

Kaiser, Ulrich; Spitz, Alexandra

Working Paper

Quantification of qualitative data using ordered probit models with an application to a business survey in the German service sector

ZEW Discussion Papers, No. 00-58

Provided in Cooperation with:

ZEW - Leibniz Centre for European Economic Research

Suggested Citation: Kaiser, Ulrich; Spitz, Alexandra (2000) : Quantification of qualitative data using ordered probit models with an application to a business survey in the German service sector, ZEW Discussion Papers, No. 00-58, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim

This Version is available at:

<http://hdl.handle.net/10419/24407>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Quantification of Qualitative Data Using Ordered Probit Models with an Application to a Business Survey in the German Service Sector [§]

by

ULRICH KAISER* and ALEXANDRA SPITZ**

November 2000

Abstract: This paper aims at providing business survey analysts with simple econometric tools to quantify qualitative survey data. We extend the traditional and commonly applied method proposed by Carlson and Parkin (1975) to capture observable survey respondent heterogeneity. We also discuss specification tests.

The empirical analysis is based on business survey data taken from the ZEW's 'Service Sector Business Survey', a quarterly business survey in the German business-related service sector carried out since 1994.

JEL classification: C25, L8

Keywords: quantification technique, ordered probit, specification tests

[§]This research was inspired by discussions with Robert Dornau and Winfried Pohlmeier. Ulrich Kaiser gratefully acknowledges financial support by the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG) under grant PF331/3-3.

*Centre for European Economic Research (ZEW), Dep. of Industrial Economics and International Management, P.O. Box 103443, D-68034 Mannheim, Germany, email: kaiser@zew.de; and Centre for Finance and Econometrics at the University of Konstanz (CoFE).

**Centre for European Economic Research (ZEW), Dep. of Industrial Economics and International Management, P.O. Box 103443, D-68034 Mannheim, Germany, email: spitz@zew.de

Non–technical summary

Whenever present day information on the development of an economy or parts of the economy is missing, information gathered from business surveys receive heightened attention. The informational content of business surveys is, however, often limited. This is especially true for surveys in which questions on the state of respondents' business are asked on an ordinal scale. To overcome this shortcoming, techniques for quantifying qualitative surveys were invented in the early fifties.

In 1975, Carlson and Parkin developed a fairly complex solution to the problem of quantifying three-category qualitative survey responses based on the normal distribution. Although their method demands some computational effort, it is the most common applied quantification technique until today.

In this paper, we interpret their methodology in an ordered probit context. This facilitates and speeds up the application since the ordered probit model is included in almost any standard econometric software package. In addition, we extend their method to take into account observable differences across firms. This improves the precision of the quantified survey results.

1 Introduction

Whenever present day information on the development of an economy or parts of the economy is missing, the public interest in information gathered from business surveys receives heightened attention. A major advantage of business surveys is that first results can usually be published within three months after the data collection period has ended. Many economists, such as Oppenländer (1997), claim that this up-to-dateness makes business surveys at least as important as official statistics. A synoptic table provided by the Centre for International Research on Economic Tendency Surveys office (CIRET, 1998) highlights the influence of business surveys: while there were 34 surveys in 15 countries collected in 1960, the number increased to 318 surveys in 57 countries by the end of 1997.

The informational content of business surveys is, however, often limited. Most surveys simply ask questions on the state of the respondents' business on an ordinal scale. A frequent question is, for example, "Did your total sales increase, decrease or remain the same in the current quarter with respect to the preceding quarter". In order to aggregate the information contained in the individual responses, balances — the share of firms reporting increased sales minus the share of firms reporting decreased sales — are calculated. In addition to the more formal aspect that the information contained in the "no change" category is neglected,¹ people may find it difficult to assess the implication of a balance of 20 percent, for example. In particular, if a time dimension is lacking, it is difficult to assess whether this value signals confidence or stagnation.

Carlson and Parkin (1975) developed a fairly simple technique to quantify the qualitative information collected in business tendency surveys. Their method has been extended in many different ways; comprehensive surveys are presented by Geil and Zimmermann (1996), Seitz (1988) and Zimmermann (1985 and 1997).

In this paper, we suggest a simple alternative to the basic Carlson and Parkin (1975) procedure, which has several advantages with respect to 'direct' tests for the crucial assumption of normality and with respect to the incorporation of individual-specific variables that allow control for observed survey respondent heterogeneity.

This paper also introduces a comparatively new dataset, the 'Service Sector Business Survey' (SSBS) to the literature. The SSBS is a quarterly business survey that is collected by the Centre for European Economic Research (Zentrum für Europäische Wirtschaftsforschung, ZEW) in cooperation with Germany's largest credit rating agency Creditreform since June 1994.² Roughly 1,100 firms of the fast growing German business-related services sector regularly take part in the SSBS. The SSBS is unique in the sense that it provides information on an increasingly important part of the German economy that is substantially underrepresented in official statistics. Hax (1998) recently criticized the lack of appropriate data on the service sector that severely hampers business cycle forecasts and economic policy advice. The lack of data for the observation of business cycles in the German business-related service sector appears to even more severe since Kaiser and Voss (2000) have shown, using Granger causality analysis, that manufacturing generally does not lead business-related services in the business cycle. That inadequate data availability on services is not only a particular German problem but also a worldwide problem, as has been stressed by Waller (1997).

