Opinion Formation in Business Surveys: Empirical Evidence from German Micro Data

Klaus Wohlrabe
Ifo Institute for Economic Research
Poschingerstrae 5
81679 Munich
wohlrabe@ifo.de

July 17, 2006

Abstract

The article contributes to a broader understanding of how firms form their opinion in business surveys. The rational expectation hypothesis states that every available information is used to form its expectations. We specify the empirical content of the public and private information set. We use a large micro data set from the business surveys of the ifo Institute in Germany. We estimate panel order probit and log-linear probability models. We find for the private information set that production expectations play the most important role for the expectations and the appraisal of the backlog of orders for the assessment of the current business situation. The results for the public information set are not clear-cut due to a possible omitted variable bias. Furthermore we find some preliminary empirical evidence for persistence in opinion formation and reverse causal effects for expectations and appraisals.

JEL: L29, C25, C42, C81

Draft version, do not quote, comments welcome

1 Introduction

Expectations play a central role both in theoretical as well as in empirical economics. The rational expectations hypothesis (REH) proposed by Muth (1969) states in original form that every public information available at time t is used to form its expectation. This idea was later extended that the information set Ω_i is given by

$$\Omega_{it} = \Psi_t \cup \Phi_{it}, \tag{1}$$

where Ψ_t denotes the public information set and Φ_{it} the private information set for an individual i at time t. Besides the REH there are several other expectation formation theories, i.e. the general class of extrapolative expectations or the simple static formation processes, Pesaran and Weale (2006) gives an overview. They all have in common that they rely on an specific information set. As the theory is very clear and general empirically it is unclear which concrete information is contained in the information set. There has been many test about the REH[1] but most of them rejected the REH. This has lead to more elaborate theories of expectation formation, e.g. Bayesian learning models and heterogeneous expectations.[2] The subscript i indicates that the information set may differ across individuals.

This paper wants to investigate whether it is possible to empirically investigate which information are contained in the information set Ω_{it} and if so which are the specific information. In order to answer these question we use a large micro data set provided by ifo Institute for Economic Research in Germany. The ifo business cycle test is one of the well known indicators for the German economy. The business cycle test consists of two main questions, *Assessment of the current business situation* and the *expectations for next six month*. Until now the time series and forecasting properties has been investigated in the empirical literature.[3] Recently the micro data has been provided for the scientific public.[4] To answer the research questions we have to ask which information is available at t. Concerning the private information set we have for a specific firm in the ifo business cycle test only

[1]See for instance James McIntosh and Low (1989), Kukuk (1994), Bonham and Cohen (2001), and G. Elliot and Timmerman (2005).

[2]See again Pesaran and Weale (2006) for a recent survey.

[3]See Hfner and Schrder (2002), Bandholz and Funke (2003) and Kholodilin and Siliverstovs (2006) among others.

[4]See Klaus Abberger and Wohlrabe (2006) for details.

the monthly information given in the questionnaire and furthermore some structural information on the enterprise. The choice of available public information is arbitrary as we do not rely on a specific theory where we conduct the testable model.

Heinz Koenig and Oudiz (1981) and Nerlove (1983) also used ifo micro data to investigate the "black box" of a firm (Nerlove). They presented simple models of expectation formation and planning of firms who report over time on both expectations and their subsequent realizations. They found that the error-learning model (a form of adaptive expectations) is the best and most parsimonious explanation of the ifo data of German manufacturing firms. Furthermore they state that price anticipations and production plans are independent of each other. In a way we extend this work as we focus on the two main questions and apply in addition to the log-probability model the ordered probit model.

We proceed as follows. First we present some expectation formation theories. Afterwards we describe the data for the private and public information set. In section 4 we present the econometric model which we use to uncover the specific information set. It is followed by the empirical results. Finally we conclude.

2 Expectation formation - Theory

First we focus on expectation formation theories and extend this to the assessment formation theory. We follow Pesaran and Weale (2006) we decompose the individual specific information set of an individual i and time t, Ω_{it} into a public information set Ψ_t, and an individual-specific private information set Φ_{it} such that

$$\Omega_{it} = \Psi_t \cup \Phi_{it} \tag{2}$$

for $i = 1, 2, \ldots, N$. Under the Muthian notion of the REH, private information plays no role in the expectation formation process, and expectations are fully efficient with respect to the public information, Ψ_t. In the case of point expectations, the optimality of the RHE is captured by the "orthogonality" condition

