more volatile than the consumer sector. Previous studies have tended to pool economic sectors and even countries, which may create some aggregation bias in the reported findings.

Our main interest is in the "production expectations for the months ahead" question, the answers to which are recorded in a trichotomous way (up, unchanged, down). In addition, we consider the questions asking for a judgment of production trends in the last month (up, unchanged, down) and for the judgment of current order books (above normal, normal, below normal).

The data used for this study are aggregate response percentages under each response category of the trichotomy. Since the results do not mention "don't know" or "missing" categories, these percentages sum to 100 in every case. For expository convenience, the series containing the percentage responses mentioning an increase will be labeled "UP", a decrease "DO" and no change "EQ".

Survey responses are data of a different kind than the "objective" data economists are used to and require consideration of their nature before one proceeds to econometric or to time series modeling. The major issues are the substantive meaning of the responses and their measurement and scaling properties.

The survey answers should not necessarily be taken at the face value of the questions: the actual meaning of the responses should be revealed, at least in part, by validation studies. Econometric and time series models on survey response data are to some extent validation attempts rather than only causal or predictive investigations. Some more limited face validity questions can be examined directly. For example, expectation and judgment questions are worded to ask for a comparative assessment of production levels, making abstraction from seasonal fluctuations, and should thus be free of seasonal and of level (trend) effects. These are properties one expects from these measures. We suggest to make these properties a matter for empirical investigation.¹ Our experience is that even these simple properties cannot be established for the survey data,

The survey results used for our study are the percentage distribution of responses over three response categories (UP, EQ, DO). Each survey question thus gives rise to three time series of percentages. The percentages are constrained to sum to 1.0 for each observation. Irrespective of the number of response categories, this situation is typical for survey work. It confronts the researcher with multiple, non-orthogonal series to represent the responses to a single survey question and, eventually, with the need to summarize such series in a meaningful and efficient way.

If the response percentages were drawn "randomly" subject to the sum constraint, their joint distribution would be Dirichlet with parameters a_{UP} , a_{EQ} , a_{DO} , and with means

¹The same remark holds for the time span in expectation or judgment formation. Combining monthly data to yield quarterly survey response indicators as in Nerlove [1983] or in Dramais and Waelbroeck [1984] should be done only after an empirical verification of the nature of the crude data, if at all.

$$t_i = a_i / (a_{UP} + a_{EQ} + a_{DO}) \quad (i = UP, EQ, DO)$$

Under this hypothesis one obtains the following useful results [Wilks 1962]

$$(a) \quad \text{var}(i) = r t_i (1 - t_i) \quad (i = UP, EQ, DO) \quad (4.1)$$

$$\text{cov}(i, j) = r t_i t_j \quad (i \neq j; i, j = UP, EQ, DO) \quad (4.2)$$

with $r^{-1} = (a_{UP} + a_{EQ} + a_{DO} + 1)$ a measure of dispersion

$$\text{hence} \quad \text{corr}(i, j) = -t_i t_j / \sqrt{t_i (1 - t_i) t_j (1 - t_j)} \quad (4.3)$$

(b) for any $i \neq j \neq k$ ($i, j, k = UP, EQ, DO$)

$$i/(i+j) \text{ is distributed independently of } k \quad (4.4)$$

Under the Dirichlet assumption the percentage series UP, EQ, DO will obviously not be uncorrelated. While the empirical correlation between any two of the three series may be nonsignificant, this will not reduce the overall degree of collinearity among the three series. This non-orthogonality leads to the question of which variables should be used in regression analysis when there are multiple indicators for a single survey question. Three strategies are available (1) use some or all of the original series, (2) use linear transformations of the original series or (3) use nonlinear transformations of these.

4.1 Use of original series

Since the original UP, EQ and DO series sum to one, all intercept regression models incorporating two series as regressors will be observationally equivalent. If only one series is used, the selection should favor the one which is most meaningful and which summarizes the most information. The distinction between directional series (UP, DO) and nondirectional series (EQ)

is relevant. The researcher will normally be interested in predicting a directional (change) dependent variable for which directional regressors have high face validity. The empirical correlations in our data show that the UP and DO series are substantially related, so that one of them picks up much of the information in the other. If one series has been selected as a regressor, the information contained in the other series should not necessarily be discarded, it can be used e.g. to build variable elasticities for the survey or other data.

4.2. Linear transformations of original series

Linear transformations of the original percentage series have been the favorite way of reducing multichotomous response percentages into single indicators, the transformation being developed a priori or empirically. A priori transformations amount to assigning a value to each response category and to computing a weighted average. In the particular case of our data, the balance transformation

$$BAL = UP - DO \quad (4.5)$$

is a favorite way to summarize the directional information. It is equivalent to a weighted average of the responses $UP=1$, $EQ=0$ and $DO=-1$. The BAL series can be complemented by the non-directional EQ series as a second variable.