¹See Ronning (1984, 1990) for details on this issue.

²Details on the sample design and the survey design are given in Kaiser et al. (2000).

We aim at closely linking quantification methodology with practical implementation and, hence, start by describing a somewhat ‘typical’ business survey, the SSBS, and proceed with a discussion of quantification methods. Finally, we present quantification results and perform specification tests.

Our discussion focuses on the standard ordered probit model. Although it is well established that quantification in an ordered probit context is feasible and simple, it is scarcely applied in practice. In this paper we demonstrate that it is worthwhile to use the ordered probit model for quantification since the inclusion of respondent-specific variables — which is infeasible in the Carlson and Parkin (1975) method — helps to increase the precision of the estimates and substantially reduces the width of the confidence bounds that correspond to the quantified survey results.

2 Data

The SSBS has steadily gained in terms of media attention since its implementation in the second quarter of 1994. It focuses on ten branches of the service sector, which are often referred to as ‘business-related services’. Although no clear-cut and broadly accepted definition of business-related services exists, researchers have agreed upon definitions based on the enumeration of certain sectors. Our definition of business-related services closely follows Hass (1995), Klodt et. al. (1997), Miles (1993) and Strambach (1995). It is displayed in the table below with the corresponding German industrial classification WZ93.³

Sector	WZ 93
Computer Services	72100, 72201-02, 72301-04, 72601-02, 72400
Tax consultancy & Accounting	74123, 74127, 74121-22
Management Consultancy	74131-32, 74141-42
Architecture	74201-04
Technical Advice & Planning	74205-09, 74301-04
Advertising	74844, 74401-02
Vehicle Rental	71100, 71210
Machine Rental	45500, 71320, 71330
Cargo handling & Storage	63121, 63403, 63401
Waste and Sewage Disposal	90001-07

Every three months, ZEW and Creditreform send out a single page questionnaire to about 3,500 firms that belong to the ten sectors listed above. The survey is constructed as a panel data set and currently covers 25 waves. It is a stratified random sample, stratified with respect to the ten sectors, five size classes (two for Eastern Germany and three for Western Germany), and regional affiliation (Eastern/Western Germany). The stratified target population thus consists of 50 cells. A sample refreshment takes place on an annual basis. Firms that have not taken part in the survey for more than six times in a row are removed from the panel. First survey results of the study and a general description of

³The WZ93 industrial classification code is a classification system developed by the German Federal Statistical Office in accordance with the European NACE Rev. 1 standard that classifies economic units according to their sector of concentration.

the survey can be found in Saebetzki (1994). Current survey results are released in the media and in ZEW publications.⁴

The SSBS starts three weeks prior to the end of a quarter. Questionnaires and a personal letter to the prospective survey respondent are sent out by mail. The questionnaires are mostly returned to the ZEW by fax. After two weeks, those firms that have not replied are sent a reminder. Altogether, the response rate amounts to about 30 percent. As a thank you for filling out the questionnaire, the participating firms receive an analysis in the form of a four page report that contains the main findings of the survey. In addition, they can obtain further information over the Internet.⁵

The questionnaire is divided into two parts. In the first part, the firms are requested to indicate on a three point Likert scale whether their sales, prices, demand, returns and number of employees have decreased, stayed the same, or increased in the current quarter in comparison to the previous quarter. Moreover, they are supposed to give an assessment for the forthcoming quarter. The second part of the survey is dedicated to current economic and political issues. Topics cover on-the-job-training, wage negotiation and dispersion of general wage agreements (Kaiser and Pfeiffer 2000; Kaiser and Pohlmeier 2000), innovation and the demand for heterogeneous labor (Kaiser, 1998a), the adjustment to demand fluctuations (Kaiser and Pfeiffer, 2000) and the implications of the introduction of the Euro on firms' export propensity (Kaiser and Stirböck, 1999).⁶

A detailed description of the data set is presented by Kaiser et al. (2000). An overview and selected survey results are reviewed in Kaiser (1999). Public use files — for scientific use only — are available upon request (write to konjunkturumfrage@zew.de).

3 Quantification methods reconsidered

People in charge of collecting business survey data are often hesitant to ask directly for sales, prices, profits, demand or employment. In practice, survey respondents are asked to give a qualitative assessment on their business development on a three or five point Likert scale. There are three main reasons for proceeding this way instead of asking for quantitative assessments. First, firms may be reluctant to report actual figures due to privacy reasons. Second, an inherent risk of asking quantitative questions is that there is a high potential of ending up with information with 'spurious precision', for example respondents may be either unable to report precise figures or they may purposely misreport the actual figures. The third reason may be the most compelling one in terms of practical relevance: it is simply easier and faster to give qualitative instead of quantitative assessments. Asking ordinal questions helps to save the respondents' time and hence helps to improve the total response rate.