$$\mathrm{E}(\xi_{t+1}|S_t) = 0 \tag{3}$$

where ξ_{t+1} is the error of expectations defined by

$$\xi_{t+1} = x_{t+1} - \mathrm{E}(x_{t+1}|\Psi_t) \tag{4}$$

and $S_t \subseteq \Psi_t$, is a subset of Ψ_t. The orthogonality condition (3) in turn implies that, under the REH, expectation errors have zero means and are serially uncorrelated.. It does not require the expectation errors to be conditionally or unconditionally homeskedastic. There has been a lot of research to establish heterogeneous rational expectations models, see Pesaran and Weale (2006) for a survey. When these models have unique solution, expectation errors of individual agents continue to satisfy the usual orthogonality conditions.
In addition to the REH, a wide variety of expectations formation has been advanced in the literature with differing degrees of informational requirements. Most of these models fall under the "extrapolative" category, where point expectations are determined by weighted averages of past realizations. A general extrapolative formula is given by

$$\mathrm{E}_i(x_{t+1}|\Omega_{it}) =_{s=0}^{\infty} \Phi_{is} x_{t-s} \tag{5}$$

where the coefficient matrices, Φ_{is}, are assumed to be absolute summable subject to the adding up condition

$$_{s=0}^{\infty}\Phi_{is} = I_k. \tag{6}$$

Finally we want to state the simplest extrapolative model, the static expectation model. In its basic form it is given by

$$\mathrm{E}_i(x_{t+1}|\Omega_{it}) = \bar{\mathrm{E}}(x_{t+1}|S_t) = x_t \tag{7}$$

and is optimal (in the mean squared error sense) if x_t follows a pure random walk model.

3 The data

3.1 The private information set

We use one of the oldest and most famous business survey is that done by the ifo Institute, Munich in Germany, every month since November 1949 for Germany. The empirical results reported in this paper are based on monthly data from January 1991 - December 2000. The monthly ifo business cycle test contains about 6800 establishment, but has to be corrected for several reasons. First we concentrate on the industry level as the most comprehensive aggregate of the economy and as the representation in this category is the largest. Furthermore we exclude firms which do not export[5], as the German industry is export oriented. All the data we use are trichotomous. The categorical data can be classified into three groups

1. variables that reflect plans or expectations (ex ante data);
2. variables referring to realizations (ex post data);
3. variables indicating evaluation evaluations or appraisals of inventories, order backlogs and the like.

Response are in the form: Increase (+), normal (=), or decrease (-); or greater than normal (+), normal (=), or less than normal (-); too large (+), about right(=), or too small (-). The data available is summarized in Table 1.

Table 1: Ifo Business-Test Variables

Variable	Plans or Expectations	Realizations	Appraisals
Business Conditions	G^*	G	-
Production	Q^*	Q	-
Inventories of Finished Products	-	L	L^a
Backlog of Orders	-	A	A^a
Domestic Selling prices	P^*	P	-
Demand	-	D	-

The ifo Business cycle test is the geometric mean of the assessment of the current business situation (G) and the expectations for next six months (G^*).[6] To give an idea of the answering behavior consider the different answering combinations of the two variables displayed in Table 2.

For the time period 1991 to 2000 we calculated how long does a firm stays in one of the 9 given states plus the persistence in each category of G and G^*. The results are given in Table 3. The largest persistence we observe when a firm both makes its cross in the middle category (2,2). On average a firm stays 5 month in this state. This goes in line when we focus on each variable separately. Again the middle category exhibits the largest persistence. We will incorporate this result in our estimations.

Table 2: Answering Combinations between Assessment and Expectations

		G^*		
		"more favorable"	"unchanged"	"less favorable"
	"good"	(1,1)	(1,2)	(1,3)
G	"satisfactorily"	(2,1)	(2,2)	(2,3)
	"poor"	(3,1)	(3,2)	(3,3)

Table 3: Persistence in the ifo Questionnaire 1991-2000

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
Mean	3.27	3.86	2.79	3.23	4.88	3.11	3.40	4.07	3.76
Sdt.	0.93	0.99	1.42	0.59	1.31	0.65	0.84	0.96	1.08
Max	6.75	7.55	7.00	6.06	12.49	5.18	7.00	8.69	7.59
	Assessment (G)			Expectations (G^*)					
	1	2	3	1	2	3			
Mean	4.84	6.49	6.39	3.99	6.10	3.95			
Std.	1.71	1.98	2.01	0.78	1.72	0.85			
Max	12.57	16.76	12.76	5.83	14.91	6.02			

3.2 The public information set

The choice of the used data is ad hoc and driven by plausible assumptions. We use the following data obtained from the EcoWIN database

- OIL Price
- Real Effective Exchange Rate
- DAX
- Production Index
- Domestic Turnover Index
- Export Turnover Index

[5]This information is available from the questionnaire.
[6]For more information on the calculation of the business climate see www.ifo.de.