Under the Dirichlet assumption outlined above, the correlation of BAL with EQ will be

$$\frac{t_{EQ}(t_{UP} - t_{DO})}{\sqrt{t_{UP}(1-t_{UP}) + t_{DO}(1-t_{DO}) + 2t_{UP}t_{DO}} \sqrt{t_{EQ}(1-t_{EQ})}} \quad (4.6)$$

Considering that UP and DO are trendless data over a sufficiently long period, t_{UP} and t_{DO} should be nearly equal and thus BAL and EQ should be nearly orthogonal under the Dirichlet assumption. While our empirical correlations

between BAL and EQ are regularly significant as discussed below, the extent of collinearity is modest enough not to preclude the use of these variables jointly as regressors.

Empirically derived linear combinations will be based e.g. on the principal components of the data. The attractiveness (orthogonality, efficient summary of information) of such transformations should be traded off against the lower face validity of indicators thus developed.

4.3 Nonlinear transformations

Many nonlinear transformations could be suggested. An intuitively attractive transformation considers the fraction of positive answers in the total of directional answers.

$$POS = UP / (UP + DO) \quad (4.7)$$

The POS series is theoretically independent from the EQ series, which can be seen as an advantage. A potential drawback is that POS, in principle, is independent of the size of the balance between UP and DO (i.e. POS has the same value for 2 % UP vs. 1 % DO as for 20 % UP vs. 10 % DO). As a result, the POS series may need to be complemented by such variables as EQ or as the absolute value of the balance. Dramais and Waelbroeck [1974] develop the Majority-variable, closely related to POS, as follows: Majority = POS if POS $\geq$.50, otherwise Majority = POS - 100%. The authors do not provide a rationale for this transformation. Compared to POS, one finds that Majority has become more of a dichotomous variable, since it will range between -1.0 and 1.0 excluding the -.5 to .5 interval. We fear that this transformation reduces the information in the data excessively, while at the same time exaggerating the correlation between two likewise transformed variables.

All the abovementioned transformations have a limited range, which makes them less suitable as dependent variables. Nonlinear transformations based on

the assumption of a continuous latent variable and a discriminial process allow us to circumvent this problem [Bechtel 1981, Carlson and Parkins 1975]. Let x_{it} represent the interval scaled attitude of individual i at time t , with $f_t(x)$ the density of x_t over the population and $F_t(x)$ the corresponding cumulative density. Let the discriminial process be such that there are two cutoff points X_{DO} and X_{UP} ($X_{DO} < X_{UP}$) on the attitude continuum and that a DO response occurs if $x_t \leq X_{DO}$ and UP response if $x_t > X_{UP}$ and an EQ response if $X_{DO} < x_t \leq X_{UP}$. Since the attitude continuum may be interval scaled, we can arbitrarily set $X_{DO} = 0$ and $X_{UP} = 1$ and write

$$DO_t = F(0) \quad (4.8)$$

$$DO_t + EQ_t = F(1) \quad (4.9)$$

Since DO, EQ and UP provide only two independent information elements, a two-parameter function $f_t(x)$ can be inferred from the data. If one opts for a symmetric two-parameter function, its central tendency and its dispersion provide us with two meaningful and potentially independent indicators. The logit relationship

$$F_t(x) = [1 + e^{\alpha_t + \beta_t x}]^{-1} \quad (4.10)$$

is an attractive formulation for a two-parameter cumulative density. Its parameters α_t, β_t can easily be inferred from the DO, EQ and UP data using (4.8) and (4.9). The central tendency of f_t is located at x_{ct} such that $F_t(x_{ct}) = .50$, namely $x_{ct} = \alpha_t / \beta_t$. The slope of $F_t(x)$ at x_{ct} serves as a measure of concentration of $f_t(x)$ around x_{ct} and equals $\beta_t / 4$. While these measures of central tendency and of concentration (denoted as CENTRAL and AGREE for the next section) are independent in principle, empirical correlations between them cannot be ruled out.