When survey respondent i answers questions on an ordinal scale, she implicitly has a threshold model in mind. She indicates increased ('+') sales if the actual change in sales, hereafter abbreviated by Y^* , is above a certain threshold μ_2 . Likewise, if the actual

⁴The ZEW sends current survey results to an interested public. Send an email to konjunkturumfrage@zew.de to receive copies.

⁵The Internet address is: <http://www.zew.de/aktuell/branchenreport/wb-BreportStart.html>

⁶In a related study, Kaiser (1998b) analyzes the impact of political events on answering patterns in business surveys.

change in sales is below a lower threshold μ_1 , she indicates decreased sales (‘—’). If the actual change is between the two thresholds, she reports unchanged (‘=’) sales figures. Clearly, these thresholds may vary across different survey respondents or groups of survey respondents and also across time. In order to clarify things, it is useful to write the threshold model formally as:

$$Y_i = \begin{cases} + & \text{if } Y_i^* > \mu_2 \\ = & \text{if } \mu_1 < Y_i^* \leq \mu_2 \\ - & \text{if } Y_i^* \leq \mu_1 \end{cases}, \quad (1)$$

where Y_i denotes the qualitative sales assessment of respondent i . Let N^+ , $N^=$ and N^- denote the number of individuals who report increased, unchanged and decreased sales figures, respectively, and let N denote the total number of survey respondents. Then the relationship between the choice probabilities and the answering shares can be summarized by the following system of equations:

$$\begin{aligned} P[Y_i = \text{‘+’}] &= P[Y_i^* > \mu_2] = N^+/N \\ P[Y_i = \text{‘=’}] &= P[\mu_1 < Y_i^* \leq \mu_2] = N^=/N \\ P[Y_i = \text{‘-’}] &= P[Y_i^* \leq \mu_1] = N^-/N. \end{aligned} \quad (2)$$

That is, the empirical probabilities to indicate increased, unchanged or decreased sales are simply equal to the shares of the respective answers.

The system of equations (2) nicely illustrates that a straightforward and simple non-parametric, e.g. distribution and parameter-free, estimator for the probability to report increased, unchanged or decreased sales simply is the share of answers for these categories. In order to quantify qualitative information, a distributional restriction concerning the choice probabilities $P[\cdot]$ has to be imposed. Let the actual sales changes Y_i^* be dependent on a constant term, β_0 and an identically and independently distributed error term ϵ_i which follows a distribution function $F(\cdot)$ with mean zero and variance σ^2 : $Y_i^* = \beta_0 + \epsilon_i$. The choice probabilities $P[\cdot]$ are hence given by:⁷

$$\begin{aligned} P[Y_i = \text{‘+’}] &= 1 - F\left(\frac{\mu_2 - \beta_0}{\sigma}\right) \\ P[Y_i = \text{‘=’}] &= F\left(\frac{\mu_2 - \beta_0}{\sigma}\right) - F\left(\frac{\mu_1 - \beta_0}{\sigma}\right) \\ P[Y_i = \text{‘-’}] &= F\left(\frac{\mu_1 - \beta_0}{\sigma}\right). \end{aligned} \quad (3)$$

The choice of the distribution function, often also referred to as the ‘link’ function is arbitrary provided that it is symmetric. However, one must test if the distributional assumption is correct. Common choices are the normal and the logistic distribution. The normal distribution leads to the ordered probit model and the logistic distribution leads to the ordered logit model.⁸ In this paper we shall consider the normal distribution only since this is the distribution function considered by Carlson and Parkin (1975).⁹ Since

⁷Since, e.g., $P[Y_i = \text{‘+’}] = P[\beta_0 + \epsilon_i > \mu_2] = P[\epsilon_i > \mu_2 - \beta_0] = 1 - P\left[\frac{\epsilon_i}{\sigma} < \frac{\mu_2 - \beta_0}{\sigma}\right] = 1 - F\left[\frac{\mu_2 - \beta_0}{\sigma}\right]$.

⁸A discussion of whether ordered or unordered models are appropriate in this context is provided by Ronning (1990).

⁹Choosing either the logistic or the normal distribution merely is a matter of convenience since the distributions are very similar to one another with the logistic distribution having more mass at the tails. It is therefore advisable to consider the logistic distribution instead of the normal distribution if the extreme choice categories, in this case ‘+’ and ‘—’ are heavily populated. The choice of the normal distribution by Carlson and Parkin (1975) was the source of wide criticism, e.g. see Maddala (1990).

increases in β_0 and σ such that the ratio β_0/σ remains constant does not affect either probability and since changes in the parameter corresponding to the constant term in the mean function and in the thresholds such that their distance remains unchanged also do not affect the probabilities, identification restrictions have to be imposed. Standard software packages such as LIMDEP and STATA both set σ to one. LIMDEP furthermore restricts the first threshold parameter to zero and estimates a constant term in the mean function while STATA sets the coefficient of the constant term to zero and estimates all threshold parameters.