- Producer Prices Index
- Ifo Incoming Orders
- lagged Ifo Business Climate

The OIL price is a proxy for the material prices in general. The real effective exchange rate represents the comparative advantage of exporting firms. We use the DAX to incorporate a financial indicator of the current situation of the economy as a whole. The indices are used to represent aggregated measures for the industry branch and may affect the opinion formation of a specific firm. The lagged Ifo business climate may be regarded as a general sentiment indicator of the economy.
Besides the Ifo indicators are these time series exhibits trends and therefore cannot be used in ordinal probit regressions. We use two different transformations. First, we take the first log differences, which can be interpreted as short-run influences on the opinion formation process. Long-run influences may be covered by the yearly log differences.

4 The econometric model

As already mentioned we do not employ a specific theoretical model from which we can deduce an econometric model. As the data is ordered trichotomous the ordinal probit model is the natural thing to use. The ordinal probit model can be used both for the private and the public information set. Furthermore in line of Heinz Koenig and Oudiz (1981) and Nerlove (1983) we use the log-probability model to investigate relationships between the ordinal variables. As will be shown in the next section that the regressions may be biased due to endogeneity we shortly describe a test for exogeneity in panels due to Peter Adams and Ribeiro (2003).

4.1 The Ordinal Probit Model

We assume the following general relationship

$$y_{it}^* = \beta' x_{it} + \gamma y_{it-1} + c_i + \epsilon_{it} \quad (i = 1, \ldots, N; t = 2, \ldots, T) \qquad (8)$$

where y_{it}^* represents either the expectation or the situation variable. The vector x contains observed variables from the questionnaire or exogenous macro variables which may be associated with the y variable. To capture state dependence , y_{it-1} is a vector of indicators in the previous month. The disturbance term ϵ_{it} is a time and firm-specific error term which is assumed to be normally distributed and uncorrelated across firms and months and uncorrelated with c_i. Furthermore ϵ_{it} is assumed to be strictly exogenous. As we dot not have a natural scale for the latent variable the variance of the idiosyncratic error term is restricted to equal one.

In our data the latent outcome y_{it}^* is not observed. Instead, we observe an indicator of the which the latent indicator falls (y_{it}). The observations mechanism can be expressed as:

$$y_{it} = j \quad \text{if} \quad \mu_{j-1} < y_{it}^* \leq \mu_j, \quad j = 1, \ldots, m \tag{9}$$

where $\mu_0 = -\infty, \mu_j \leq \mu_{j+1}, \mu_m = \infty$. Given the assumption that the error term is normally distributed, the probability of observing the particular category of expectations or assessment reported by firm i at time t, conditional on the regressors and the individual effect, is:

$$P_{itj} = P(y_{it} = j) = \Phi(\mu_j - \beta' x_{it} - \gamma y_{it-1} - c_i) - \Phi(\mu_{j-1} - \beta' x_{it} - \gamma y_{it-1} - c_i) \tag{10}$$

where $\Phi(\cdot)$ is the standard normal distribution function. From equation (10) we see that it is impossible to separately identify an intercept and the linear index (β_0) and the cut points (μ), the model only identifies ($\mu_j - \beta_0$). We solve this problem by setting $\beta_0 = 0$ (an alternative would be to set $\mu_1 = 0$). By extension, it is clear, that, without a priori restrictions, the individual effect (c_i) cannot be distinguished from an individual-specific cut point shift. The same argument applies to the impact of the regressors on y_{it}^* are a linear function of regressors. This should be borne in mind when interpreting the results presented below.

To implement the random effects estimator the individual effect can be integrated out, using the assumption that its density $\mathrm{N}(0, \sigma_c^2)$, to give the sample log-likelihood function

$$\ln L = \sum_{i=1}^{n} \left\{ \ln \int_{-\infty}^{+\infty} \prod_{t=1}^{T} (P_{itj}) \left[1/(\sqrt{2\pi}\sigma_c^2) \exp(-c^2/2\sigma_c^2) \right] dc \right\} \tag{11}$$

The likelihood contains a univariate integral which can be approximated by Gauss-Hermite quadrature. We allow for the possibility that the observed regressors may be correlated with the individual effect. As we deal with a non-linear dynamic model we have to take account of the problem of initial conditions as described by **?**. First, the initial observations have to be exogenous variables. This invalid when the error process is not serially independent and the first observation is not the true initial outcome of the process. The second assumption is that the process is an equilibrium such that the marginal probabilities have approached their limiting values and can therefore be assumed time-invariant. This assumption is not fulfilled if non-stationary variables serve as regressors. Wooldridge (2005) suggests a simple solution to the first problem by modeling the distribution of the unobserved effect conditional on the initial value and any exogenous variable. So we have to parameterize the distribution of the individual effect:

$$c_i = c_0 + c_1' y_{i1} + \alpha_2' \bar{x}_i + u_i. \tag{12}$$

We estimate the pooled ordered probit and random effects ordered probit model using the STATA program reaprob.ado, written by Guillaume R. Frechette.[7]