4.4 Study of empirical correlations

The previous discussion shows that varied set of indicators can be suggested to represent multichotomous survey responses. Some empirical results are helpful in order to reduce the complexity of the problems. These results are based on a study of correlations between the variables UP, DO, EQ, BAL, POS, MAJORITY, CENTRAL and AGREE for the question on production expectations as measured for three industrial sectors and five countries. Since all of these correlations are based on more than 100 observations in general, they are fairly reliable. Such a study shows the following results:

1. the correlations between BAL and CENTRAL range from .97 to .999.
2. the correlations between BAL and POS range from .91 to .99 with an average of .97. These results indicate that one should be indifferent between BAL, CENTRAL or POS.
3. the correlations between BAL and MAJORITY range between .74 and .92, with an average correlation of .87. While BAL and MAJORITY are substantially correlated, they are not virtually identical. This may be due to the fact that they contain different systematic information; we suggest that it may also be the result of downgrading of information in the MAJORITY transformation.
4. two nondirectional measures, EQ and AGREE have been investigated. When choosing either one to complement BAL or MAJORITY, one will most likely select the least collinear complementary variable. The correlations show EQ to be less collinear in general with either BAL or MAJORITY. The correlations of EQ with BAL range between .04 and .57 in absolute value, with an average (absolute value) around .2; the correlations of EQ with MAJORITY range between .01 and .39 with an average around .15.

The empirical data indicate that there is no clear evidence in favor of one particular set of variables or transformations. The practice of using Balance series in order to indicate direction, eventually supplemented by the EQ series as nondirectional data is certainly not worse than any other approach and has the advantage of remaining close to the meaning of the original survey data. As our work also uses attitudinal data as dependent variables, we prefer the central tendency transform which is not constrained in range and is almost perfectly correlated with balance data.

5. Empirical Results

5.a. Forecasting Production Judgments

The first criterion of usefulness of survey expectations is that they predict subsequent judgments, or manufacturers' perceptions of production activity. In a Granger causal context, we define an information set of three survey questions: production judgments, order judgments and production expectations. Thus, the implicit model is:

$$\text{CURRENT PRODUCTION JUDGMENT} = f [\text{PAST PRODUCTION JUDGMENTS, PAST ORDER JUDGMENTS, PAST EXPECTATIONS}]$$

and is hypothesized to outperform the univariate model

$$\text{CURRENT PRODUCTION JUDGMENT} = g [\text{PAST PRODUCTION JUDGMENTS}]$$

The inclusion of past orders in the information set is justified because orders are logically related to future production. Since this information can only be obtained via the survey, this variable is treated in the same way as the expectations.

Following the results of the scaling investigation, the survey results were transformed to an overall sentiment level (central tendency) and a sentiment homogeneity (agreement) score. The logit-derived central tendency was used because it is not range-constrained, and thus more appropriate in a regression context than the balance data. The sentiment homogeneity variable was initially included in the models, but never contributed significantly to the results. In addition, this variable causes collinearity problems, presumably because it lacks variance over time. For example, the percent "status quo" opinions in Germany centers around 74, with a very small standard deviation. Therefore, all opinion variables are measured with one instrument: the central tendency of the sentiment.

Univariate Box-Jenkins analysis on the survey series revealed that they are stationary, and strongly autoregressive. Thus, for multivariate modeling it is desirable to apply a filter to the data, in this case an AR(1) filter because all series share this pattern. The filter parameter was set at 0.75. The effect of prefiltering on collinearity in the models is dramatic; for example, the tolerance level of some explanatory variables increased from .05 to .80, which benefits the reliability of the OLS estimates.

The specification of lag orders in the models was done as follows: (1) the production judgment series were whitened by an AR(3) process in conjunction with a seasonal AR(1) of length 12^a. Although this model may not be the most parsimonious in all cases, it is the most parsimonious common

^aIt is interesting to observe that a mild seasonal pattern in the production judgment data exists, even though the survey explicitly requests seasonally adjusted opinions.

univariate model for the five countries and three sectors in the sample;

(2) The three most recent lags of expectations were included on a priori grounds: the survey expectations are reported for the next quarter, thus up to three lagged expectations could be related to current production^b;

(3) Likewise, one lag for orders was included, because that question invites a month-by-month comparison of ordering activity. Adding higher order lags did not improve the results.

In conclusion, the models were estimated as follows:

Univariate:

$$FP_t = b_0 + (b_1L + b_2L^2 + b_3L^3 + b_{12}L^{12})FP_t + u_t, \quad (5.1)$$

Multivariate:

$$FP_t = \beta_0 + (\beta_1L + \beta_2L^2 + \beta_3L^3 + \beta_{12}L^{12})FP_t + (\gamma_1L + \gamma_2L^2 + \gamma_3L^3)FE_t + \delta_1L F0_t + v_t, \quad (5.2)$$

where $FP_t = (1 - .75 L)P_t$ (production central tendency)

$FE_t = (1 - .75 L)E_t$ (expectations central tendency)

$F0_t = (1 - .75 L)O_t$ (orders central tendency)

The in-sample results (pre-1982) are reported in Table 4, and the out-of-sample results (1982-83) in Table 5. The findings strongly support the hypothesis that expectations Granger cause production judgment: (1) Wald's Chi-square test is significant in 13 of 15 cases at $p < 0.05$ for the in-sample analysis, (2) the expectations improve the forecasting error in 14 out of 15 predictions for 1982-1983. The range of forecasting improvement is 5 to 43 percent, averaging around 20 percent.