If both thresholds are known, the constant term in the mean function β_0 and the standard deviation of the error term σ can be estimated. In this case, quantification by an ordered probit model with known thresholds and the Carlson and Parkin (1975) approach are exactly identical. In fact, such an ordered probit model *is* the Carlson and Parkin method expressed in an alternative way. In the ordered probit context, the estimated parameter $\hat{\beta}_0$ denotes the quantified sales growth rate and the estimated parameter $\hat{\sigma}$ denotes the standard error of the quantified sales growth rate.

An extension of this basic quantification method that uses ordered probit models for one single survey to repeated surveys is straightforward. Let t denote the point in time in which individual i and its survey response is observed and let D_{it} denote a dummy variable that is coded ‘one’ if individual i took part in the t th survey. In order to find quantified sales changes for each of the $t = 1, \dots, T$ survey waves in an ordered probit context, the latent variable is specified as $Y_{it}^* = \sum_{t=1}^T \alpha_t D_{it} + \epsilon_{it}$. The threshold model is then given by:

$$Y_{it} = \begin{cases} + & \text{if} & Y_{it}^* = \sum_{t=1}^T \alpha_t D_{it} + \epsilon_{it} > \mu_2 \\ = & \text{if} & \mu_1 < Y_{it}^* = \sum_{t=1}^T \alpha_t D_{it} + \epsilon_{it} \leq \mu_2 \\ - & \text{if} & Y_{it}^* = \sum_{t=1}^T \alpha_t D_{it} + \epsilon_{it} \leq \mu_1 \end{cases} . \quad (4)$$

The constant term β_0 is now made wave-specific by the inclusion of the dummy variables D . Estimates of the α_t ’s represent the quantified sales changes at time t . Estimates for the standard error of the quantification can be obtained by specifying the standard error of the disturbance term as $\sigma_{it} = \exp(\sum_{t=1}^T \gamma_t D_{it})$, where γ_t are the estimated parameters.¹⁰

As opposed to the linear regression model in which the estimated parameters retain their consistency even when the error terms are non-normal, not identical and not independent, the parameters of the ordered probit model become inconsistent in these cases. Specification tests are therefore advisable though rarely used in applied econometric work. We will return to this issue after having presented quantification results in Section 4.

The standard error of quantified ordinal information usually is much lower when survey respondents give an assessment on *overall* economic issues compared to the situation when they judge their *own* business condition. In both cases, the variance in the answers is attributable to heterogeneity across the survey respondents. However, though opinions on the state of the overall economy may of course differ among survey respondents, the deviation of judgements on the state of their *own* businesses are likely to be much larger. In fact, variations of these opinions may be dependent upon firm size, regional affiliation (Eastern/Western Germany) or sector affiliation. It is thus straightforward to incorporate these differences within the specification of the standard deviation of the error term.

¹⁰The exponential function is taken in order to avoid negative standard deviations.

Let SC_{ik} denote the k th firm size class of respondent i , let $East$ denote a dummy variable for Eastern German firms and let $Sector_l$ denote the l th sector. The standard error of the disturbance term is then given by $\sigma_i = \exp\left(\sum_{t=1}^T \gamma_t D_{it} + \sum_{k=1}^{K-1} \delta_k SC_{ik} + \kappa East_i + \sum_{l=1}^{L-1} \zeta_l Sector_l\right) = \exp(\boldsymbol{\theta} \mathbf{z}_i)$ for $i = 1, \dots, N$, where the k th size class and the l th sector are the reference groups. Likewise, it seems reasonable that the same set of variables affects not only the variation of individual responses but also the growth rate and thus the choice of the answering category so that $Y_{it}^* = \sum_{t=1}^T \alpha_t D_{it} + \sum_{k=1}^{K-1} \rho_k SC_{ik} + \tau East_i + \sum_{l=1}^{L-1} \varphi_l Sector_l + \epsilon_{it} = \mathbf{x}_i \boldsymbol{\beta}$.

The inclusion of the explanatory variables is equivalent to moving the threshold parameters μ around. This implies that if explanatory variables such as firm size and regional and sector affiliation are included in the specification, this is equivalent to specifying group-specific threshold parameters.

It is straightforward to obtain sector-specific sales growth rates for example by simply interacting the wave dummy-variables with the sector dummy-variables. The coefficients obtained from such an estimation reflect the wave-specific and sector-specific sales growth rates.

Another extension of the standard ordered probit model as described in this section is the ordered panel probit model. Many business surveys are constructed as panel data sets and it seems advisable to explicitly use this additional information. The main advantage of panel data models is that unobserved heterogeneity of the individuals i can be taken into account. In this case, the error term ϵ_{it} is specified as the sum of two components: $\epsilon_{it} = \xi_i + \nu_{it}$. The term ξ_i is assumed to be a time independent individual-specific random variable, reflecting unobserved firm heterogeneity while ν_{it} is assumed to be an error term that is independent both among individuals and over time. Both error terms are assumed to be normally distributed with zero means.