4.2 The log-linear probability model

We present the basic idea of the log-probability model to make the paper self-contained. For further details see any textbook on multivariate statistics. Suppose we have two variables A and B and set up the following contingency table

	B		
A	N_{11}	N_{12}	$N_{1.}$
	N_{21}	N_{22}	$N_{2.}$
	$N_{.1}$	$N_{.2}$	N

Assuming that N_{ij} is distributed multinomial we can show via maximum likelihood

$$\hat{P}_{ij} = \frac{N_{ij}}{N}. \tag{13}$$

Assuming that A and B are independent we can write

$$P_{ij} = P_{i.} P_{.j} \tag{14}$$

[7]Stata Technical Bulletin 59, January 2001

so we have

$$\hat{P}_{ij} = \frac{N_{i.}}{N}\frac{N_{.j}}{N}. \tag{15}$$

We can rewrite this equation to

$$\log P_{ij} = \log N_{i.} + \log N_{.j} - 2\log N \tag{16}$$

This is similar to ANOVA, where $\log N$ is the average mean of all observations, thus we rewrite under the assumption of independence

$$\log P_{ij} = \alpha_1(i) + \alpha_2(j) + \mu \tag{17}$$

where $\alpha_1(i)$ and $\alpha_2(j)$ are the variables A and B respectively and μ is standardization parameter. If the variables are not independent we have to add an interaction parameter

$$\log P_{ij} = \alpha_1(i) + \alpha_2(j)\beta_{12}(i,j) + \mu \tag{18}$$

To identify the parameters the following restrictions must be fulfilled

$$\sum_i \alpha_1(i) = \sum_j \alpha_2(j) = 0 \quad \sum_i \beta_{12}(i) = \sum_j \beta_{12}(j) = 0 \tag{19}$$

We can generalize this idea to $A = A_1, \ldots A_q$ variables with values $i_1 = 1, \ldots, I_1, i_2 = 1, \ldots I_2, i_q = 1, \ldots, I_q$. Thus we obtain

$$\begin{aligned} \log P_{i_1,\ldots i_q} &= \mu + \alpha_1(i_1) + \ldots + \alpha_q(i_q) \\ &\quad +\beta_{12}(i_1, i_2) + \ldots + \beta_{q-1,q}(i_{q-1}, i_q) \\ &\quad + \ldots \\ &\quad +\omega_{1,\ldots,q}(i_1, \ldots, i_q) \end{aligned} \tag{20}$$

with the restrictions

$$\begin{gathered} \alpha_1(.) = \alpha_2(.) = \alpha_q(.) = 0 \\ \beta_{12}(i_1, .) = 0, \beta_{12}(., i_1) = 0, \ldots \beta_{(q-1)q}(., i_q) = 0 \\ \ldots \\ \omega_{1,\ldots,q}(i_1, \ldots i_{q-1}, .) = 0, \ldots \omega_{1,\ldots,q}(., i_2, \ldots i_q) = 0. \end{gathered}$$

There are several ways to estimate such a model. The first approach is to estimate a general saturated model (unconditional estimation). Then eliminate all insignificant variables and reestimate the model until all variables are significant. A second approach exploits specific dependency structures, conditional estimation. Due to the restrictions in the estimation some of the variables a collinear, i.e. redundant in the estimation procedure. We can their parameter value from the other ones but cannot state any significance level.

4.3 Test for exogeneity in questionnaires

to be completed

5 Empirical Results

5.1 The private information set

5.1.1 The ordinal probit model

The private information available at time t is given by the questions in the questionnaire. If we assume that the two main questions *Assessment* and *Expectations* are exogenenous we can employ the ordinal probit model given by equation (8). In other words we assume that a specific enterprise deduces the opinion formation for the two main questions from the other answers in the questionnaire. First, we estimate both a pooled model and the RE panel probit model without lagged dependent variables. The estimated coefficients are not directly comparable to those reported for the pooled models due to different scaling of the error variance. The pooled ordered probit assumes that the error term as a whole is distributed N(0,1) for identification of β. The random effects ordered probit restricts ϵ_{it} to be N(0,1), so that the overall error variance equals $(\sigma_u^2 + 1)$. This implies different scaling of the estimated coefficients in the two models. The results are tabulated in Table 4. All coefficients are significant. Despite different scaling the results between pooled and the RE model only differ slightly. Although after allowing for heterogeneity substantially improves the fit of the model as evidenced by the change in log-likelihood. For *Assessment* approximately 29% and for *Expectations* 21% of the latent error variance is attributable to unobserved heterogeneity, as measured by the intra-class correlation coefficient ρ. On the basis of this result we do not show the average partial effects as the regressors