^bAlternatively, Dramais and Waelbroeck (1984) replace the production series by three-month moving averages. We found that approach to be problematic when time-series methods are used.

In conclusion, there is strong evidence of internal consistency of the survey data: manufacturers' current judgments of orders and current production expectations are useful in predicting next period's production judgments. Also, the role of the expectations is stronger than that of the orders: the expectations parameters and their t-ratios are generally higher (both series have about equal means and variances).

5.b. Forecasting production accounts

If manufacturers' assessments of production activities are accurate, and if the survey sample is adequate, the results of the analysis so far should apply to objectively measured production data. However, it is not obvious that such will be the case, because there may be systematic bias in the surveys over time. For example, a manufacturer may report increased production because he recalls reporting an optimistic expectation last month and does not want to appear inconsistent. Thus, a replication of the models on a different data source is at order.

National accounts production data were obtained in the form of time series of indices for the same sectors and countries, where December 1975 is scaled as a base index 100. Although seasonally adjusted data were also available, the analysis is done on raw numbers. It is preferable to analyze the time-series properties of these raw data and, if necessary, to use an appropriate transformation, for example a stochastic filter. The results of multivariate modeling may be substantially different between original and seasonally adjusted data (e.g. Feige and Pearce 1979).

The production indices exhibit strong seasonal behavior, as expected, and also a trend, so that regular and seasonal differencing was necessary to achieve stationarity. The transformed series were correlated with the production judgment data in order to obtain a first, crude idea of the correspondence between the two. As illustrated in Table 3, 14 out of 15

correlations are not significantly different from zero, and the only significant result has a negative sign. Thus there is overwhelming evidence that production judgments and production accounts are widely different measures of economic activity. Therefore the additional test on the value of production expectations is not trivial.

The design of the Granger test on production accounts data is similar to the one used in the survey analysis. The production indices (after differencing) are whitened by an AR(3) process plus a seasonal AR(1) component of length 12. Thus the models are:

-univariate:

$$I_t = b_0 + (b_1L + b_2L^2 + b_3L^3 + b_{12}L^{12})I_t + u_t \quad (5.3)$$

- multivariate:

$$I_t = \beta_0 + (\beta_1L + \beta_2L^2 + \beta_3L^3 + \beta_{12}L^{12})I_t + (\gamma_1L + \gamma_2L^2 + \gamma_3L^3)FE_t + \delta_1L FO_t + v_t \quad (5.4)$$

$$\text{where } I_t = \text{INDEX}_t - \text{INDEX}_{t-1} - \text{INDEX}_{t-12} + \text{INDEX}_{t-13}$$

The results are summarized in the bottom sections of Tables 4 (in-sample) and 5 (forecasting). For the in-sample analysis, although the residual variances are generally lower when expectations are included, Wald's Chi-square test is not significant in 14 out of 15 cases. The forecasting test confirms this pattern, showing only 4 out of 15 cases where a forecasting improvement occurred. The conclusion, then, is that Granger causality of survey expectations on production accounts cannot be established.

5.c. The formation of survey expectations

Within the information set containing survey and accounts data it is possible to perform a limited analysis of factors influencing production

expectations. The task is different from some published empirical investigations of rationality, because quantitative expectations that are directly comparable to the actual outcomes are not available. For example, Friedman (1980) was able to compare actual interest rates to point forecasts, which are measures of interest rate expectations.

Since the results on the predictive value of survey expectations are mixed, two questions arise: (1) do the expectations make an efficient use of historical information?, and (2) are they updated as new information becomes available? A positive answer to both would be a necessary, but not sufficient condition for rationality.

Methodologically, both questions can be cast in a Granger causality framework. First, consider the univariate model

$$E_t = b_0 + \sum_{k=1}^{\infty} b_k E_{t-k} + u_t \quad (5.5)$$

$$E_t = \beta_0 + \sum_{k=1}^{\infty} \beta_k E_{t-k} + \sum_{l=1}^{\infty} \gamma_l I_{t-l} + v_t \quad (5.6)$$

If model (5.6) outperforms (5.5), we would have evidence of insufficient use of historical information. Indeed, if the expectations are rational, then they incorporate all previous production information, so production accounts data should not Granger cause expectations.