The ordered probit model, as discussed above, is a so-called ‘pooled’ ordered probit model. That is, we do not take into account the additional information contained in our panel data set by assuming the error term ϵ_{it} to be independent and identically distributed with a mean of zero and variance σ^2 for all individuals i and over time t .

Two principles for estimating panel data models exist: the ‘fixed effects’ and the ‘random effects’ approaches. Fixed effects estimation assumes the presence of an individual-specific effect ξ_i and independence of the error term component ν_{it} . In this nonlinear specification, the fixed effects ξ_i and the coefficients α_t are unknown parameters and have to be estimated. In this case, the maximum likelihood estimator is only consistent when T tends to infinity. When T is finite, as is usually the case, the *incidental parameter problem* (Neyman and Scott, 1948) occurs: there is only a limited number of observations of Y_{it} for each individual i , $t = 1, \dots, T$, that contain information about ξ_i . Furthermore, an increase of the cross-sectional units, N , provides no information about ξ_i , but it increases the number of parameters ξ_i . The result is that any estimation of ξ_i is meaningless if T is finite, even if N is large. Unfortunately, the maximum likelihood estimators α_t and ξ_i cannot be separated in the nonlinear qualitative response models as is the case for linear models. When T is finite, the inconsistency of the estimated ξ_i is transmitted into the estimation of α_t . Chamberlain (1984) suggested an approach to remove the unobserved heterogeneity in multinomial logit models.¹¹ Such an estimator does not exist, however,

¹¹This approach is based on a conditional likelihood approach proposed by Anderson (1970, 1973). The baseline idea is to remove the incidental parameters by writing the multinomial logit model in terms of a

for ordered panel data models due to the existence of the threshold parameters.¹² Random effects estimation in the ordered probit context is feasible, even in standard software packages such as LIMDEP. Instead of estimating N parameters ξ_i as in the fixed effect model, only the mean and variance are estimated. It only leads to efficiency gains if significant random effect are present, e.g., if the error components ν_{it} are correlated over time. The pooled panel ordered probit estimator, however, retains its consistency.¹³ For the sake of brevity, we will therefore not discuss the random effects ordered probit model in further detail. Comprehensive discussions are presented by Hamerle and Ronning (1995) as well as Tutz and Hennevogl (1996). A recent application of the random effects ordered probit model is presented in Kaiser and Pfeiffer (2000).

To summarize, quantification of qualitative survey data by ordered probit models has two main advantages: (i) it allows for group-specific thresholds by the inclusion of explanatory variables and (ii) it allows one to explicitly take into account the variation of survey responses among the responding individuals. Further advantages are that tests for normality and heteroscedasticity can fairly easily be implemented and tests of identity of sales changes in individual quarters can be easily conducted by using a Wald test. The latter two topics will be discussed in further detail below.

4 Quantification results

A key question in any quantification context is the derivation of the threshold values. Carlson and Parkin (1975) estimated thresholds by assuming long-term unbiasedness.¹⁴ It is common practice to directly ask the survey respondents for the minimum value to which actual sales have to increase (decrease) before they report increased (decreased) sales figures once and then to assume that these values remain constant during the next couple of months or years.¹⁵

Proceeding this way is, however, not a sensible approach for the SSBS since this data set is not well balanced, for example the fluctuation of responding firms is quite large so that a considerable share of firms that has answered in survey wave t when it is asked for the individual thresholds is likely not to answer at $t + s$ and vice versa.¹⁶

Therefore, the threshold parameters were obtained from another data set which was also compiled by the ZEW, the Mannheim Innovation Panel in the Service Sector (MIP-S).¹⁷ The MIP-S covers very similar sectors as the SSBS and has up to now been conducted four times, in 1995, 1997, 1998 and 1999. In 1997, the participating firms were asked

conditional maximum likelihood function. In probit models, the conditional maximum likelihood method does not remove the individual specific effects, however.

¹²Also note that time-invariant variables such as sector or regional affiliation have to be removed from the specification since they are absorbed by the fixed effect.

¹³This is, in fact, a main reason that the application of the random effect model is scarce in the empirical literature.

¹⁴That is, they estimated the threshold by scaling the estimated industry-wide inflation rate so that the sample average of the estimated series are equal to the actual observed rate of buying-price inflation.

¹⁵Threshold values for the well known ZEW Financial Market Test (for more information see <http://www.zew.de/projekte/epl?action=detail&nr=6&lang=eng>) are, e.g., obtained that way.

¹⁶See Kaiser et al. (2000) for more information on the stability of the panel data set.