are scaled identically.[8]
For the assessment of the current business situation the appraisal of the backlog of orders plays the most important role indicated by the largest coefficient of 1.18 and 1.26, respectively. There results show that both realization as well as expectations and appraisals play an important role in the opinion formation process. Similar results we obtain for the expectations. The production expectation play the most important role for the formation of expectations. The other variables are significant but have much less influence.
We rerun the regression including the lagged dependent variable as a regressor. As shown in the data section there is a high persistence in the answering behavior of the firms. Table 5 reports the results. The difference between the pooled and the RE estimation is again quite small. The value of the log-likelihood is again higher which points out to the fact that this model is more appropriate compared to the model without an autoregressive term. The lagged variables dominate all the other variables. The magnitude of the coefficients decreased in all cases. The appraisal of the backlog of orders has got the most influence on opinion formation. The price expectations are not significant anymore. For the expectation formation the state dependence is lower as in the assessment case. The production versus the previous month turned to be negative and significant. One possible story behind this could be that is a possible sign for future counteraction. The appraisal of backlog of orders does not play any role in the expectation formation.

5.2 The log-probability model

By applying the log-probability model we focus on the conditional model. Estimation of the model with all variables in the model is time consuming and cannot be reported appropriately in a paper. For instance if we would include all 11 variables we would have 10 bivariate effects and 900 possible trivariate effects and so on. We focus on the most influential variables from Table 1, the appraisal of the backlog of orders, the expected production and the lagged value of the corresponding variable in question. We estimate the model for each month average the results. Table 6 presents the results for the assessment of the current situation. The table has to be read as follows. On the horizontal axis are the categories "good", "satisfactorily" and "poor" for the assessment of the current situation. On the vertical line the values

[8]The average partial effects can be obtained from the authors upon request.

of the corresponding variable in the bivariate interaction. As there is no reference value we can only compare relative to each other. The value 5.855 for the bivariate interaction between G and A^a indicates a high probability of marking "good" for G and "comparatively large" for A^a. The effect is even stronger for "poor" and "too low", respectively with a value of -14.468. Furthermore it is unlikely that a firms marks "good" and "too low" (-9.797). The advantage of the log-probability model is its opportunity to identify relationships between specific categories of two variables in a questionnaire. Generally speaking the model approve the results obtained from the ordinal probit model. Again, the appraisal of the backlog of orders is the most important influence for the assessment of the current situation. A further tool to investigate the strength between two ordinal variables is the Goodman-Kruskal Gamma[9] (γ), which also used Nerlove (1983) to test expectation formation theories. Like standard correlation coefficients it is bounded between 1 and -1. The association between G and A^a is with 0.838 the highest and supports again the already stated results.
Table 7 shows the results for the expectations of firms. We do not go into details as the results are similar in interpretation as for the assessment. Again the production expectations play the most important role for the expectation formation process.
We end this section with the bivariate association between our two main variables and the other variables from the questionnaire. Table 8 presents the Goodman-Kruskal gamma over the whole sample period. The results endorse the previous findings from the orderet probit estimation presented in Table 6 and 7.

5.3 The public information set

In order to investigate the public information set we estimate model (8) both for monthly and yearly differences. We include time dummies to capture macroeconomic effects. Table 9 states the results for the macroeconomic time series in monthly differences. These can be interpreted as short run influences on the opinion formation process of firms. As the scaling of the variables is arbitrary one has to be careful in interpreting the magnitude of the coefficients.
We therefore calculated the marginal average effects for each category and

[9] See Goodman and Kruskal (1979).

variable. The results are presented in Table 10. The figures in bold face are significant at the 5 % level. The results are not straightforward and do not always show the expected sign. The largest influence for the category "good" for the assessment are the producer prices but we would expect a negative sign meaning higher prices would decrease the probability marking "good". Further variables with unexpected signs are the export turnover and the production index. The positive signs for the incoming orders, the DAX and the lagged Ifo climate indicate a kind of herding behavior. In the middle category almost no variable is significant which is plausible because it does not exhibit large variance. For the "poor" case we have just the opposite signs compared to the "good" case. Here we have to state the opposite interpretations.
Similar results apply for the expectations. The largest influence has still the producer prices but again with the "wrong" sign. Implausible signs we have for the production index and the ifo incoming orders. A rising oil price do play a role in the expectation formation with goes in line with the large press coverage on raw materials. A central role play also the domestic turnover and the export.
Table 11 shows the estimation results for ordinal probit model with yearly differences. Due to differences in scaling and timeliness we do not compare the coefficients among each other and to the monthly differences. We report in Table 12 the average partial effects. There are less significant variables compared to the short run effects. Again the producer prices have the largest effect. A lot of coefficients sill exhibit the unexpected sign, as the export for the assessment and the oil price and the incoming orders for the expectations. For the middle category no coefficient is significant.
If we summarize the results for the public information set we have to state that these are not clear-cut. We obtained in the average partial effects many coefficients with an unexpected sign. These results can be a result of an omitted variable bias. We cannot assure that we have included all possible public information. The problem could increase as the ordinal probit model is non-linear. We leave this for further research.