Second, consider the multivariate model

$$E_t = \gamma_0 + \sum_{k=1}^{\infty} \gamma_k E_{t-k} + \delta I_t + w_t \quad (5.7)$$

If the manufacturers effectively incorporate new (i.e. current) production information in their forecasts, then δ should be significant in the equation. An extension of this idea is to consider three forms of new information:

- the current production index I_t ,
- the current production judgment P_t ,
- the current order judgment O_t .

Strictly speaking, I_t is not immediately available in numerical form, but one can assume that it is felt immediately by those closest involved with industrial activity, the manufacturers. Since I_t and P_t are uncorrelated, the question arises as to whether the manufacturers react to an objective measure of production or to their perception of actual production or both.

In order to avoid specification bias, the hypotheses of efficient use of historical and new information are tested jointly, i.e. using the model^a

$$E_t = \beta_0 + \sum_{k=1}^{\infty} \beta_k E_{t-k} + \sum_{l=1}^{\infty} \gamma_l I_{t-l} + \delta I_t + \psi P_t + \lambda O_t - u_t \quad (5.8)$$

The maximum lag orders for E_{t-k} and I_{t-l} were set at 3, which is consistent with previous analyses. Experiments with higher-order lags revealed no additional effects.

The results are summarized in Table 6, which contains t-values of the regression parameters, for ease of exposure. Generally they indicate that historical information is used efficiently, and that new information updating occurs:

- (1) only 7 out of 45 lagged production accounts parameters are significant.

In other words, previous production levels are generally incorporated in previous expectations levels. It is interesting to observe that 5 out of 7 exceptions occurred in the intermediate sector.

^aThe single-equation approach does not suffer from simultaneous-equation bias (between production and expectations), because the system is recursive and there is no autocorrelation in the residuals.

- (2) the updating variables (current production accounts, current orders judgment and current production judgment) contribute significantly to current expectations levels. Fourteen out of 15 regressions have at least one significant updating effect. Interestingly, current order judgments are generally the most strongly associated with expectations. Also, it appears that the survey data are used more strongly than the production accounts data.

6. Conclusions and Discussion

This paper has presented an empirical evaluation of the value of survey reported production expectations data. At the methodological level, we have compared various methods for scaling survey expectations. Also, we have stressed the importance of examining the time-series properties of the data in building evaluation models. Specifically, we recommend that tests of the forecasting value of expectations be based on the concept of Granger causality.

At the empirical level, our recommendations were used on EEC production expectations and order judgments for three sectors in five member countries. The main findings are (1) survey expectations and order judgments Granger cause production judgments, (2) survey expectations do not Granger cause production accounts, and (3) survey expectations make an efficient use of historical and new production and orders information. The findings are generally consistent across countries and economic sectors; one exception is that the intermediate sector is possibly less efficient in using historical information.

The apparent conflict in our results on the forecasting value of survey expectations is intriguing and important. It is not possible to state a priori that either production judgments or production accounts are better measures of

the true economic activity in a sector. And yet their intercorrelation is essentially zero, so one of them is an invalid measure or they measure different aspects of economic activity. We leave the answer open to the reader, but its implications are clear: if the economy is measured by manufacturers' opinions, then the expectations data make a valuable contribution in forecasting. If national accounting data are used, then the expectations are essentially useless.

Although this study is the first to use a Granger causality test on multiple measures of the economic variable to be forecast, it leaves a few areas of needed research unanswered. First, the use of a multi-country, multi-sector design may enhance one's statistical confidence in the results, but the choice of countries and sectors was influenced by data availability. An investigation of the poolability of the data is needed. Second, a validation of our findings on consumer sentiment data would also be useful. Finally, we have not explored some more sophisticated modeling techniques such as varying coefficients models which recognize that the predictive value of expectations may be contingent.

Table 1. Overview of research questions concerning survey based economic attitudes

Dependent variable	current judgment	current expectation*	current accounts
current judgment		1,F,C Does current expectation depend on current judgment?	2,C
current expectation	1,C Do current judgment and expectation measure a common underlying factor?		not applicable
current accounts	2,C How accurately do judgment and current realizations correspond?	7,F Do current expectations depend on current accounts?	
past judgment	3,F Is current judgment explained by its own past?	8,F Do current expectations depend on past judgment?	10,P How well are current accounts predicted by past judgment?
past expectation	4,P How well do past expectations predict current judgment?	9,F How well are current expectations explained by their own past?	11,P Are current accounts predicted by their past expectations?
past accounts	5,F Do current judgments depend on the history of past accounts	7,F Do current expectations depend on past accounts?	12,P,F Do present accounts depend on their own past?

*Current expectation denotes the expectation of a future event, formulated today.