¹⁷A thorough description of this data set is presented by Janz et al. (2000).

to indicate on a five point scale whether their sales improved, remained unchanged or decreased within the last three years. Due to the panel structure of this data set, we were able to compare this qualitative assessment with the actual changes in total sales. We have calculated the median changes in sales — corrected to take into account that the SSBS asks for quarterly sales changes — for those firms that reported increased (decreased) sales figures as the upper (lower) threshold parameters. The respective value for μ_2 is 1.3 and the value for μ_1 is -0.5 percent. That is, we have found evidence for the presence of asymmetric thresholds: actual sales changes have to exceed a considerably higher threshold before firms report increased sales figures than the other way around. Besides the obvious psychological explanation that people tend to overstate bad economic or personal situations compared to good ones, Batchelor (1986) argues that individuals’ answers may be subject to strategic behaviour, e.g. firms are more likely to report pessimistic results, in the hope of getting subsidies for their industry. Our finding of asymmetric thresholds supports the criticisms of the Carlson and Parkin (1975) approach, which assume symmetric thresholds.¹⁸

Positive sales changes						
Wave	minimum	10%	median	90%	mean	std. dev.
20	1.1	4.0	10.0	20.0	11.4	7.5
21	0.9	3.0	10.0	20.0	11.2	7.3
22	1.0	4.0	10.0	20.0	11.0	6.7
23	1.1	3.5	10.0	20.0	11.2	7.0
24	1.3	3.0	10.0	21.0	11.3	7.3
25	0.9	3.0	10.0	20.0	10.9	7.5
<i>mean</i>	<i>1.1</i>					

Negative sales changes						
Wave	maximum	10%	median	90%	mean	std. dev.
20	-0.5	-25.0	-10.0	-5.0	-13.7	8.0
21	-0.8	-24.0	-10.0	-4.2	-11.5	7.6
22	-0.9	-20.2	-10.0	-3.0	-11.6	7.5
23	-0.6	-20.0	-10.0	-3.9	-12.3	7.4
24	-0.6	-25.0	-10.0	-5.0	-13.9	8.0
25	-0.7	-25.0	-10.0	-5.0	-12.9	7.0
<i>mean</i>	<i>-0.7</i>					

TABLE 1: Descriptive statistics of the actual sales changes reported in the SSBS

In order to compare the thresholds derived from the MIP-S and the SSBS, Table 1 displays descriptive statistics of the actual sales changes reported by the firms interviewed in the SSBS since wave 20. The minimum value corresponding to the positive sales changes (upper panel) can be regarded as the bound *above* which firms indicate increased sales changes. The mean minimum (maximum) value of the actual sales changes reported by firms with increased (decreased) sales changes are 1.1 (-0.7) so that they compare well to

¹⁸Other studies explain the existence of the ‘stay the same’ category by considerations concerning the cost-intensive information acquisition process (Fishe and Idson, 1989).

our estimated thresholds of 1.3 and -0.5.

A crucial assumption of the Carlson and Parkin (1975) approach is that the threshold parameters are time-invariant. This also is a source of wide criticism. Batchelor (1986) argues that the threshold parameters should be allowed to be a function of the size and variability of the stimulus. Some empirical papers investigate the appropriateness of varying threshold parameters in the field of inflation expectations. In this context, Seitz (1988) does not find that the threshold parameters are dependent on the level or variance of inflation. Dasgupta and Lahiri (1992) demonstrate that, in the case of inflation expectations, although varying thresholds help to capture extreme values better, they do not improve the resulting quantitative series.¹⁹ Having a glance at Table 1 shows, that time-invariant thresholds might be a sensible choice here: the 10, 50 and 90 percent percentiles as well as means and standard errors of the actual sales changes do not differ much through time.

	Coeff.	Std. err.		Coeff.	Std. err.
	Conditional mean			Conditional variance	
α_1	1.1784	0.0965	σ_1	2.0049	0.4749
α_2	1.0411	0.0687	σ_2	1.5533	0.3246
α_3	1.2631	0.0676	σ_3	1.5128	0.3157
α_4	0.6948	0.0763	σ_4	1.5878	0.3646
α_5	0.9332	0.0626	σ_5	1.5443	0.2961
α_6	0.8596	0.0628	σ_6	1.4731	0.2943
α_7	0.8967	0.0649	σ_7	1.4915	0.3046
α_8	0.2012	0.0722	σ_8	1.7390	0.3549
α_9	0.5316	0.0661	σ_9	1.6052	0.3172
α_{10}	0.4370	0.0646	σ_{10}	1.5049	0.3040
α_{11}	0.6707	0.0637	σ_{11}	1.4122	0.2944
α_{12}	0.1455	0.0823	σ_{12}	1.7710	0.4058
α_{13}	0.7503	0.0603	σ_{13}	1.7233	0.2943
α_{14}	0.7565	0.0630	σ_{14}	1.7636	0.3098
α_{15}	0.8022	0.0626	σ_{15}	1.7325	0.3057
α_{16}	0.5690	0.0682	σ_{16}	1.8531	0.3413
α_{17}	0.8997	0.0606	σ_{17}	1.6252	0.2902
α_{18}	0.7522	0.0571	σ_{18}	1.5008	0.2689
α_{19}	0.9674	0.0574	σ_{19}	1.5978	0.2735
α_{20}	0.3785	0.0629	σ_{20}	1.7069	0.3080
α_{21}	0.7589	0.0492	σ_{21}	1.4557	0.2294
α_{22}	0.7428	0.0530	σ_{22}	1.4677	0.2479
α_{23}	0.9973	0.0548	σ_{23}	1.5622	0.2593
α_{24}	0.5529	0.0642	σ_{24}	1.7346	0.3155
α_{25}	0.8590	0.0521	σ_{25}	1.4218	0.2420