6 Conclusions

This paper contributes to a broader understanding of how opinions are formed in regular business surveys. The rational expectation hypothesis states that every firm uses all available information to form its expectations. The information set can be divided into a public and private information set. We have tried to investigate what does the information set empirically contains. We utilized the ordered probit and the log-probability model to answer the research questions. We used the micro data set from the ifo business cycle test to investigate the private information set. We focussed on the two main questions for the assessment of the current business situation and the expectations for the next six months. A questionnaire is the only available information we have from a specific firm. If we assume that the two main questions are exogeneous we can regress the other variables on them. As a result we obtain that every variables do have an influence in the opinion formation process. For the assessment the backlog of orders plays the most important role. For the expectation formation the firms follow closely the production expectations. The inclusion of autoregressive lags revealed a strong persistence of the answering behavior. This turned out to be the largest influence compared to the other variables in the questionnaire. These results we obtained both from the ordered probit model as well as from the log-probability model. The latter one allows us to investigate the strength of influence between the trichotomous categories among two variables in question.

The results for public information set are not clear cut. For both the monthly and yearly differences the producer prices play the most important role but the coefficient does not exhibit the expected sign. Average partial effects show some plausible and some implausible results. This could be due to an omitted variable bias in the non-linear model. A possible solution could be to extent the data basis or to account for this bias. We leave this for further research.

References

Bandholz, H., and M. Funke (2003): "In search of leading indicators of economic activity in Germany," *Journal of Forecasting*, 22, 277–297.

Bonham, C., and R. Cohen (2001): "To Aggregate, Pool, or Neither: Testing the Rational Expectation Hypothesis Using Survey Data," *Journal of Business and Economic Statistics*, 19, 278–291.

G. Elliot, I. K., and A. Timmerman (2005): "Estimation and Testing of Forecast RAtionality under Flexible Loss," *Review of Economic Studies*, 72, 1107–1125.

Goodman, L., and W. Kruskal (1979): *Measures of Association in Cross-Classifications*. Springer-Verlag, New York.

Heinz Koenig, M. N., and G. Oudiz (1981): "On the formation of price expectations: An analyis of business test data by log-linear probability models," *European Economic Review*, 16(1), 103–138.

Hfner, F., and M. Schrder (2002): "Prognosegehalt von ifo-Geschftserwartungen und ZEW-Konjunkturerwartungen: Ein konometrischer Vergleich," *Jahrbcher fr Nationalkonomie und Statistik*, 222, 316–336.

James McIntosh, F. S., and W. Low (1989): "A qualitatitive response analyis of UK firms employment and output decisioncs," *Journal of Applied Econometrics*, 4(3), 251–264.

Kholodilin, K. A., and B. Siliverstovs (2006): "On the Forecasting Properties of the Alternative Leading Indicators for the German GDP: Recent Evidence," *Jahrbcher fr Nationalkonomie und Statistik*, 226, 234–259.

Klaus Abberger, Sascha Becker, B. H., and K. Wohlrabe (2006): "Mikrodaten im ifo Institut fr Wirtschaftsforschung - Bestand, Verwendung und Zugang," *mimeo*.

Kukuk, M. (1994): "Haben Unternehmer rationale Erwartungen? Eine empirische Untersuchung," *Ifo Studien*, 40, 111–125.

Muth, J. (1969): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315–335.

Nerlove, M. (1983): "Expectations, plans, and realizations in theory and practice," *Econometrica*, 51(5), 1251–1280.

Pesaran, M. H., and M. Weale (2006): *Handbook of Economic Forecasting*chap. Survey Expectations, p. forthcoming. North-Holland.

Peter Adams, Michael D. Hurd, D. M. A. M., and T. Ribeiro (2003): "Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics*, 112, 3–56.

Wooldridge, J. M. (2005): "Simple solutions to the initial conditions problem for dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of Applied Econometrics*, 20(1), 39–54.