Table 2. Correlation matrix of expectations series

<u>Country</u>	<u>Sector</u>	<u>Correlation</u>			<u>Mean percentage</u>		
		<u>UP-EQ</u>	<u>UP-DO</u>	<u>EQ-DO</u>	<u>UP</u>	<u>EQ</u>	<u>DO</u>
Belgium	CON	-.25	-.76	-.43	.19	.57	.23
	INV	.16	-.74	-.78	.15	.58	.26
	INT	.15	-.76	-.76	.15	.60	.25
France	CON	-.62	-.76	-.05	.28	.60	.12
	INV	-.27	-.77	-.41	.25	.59	.16
	INT	-.17	-.78	-.48	.21	.64	.15
Germany	CON	-.48	-.63	-.38	.14	.74	.12
	INV	-.12	-.75	-.57	.13	.75	.12
	INT	-.14	-.65	-.66	.13	.74	.13
Holland	CON	-.61	-.35	-.53	.14	.77	.10
	INV	-.56	-.28	-.63	.10	.78	.12
	INT	-.58	-.42	-.50	.17	.71	.12
Italy	CON	-.38	-.62	-.48	.20	.62	.17
	INV	-.53	-.69	-.25	.18	.66	.16
	INT	-.29	-.65	-.54	.18	.67	.16

Table 3. Correlations between production judgment and accounts data*

	BELGIUM (N=131)	FRANCE (N=164)	GERMANY (N=143)	HOLLAND (N=109)	ITALY (N=115)
CONSUMER SECTOR	.10	.03	.09	-.04	-.04
INVESTMENT SECTOR	-.04	-.03	-.01	.04	-.14
INTERMEDIATE SECTOR	.08	-.04	.10	.09	-.21 ^a

*the accounts data are differenced (regular and seasonal) to achieve stationarity

^asignificant at $p < .05$

Table 4. Production models with and without expectations: 1969-1981

		<u>BELGIUM</u>		<u>FRANCE</u>		<u>GERMANY</u>		<u>HOLLAND</u>		<u>ITALY</u>	
Survey Reported Production	Consumer	.148 ^a	13.62 ^c	.232	5.37	.500	11.20	.194	12.37	.224	18.12
		.411 ^b	9.17 ^{d*}	.456	3.91*	.617	7.43*	.246	12.12	.296	15.16*
	Investment	.156	8.76	.338	3.45	.340	5.79	.238	24.96	.147	16.67
		.339	7.04*	.508	2.64*	.544	4.10*	.289	24.37	.361	12.83*
	Intermediate	.174	16.68	.380	5.63	.379	14.80	.172	21.71	.128	14.53
		.402	12.69*	.622	3.53*	.618	9.36*	.382	17.02*	.338	11.33*
National Production Accounts	Consumer	.374	26.70	.475	34.68	.355	22.52	.411	30.80	.154	80.58
		.421	25.34	.417	35.49	.389	22.04	.425	35.04	.203	79.39
	Investment	.433	13.01	.379	19.12	.429	14.95	.393	8.72	.292	39.12
		.501	11.86*	.407	18.77	.455	14.73	.435	8.62	.317	39.52
	Intermediate	.313	24.63	.349	10.54	.189	7.62	.305	15.44	.115	20.56
		.372	24.29	.372	10.45	.234	7.43	.365	14.63	.209	19.24

^a R^2 for the univariate model (on filtered or differenced series)

^b R^2 for the model including expectations

^c Residual variance for the univariate model (X 1000 for survey reported production)

^d R^2 for the model including expectations

* Denotes significant variance reduction using Wald's Chi-square test, at $p < .05$

Table 5. Mean squared forecast errors: 1982-1983

		<u>BELGIUM</u>	<u>FRANCE</u>	<u>GERMANY</u>	<u>HOLLAND</u>	<u>ITALY</u>
Survey Reported Production (X 1000)	Consumer	14.38 ^a	7.02	4.66	3.72	8.58
		11.36 ^b	5.52	5.32	3.52	7.26
	Investment	5.22	4.18	10.14	4.69	12.06
		4.27	3.20	8.10	3.94	9.85
	Intermediate	11.82	14.38	12.45	12.72	8.57
		9.60	10.20	7.05	8.31	6.97
National Production Accounts	Consumer	20.91	48.92	15.03	18.09	46.98
		22.48	50.54	14.87	18.39	47.22
	Investment	8.30	10.72	12.44	11.89	32.18
		11.20	11.62	12.99	13.44	34.20
	Intermediate	25.73	3.57	7.64	21.00	11.74
		21.06	3.66	6.96	19.69	16.16