TABLE 2: Ordered Probit estimation results: baseline model

¹⁹The authors used the Producer Price Index for intermediate materials and components for manufacturing as their benchmark for the quantification results of the National Association of Purchasing Managers survey.

Our baseline estimation is the one in which wave dummy variables are included in the quantification only. Results are shown in Table 2. Table 2 displays the estimated sales growth rates α and the corresponding standard errors σ (instead of the vector of parameters γ).²⁰ Each of the coefficients in the mean function and the variance function are highly significantly different from zero except for α_{12} . The weak significance of this wave dummy variable related to the 12th wave, the first quarter of 1997, implies that this is the quarter where sales growth was lowest (0.1455 percent). Inversely, the highest sales growth is dated back to the fourth quarter of 1994 (third wave, α_3 , 1.2631 percent). The standard deviation of the error term σ reflects the heterogeneity of the firms participating in the SSBS so that it is rather surprising that the precision of the quantification is quite low. The heterogeneity of sales growth rate was largest in the second quarter of 1994, which might simply reflect that firms had to get used to the SSBS questionnaire. Interestingly, heterogeneity of growth rates was lowest in the fourth quarter of 1996 (11th wave, 1.4122 percent) and hence coincides with a remarkable increase in the sales growth rates.

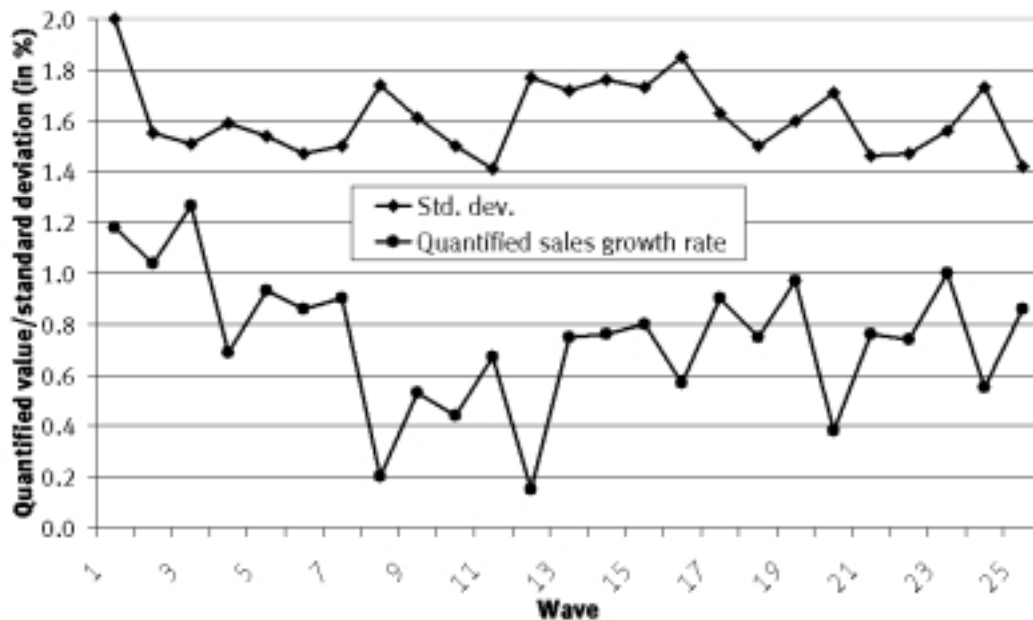


FIGURE 1: Quantified Sales Growth Rates and Corresponding Standard Errors

Quantified sales growth rates vary considerably across the period of investigation. There are two reasons for this pattern: (i) expansion factors have not been attached to the individual respondents and (ii) the figures have not been seasonally adjusted. The first issue

²⁰The corresponding asymptotic standard errors for σ were obtained using the ‘Delta’-method (Greene 1997, ch.6.7.5). All estimation results displayed in this paper are obtained using our own procedure programmed for the standard software package STATA6.0. The program code (implemented as an ‘ADO’-file) can be downloaded from the internet at <ftp://ftp.zew.de/quant.ado>. GAUSS files can be downloaded from <ftp://ftp.zew.de/quant.prg>. The standard software package LIMDEP allows for ordered probit estimation with known thresholds without requiring its own programming efforts.

can easily be implemented in maximum likelihood procedures,²¹ the second topic can be tackled using familiar seasonal adjustment methods.²² In order to keep things as simple as possible, both issues are not considered here.