Table 4: Ordered probit models - Sample 1991-2000

	Assessment		Expectations	
	Pooled Probit	RE Probit	Pooled Probit	RE Probit
Q	0.402 (0.005)	0.441 (0.006)	0.029 (0.005)	0.043 (0.005)
$Q*$	0.131 (0.005)	0.130 (0.006)	1.229 (0.005)	1.192 (0.006)
L^a	0.431 (0.005)	0.558 (0.006)	0.027 (0.005)	0.012 (0.006)
D	0.121 (0.005)	0.137 (0.006)	0.328 (0.005)	0.309 (0.006)
A	0.010 (0.005)	0.009 (0.006)	0.147 (0.005)	0.141 (0.005)
A^a	1.180 (0.005)	1.259 (0.006)	0.018 (0.004)	0.042 (0.005)
P	0.175 (0.007)	0.202 (0.007)	0.111 (0.006)	0.127 (0.007)
$P*$	0.043 (0.006)	0.093 (0.006)	0.195 (0.006)	0.215 (0.006)
Cut1	4.146 (0.019)	4.594 (0.026)	3.008 (0.018)	2.934 (0.022)
Cut2	6.396 (0.021)	7.263 (0.029)	5.347 (0.019)	5.521 (0.023)
ρ		0.287 (0.006)		0.205 (0.003)
NT	261272	261272	261272	261272
L	-185897	-162038	-183209	-169646

Standard errors are in in parentheses. For definition of variables see Table 1.

Table 5: Ordered probit models with lagged dependent variable - Sample 1991-2000

	Assessment		Expectations	
	Pooled Probit	RE Probit	Pooled Probit	RE Probit
$t-1$	1.883 (0.007)	1.672 (0.008)	1.125 (0.006)	0.935 (0.007)
Q	0.392 (0.005)	0.401 (0.006)	-0.029 (0.005)	-0.012 (0.005)
$Q*$	0.102 (0.005)	0.099 (0.006)	0.980 (0.005)	1.010 (0.006)
L^a	0.292 (0.005)	0.310 (0.006)	0.015 (0.005)	0.005 (0.006)
D	0.105 (0.005)	0.107 (0.006)	0.272 (0.005)	0.272 (0.006)
A	0.030 (0.005)	0.010 (0.006)	0.106 (0.005)	0.110 (0.005)
A^a	0.846 (0.005)	0.910 (0.006)	-0.005 (0.004)	0.006 (0.005)
P	0.175 (0.007)	0.232 (0.007)	0.074 (0.006)	0.102 (0.007)
$P*$	0.003 (0.006)	0.005 (0.006)	0.178 (0.006)	0.201 (0.006)
Cut1	6.129 (0.019)	6.387 (0.026)	4.100 (0.018)	3.863 (0.022)
Cut2	9.371 (0.021)	9.862 (0.029)	6.850 (0.019)	6.756 (0.023)
ρ		0.178 (0.006)		0.126 (0.003)
NT	211322	211322	211322	211322
L	-163731	-153214	-116200	-99646

Standard errors are in in parentheses. For the definition of variables see Table 1.

Table 6: Estimates of the conditional Model $P(G|G(-1), A^a, Q^*)$

		Variable at				
		Level 1		Level 2		Level 3
Main Effect G		-4.818	(0.446)	-1.234	(0.171)	6.052
Bivariate Interaction $G \times G(-1)$		1.542	(0.395)	0.969	(0.348)	-2.511
		0.915	(0.244)	0.797	(0.186)	-1.797
		-2.457		-1.766		-4.308
γ	0.428					
Bivariate Interaction $G \times A^a$		5.855	(0.524)	3.942	(0.284)	-9.797
		2.299	(0.477)	2.372	(0.144)	-4.671
		-8.154		-6.314		-14.468
γ	0.838					
Bivariate Interaction $G \times Q^*$		1.258	(0.449)	0.897	(0.391)	-2.155
		0.404	(0.269)	0.532	(0.193)	-0.936
		-1.662		-1.429		3.091
γ	0.457					

Averages from 1991-2000. Standard errors in parentheses.

Table 7: Estimates of the conditional Model $P(G^*|G^*(-1), A^a, Q^*)$

		Variable at				
		Level 1		Level 2		Level 3
Main Effect G		-3.343	(0.352)	-1.182	(0.164)	4.525
Bivariate Interaction $G^* \times G^*(-1)$		1.570	(0.391)	0.913	(0.246)	-2.070
		0.989	(0.341)	0.794	(0.185)	-1.783
		-2.559		-1.707		4.266
γ	0.428					
Bivariate Interaction $G^* \times A^a$		1.143	(0.602)	*0.350*	(0.230)	-1.493
		1.381	(0.565)	0.476	(0.186)	-1.857
		-2.524		-0.826		-3.350
γ	0.370					
Bivariate Interaction $G^* \times Q^*$		5.627	(0.509)	2.842	(0.352)	-8.469
		3.048	(0.418)	2.440	(0.174)	5.488
		-8.675		-5.282		-13.957
γ	0.810					

Averages from 1991-2000. Standard errors in parentheses.