^a MSFE for the base model (univariate), retransformed to original levels

^b MSFE for the model including expectations

Table 6. Expectation formation models*

Explanatory variable	Time index	BELGIUM			FRANCE			GERMANY			HOLLAND			ITALY		
		CON	INV	INT	CON	INV	INT	CON	INV	INT	CON	INV	INT	CON	INV	INT
Expectations	T-1	.99	1.46	.73	1.06	.16	3.35 ^a	-1.03	4.02 ^a	3.71 ^a	-2.25 ^b	-2.89 ^a	-3.65 ^a	1.01	-1.24	-.24
	T-2	.57	2.87 ^a	1.09	-.51	.18	.48	-2.68 ^a	1.44	-.28	-1.33	-2.84 ^a	-2.27 ^b	-.16	1.33	.09
	T-3	-.47	-1.63	-.04	-.16	-.08	-.17	-5.66 ^a	-3.08 ^a	-3.05 ^a	2.81 ^a	1.10	-1.13	-.59	-.92	.09
Production accounts	T	-.23	-.32	2.37 ^b	-2.49 ^b	.83	2.95 ^a	.48	-.25	1.93 ^c	.57	.25	1.37	.78	-.11	2.29 ^b
Orders judgment	T	2.79 ^a	3.89 ^a	2.96 ^a	4.13 ^a	4.01 ^a	4.91 ^a	6.31 ^a	2.41 ^b	2.51 ^b	2.24 ^b	.77	1.92 ^c	1.40	2.26 ^b	2.92 ^a
Production judgment	T	.89	-1.13	1.38	.69	1.96 ^b	-1.21	-2.80 ^a	-1.15	-2.65 ^a	.02	2.21 ^b	2.15 ^b	-.36	-4.12 ^a	-1.27
Production accounts	T-1	.56	.62	1.31	-1.12	1.16	2.11 ^b	1.43	.72	2.83 ^a	.52	.52	1.97 ^b	-.01	-.56	.24
	T-2	1.21	1.63	1.44	-.51	.61	1.10	1.00	1.08	2.30 ^b	.79	1.12	1.00	1.05	.85	1.39
	T-3	-1.94 ^c	1.06	1.95 ^c	-.62	-.94	.87	-.39	-.59	1.44	1.75 ^c	-.19	1.16	-.72	-.21	.59
R ²		.23	.28	.39	.27	.34	.40	.37	.26	.29	.21	.23	.19	.09	.22	.19
N		121	121	121	160	160	160	140	140	140	103	103	103	106	106	106
F		3.7 ^a	4.8 ^a	7.8 ^a	6.2 ^a	8.6 ^a	11.3 ^a	8.46 ^a	5.12 ^a	5.98 ^a	2.80 ^a	3.12 ^a	2.35 ^b	1.04	3.09 ^a	2.45 ^b

*The regression parameters are represented by their t-values and overall significance levels a(p < .01), b(p < .05) and c(p < .1). The survey data are filtered by regular and seasonal differencing.

Bibliography

- Adams, F.O., and V.O. Duggal (1974), "Anticipation Variables in an Econometric Model: Performance of the Anticipation Version of Wharton Mark III," International Economic Review 15, June, 267-84.
- Adams, G.F. and L.R. Klein (1972), "Anticipations Variables in Macro-Econometric Models," B. Strumbel, J.N. Morgan and E. Zahn (eds.), Human Behavior in Economic Affairs, Amsterdam: Elsevier/
- Ahlers, D. and J. Lakonishok (1983), "Economists' Consensus Forecasts," Management Science, 25, 10, 1113-25.
- Bechtel, G.G., (1981), "Measuring Subjective Social Indicators," Journal of Economic Psychology, 1, 165-81.
- Belsley, D., Kuh, E. and Welsch, R. (1980), Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: Wiley.
- Brown, B.W. and S. Maital (1981), "What Do Economists Know? An Empirical Study of Experts' Expectations," Econometrica, 49, 2 (March) 491-504.
- Carlson, J.A. and M. Parkins (1975), "Inflation Expectations," Economica, 42 (May), 123-38.
- Ciret (1982), International Table on Trade Cycle Surveys, Munchen, IFO, October.
- Commission of the European Communities, "Results of the Business Survey Carried Out Among Managements in the Community," Brussels, various issues.
- Dunkelberg, W.C. (1972), "The Impact of Consumer Attitudes on Behavior: A Cross-Section Study," in B. Strumbel, J.N. Morgan and E. Zahn (eds.), Human Behavior in Economic Affairs, Amsterdam: Elsevier.
- Dramais, A. and E. Waelbroeck-Rocha (1984), "A Probabilistic Approach to the Quantification of Business Survey Data," Journal of the American Statistical Association, forthcoming.