Table 8: Goodman-Kruskal Gamma 1991-2000

	Q	Q^*	L^a	D	A	A^a	P	P^*
Business conditions	0.601	0.439	0.653	0.494	0.530	0.870	0.399	0.279
Business expectations	0.445	0.828	0.298	0.559	0.522	0.356	0.333	0.353

Table 9: Ordered probit models - Monthly Differences - Sample 1991-2000

	Assessment		Expectations	
	Pooled Probit	RE Probit	Pooled Probit	RE Probit
OIL	*0.023* (0.070)		0.354 (0.071)	
EXCHANGE	3.662 (0.521)		*0.691* (0.525)	
DAX	-0.308 (0.082)		-0.398 (0.084)	
PROD	1.328 (0.441)		1.511 (0.449)	
TURNOVER-D	-2.377 (0.416)		-2.615 (0.423)	
TURNOVER-EX	2.213 (0.284)		-1.036 (0.290)	
PROD-PRICE	-9.251 (1.849)		-14.502 (1.915)	
IFO-ORDER	-0.004 (0.000)		0.013 (0.001)	
IFO-Climate(-1)	-0.016 (0.000)		-0.017 (0.000)	
Cut1	-1.301 (0.082)		0.183 (0.083)	
Cut2	0.317 (0.082)		2.115 (0.083)	
NT	381514		381514	
Log Likelihood	-360726		-329731	

Time dummies included. Standard Errors in parentheses.

Table 10: Average Partial Effects - Monthly Differences - Sample 1991-2000

	Assessment			Expectations		
	good	satisfactorily	poor	more favorable	unchanged	less favorable
OIL	-0.005	-0.002	0.008	**-0.076**	-0.024	**0.101**
EXCHANGE	**-0.847**	-0.392	**1.239**	-0.149	-0.047	0.196
DAX	**0.071**	0.033	**-0.104**	**0.086**	0.027	**-0.113**
PROD	**-0.307**	-0.142	**0.449**	**-0.324**	**-0.102**	**0.429**
TURNOVER-D	**0.550**	0.254	**-0.804**	**0.565**	**0.178**	**-0.743**
TURNOVER-EX	**-0.512**	-0.237	**0.749**	**0.224**	**0.07**	**-0.294**
PROD-PRICE	**2.139**	**0.990**	**-3.13**	**3.133**	**0.985**	**-4.119**
IFO-ORDER	**0.001**	0.000	**-0.001**	**-0.003**	-0.001	**0.004**
IFO-CLIMATE(-1)	**0.003**	**0.002**	**-0.005**	**0.004**	**0.001**	**-0.005**

Table 11: Ordered probit models - Yearly Differences - Sample 1991-2000

	Assessment		Expectations	
	Pooled Probit	RE Probit	Pooled Probit	RE Probit
OIL	0.080 (0.035)		-0.423 (0.036)	
EXCHANGE	*0.397* (0.362)		2.670 (0.369)	
DAX	*-0.072* (0.059)		-0.122 (0.061)	
PROD	*0.142* (0.494)		-1.445 (0.504)	
TURNOVER-D	-2.154 (0.449)		*0.055* (0.459)	
TURNOVER-EX	0.794 (0.273)		-2.228 (0.282)	
PROD-PRICE	-9.807 (0.903)		13.501 (0.926)	
IFO-ORDER	-0.005 (0.000)		0.006 (0.001)	
IFO-Climate(-1)	-0.013 (0.001)		-0.014 (0.001)	
Cut1	-1.599 (0.101)		-0.353 (0.101)	
Cut2	0.033 (0.101)		1.575 (0.101)	
NT	381514		381514	
Log Likelihood	-321204		-295449	

Time dummies included. Standard Errors in parentheses.

Table 12: Average Partial Effects - Yearly Differences - Sample 1991-2000

	Assessment			Expectations		
	good	satisfactorily	poor	more favorable	unchanged	less favorable
OIL	**-0.017**	-0.1	0.028	**0.912**	0.028	**-0.120**
EXCHANGE	-0.086	-0.052	0.138	**-0.579**	-0.178	**0.758**
DAX	0.015	0.009	-0.025	**0.026**	0.008	**-0.035**
PROD	-0.031	-0.018	0.049	0.314	0.097	**-0.410**
TURNOVER-D	**0.465**	0.284	**-0.749**	-0.012	-0.004	0.016
TURNOVER-EX	**-0.172**	-0.105	**0.276**	**0.484**	0.149	**-0.632**
PROD-PRICE	**2.117**	1.293	**-3.41**	-2.931	-0.902	**3.833**
IFO-ORDER	0.001	0.001	**-0.002**	**-0.001**	0.000	**0.002**
IFO-CLIMATE(-1)	**0.003**	0.002	**-0.004**	**0.003**	0.001	**-0.004**