- Feige, E.L. and D.K. Pearce (1979), "The Casual Causal Relationship between Money and Income: Some Caveats for Time Series Analysis," The Review of Economics and Statistics, 61, 4 (November), 521-33.
- Friedman, B.M. (1980), "Survey Evidence on the 'Rationality' of Interest Rate Expectations," Journal of Monetary Economics, 6, 453-65.
- Friend, I. and W. Thomas (1970), "A Reevaluation of the Predictive Ability of Plant and Equipment Anticipations," Journal of the American Statistical Association, 65, 330, 510-519.
- Geweke, J., R. Meese and W. Dent (1983), "Comparing Alternative Tests of Causality in Temporal Systems," Journal of Econometrics, 21, 161-194.
- Hymans, S.H. (1970), "Consumer Durable Spending: Expectation and Prediction," Brookings Papers on Economic Activity.
- Jorgensen, S.W., T. Hunter and M. Vadici (1970), "The Predictive Performance of Econometric Models of Quarterly Investment Behavior," Econometrica, 38, 213-24.
- Joyeux, R. (1980), "Relationships between Economic Time-Series and their Anticipations," Quarterly Review of Economics and Business, 20, 1 (Spring), 7-15.
- Juster, F.T. and P. Wachtel (1972), "Inflation and the Consumer," Brookings Papers on Economic Activity, 1, 71-114.
- Katona, G. (1975), Psychological Economics. New York: American Elsevier.
- Liebling, H.I. and T.M. Russell (1969), "Forecasting Business Investment by Anticipation Surveys and Econometric Models in 1962-1969," Proceedings of the Business and Economic Statistics Section, American Statistical Association, 250-55.
- Liebling, H.I., P.T. Bidwell and K.E. Hall (1975), "The Recent Performance of Anticipation Surveys and Econometric Model Projections of Investment Spending in the United States," Journal of Business, 49, 451-77.

- McNeil, T. (1974), "Federal Programs to Measure Consumer Purchase Expectations, 1946-1973: A Post-Mortem," Journal of Consumer Research, 1 (December), 7-10.
- Mishkin, F.S. (1981), "Are Market Forecasts Rational," American Economic Review, 71, 3 (June), 195-36.
- Modigliani, F. and H.M. Weingartner (1958), "Forecasting Uses of Anticipatory Data on Investment and Sales," Quarterly Journal of Economics, 72, (February), 23-54.
- Mullineaux, D.J. (1978), "On Testing for Rationality: Another Look at the Livingston Price Expectations Data," Journal of Political Economy, 86, 2, 329-336.
- Nerlove, M. (1983), "Expectations, Plans and Realizations in Theory and Practice," Econometrica, 51, 1251-79.
- Papadia, F. and V. Basano (1981), "EEC-DG II Inflationary Expectations: Survey Based Inflationary Expectations for the EEC Countries," Commission of the European Communities Economic Papers, 1, May, Brussels.
- Praet, P. and J. Vuchelen (1983), "The Contribution of E.C. Consumer Surveys in Forecasting Consumers Expenditures: An Econometric Analysis for Four major Countries," Free University of Brussels, Institute of European Studies, Mimeograph.
- Praet, P. (1984), "The Integration of EEC Qualitative Consumer Survey Results in Econometric Modelling: An Application to the Consumption Function," Commission of the European Communities, Economic Papers 30, February, Brussels.
- Praet, P. (1984), "Endogenizing Consumers' Expectations in Four Major EEC Countries," Free University of Brussels, Institute of European Studies, Mimeograph.

- Rippe, R.D. and Wilkinson (1974), "Forecasting Accuracy of the McGraw-Hill Anticipatory Data," Journal of the American Statistical Association, 69 (December), 849-858.
- Rippe, R., M. Wilkinson and D. Morrison (1976), "Industrial Market Forecasting with Anticipations Data," Management Science, 22, 6 (February), 639-50.
- Shapiro, H.T. (1972), "The Index of Consumer Sentiment and Economic Forecasting: A Reappraisal," in B. Strumbel, J.N. Morgan and E. Zahn (eds.), Human Behavior in Economic Affairs, Amsterdam: Elsevier.
- Vanden Abeele, P. (1983), "The Index of Consumer Sentiment: Predictability and Predictive Power in the EEC," Journal of Economic Psychology, 3, 1-17.
- Wilks, S.S. (1962), Mathematical Statistics, New York, Wiley.
- Zarnowitz, V. (1979), "An Analysis of Annual and Multiperiod Quarterly Forecasts of Aggregate Income, Output and Price Level," Journal of Business, 52 (January), 1-33.
- Zimmermann, K.F. (1984), "On Rationality of Business Expectations: A Micro Analysis of Qualitative Responses," Discussion Paper No. 270-84, Institut für Volkswirtschaftslehre und Statistik, Universität Mannheim.