Table 3 displays estimation results of the extended model. In addition to the set of the wave dummy variables, we include control variables for observable firm heterogeneity. These variables include two firm size dummy variables (1–50 and over 100 employees with firms that have between 51 and 100 employees serving as the base category), a dummy variable for Eastern Germany and nine sector dummy variables (the sectors listed in section 2 have waste and sewage disposal as a base category). A comparison of both results shows only slight and unsystematic effects on the quantified sales growth rates. The standard errors $\hat{\sigma}_t$ of the quantified sales growth rates, however, are considerably reduced as displayed in Figure 2.²³ In order to retain the visibility of the firm size, the regional and the sector affiliation effect, Table 3 directly displays the coefficients of the wave dummy and the observable firm heterogeneity variables and not, as in Table 2, the values of σ .

The coefficients related to the mean function are all significantly different from zero at the one percent significance level except for the wave dummy variable related to the 12th wave, which is insignificant, and for the dummy variable for technical planning, which is significant at the ten percent level only.

The estimation results for the mean function indicate that larger firm are more likely to grow than smaller firms. Eastern German firms usually have smaller sales growth rates than their Western German competitors. Growth rates are smallest for Management consultancy and Computer services, and are smallest for Architecture.

The estimation results for the conditional variance indicate that the heterogeneity of the business development is largest in a firm with 50–100 employees; a U-shaped effect of firm size on the variance is present. Eastern German firms do not significantly differ from their Western German competitors in the variation of survey answers. The variability of survey responses is smallest for tax consultants and largest for advertising firms.

The wave, size class and sector dummies are also jointly highly significant both in the conditional mean and the conditional variance.

The additional explanatory variables in the mean and in the variance are highly significant from zero as a Likelihood ratio test shows ($\chi^2_{12} = 985.91$ with critical values 18.55, 21.03 and 26.22 at the 10, 5 and 1 percent significance level, respectively).

²¹The STATA-ADO file, which can be downloaded from the internet allows the inclusion of such expansion factors.

²²See Kaiser and Buscher (1999) for a suggestion to seasonally adjust short-time series.

²³In order to maintain the comparability of results, the standard errors of the extended model displayed in Figure 2 refer to a model that included the additional explanatory variables in the variance function only.

	Coeff.	Std. err.		Coeff.	Std. err.
Conditional mean			Conditional variance		
	α_1	1.1908		γ_1	-3.9195
	α_2	1.0514		γ_2	-4.1742
	α_3	1.2316		γ_3	-4.2467
	α_4	0.6651		γ_4	-4.1990
	α_5	0.9444		γ_5	-4.2522
	α_6	0.8623		γ_6	-4.2812
	α_7	0.8928		γ_7	-4.2719
	α_8	0.1986		γ_8	-4.1720
	α_9	0.5213		γ_9	-4.2144
	α_{10}	0.4308		γ_{10}	-4.2487
	α_{11}	0.6338		γ_{11}	-4.3643
	α_{12}	0.1386		γ_{12}	-4.0956
	α_{13}	0.6804		γ_{13}	-4.1760
	α_{14}	0.6899		γ_{14}	-4.1571
	α_{15}	0.7328		γ_{15}	-4.1911
	α_{16}	0.4996		γ_{16}	-4.1188
	α_{17}	0.7974		γ_{17}	-4.2776
	α_{18}	0.6820		γ_{18}	-4.3242
	α_{19}	0.8658		γ_{19}	-4.2869
	α_{20}	0.2961		γ_{20}	-4.1966
	α_{21}	0.6567		γ_{21}	-4.3067
	α_{22}	0.6767		γ_{22}	-4.3146
	α_{23}	0.8908		γ_{23}	-4.2660
	α_{24}	0.4377		γ_{24}	-4.1738
	α_{25}	0.7576		γ_{25}	-4.3276
1–50 employees	-0.1139	0.0303	1–50 employees	0.0687	0.0227
> 100 employees	0.1568	0.0380	> 100 employees	0.0530	0.0284
Eastern Germany	-0.2943	0.0271	Eastern Germany	-0.0317	0.0200
Comp. services	0.6142	0.0526	Comp. services	0.1021	0.0375
Tax cons.	0.3280	0.0471	Tax cons.	-0.1528	0.0367
Management cons.	0.6515	0.0569	Management cons.	0.0950	0.0405
Architecture	-0.3726	0.0491	Architecture	-0.0501	0.0370
Technical advice	-0.0781	0.0432	Technical advice	-0.0366	0.0324
Advertising	0.2421	0.0580	Advertising	0.1745	0.0407
Vehicle rental	0.2529	0.0675	Vehicle rental	0.1657	0.0469
Machine rental	0.1742	0.0592	Machine rental	0.1561	0.0419
Cargo handling	0.2604	0.0493	Cargo handling	0.0045	0.0370

TABLE 3: Ordered Probit estimation results: extended model

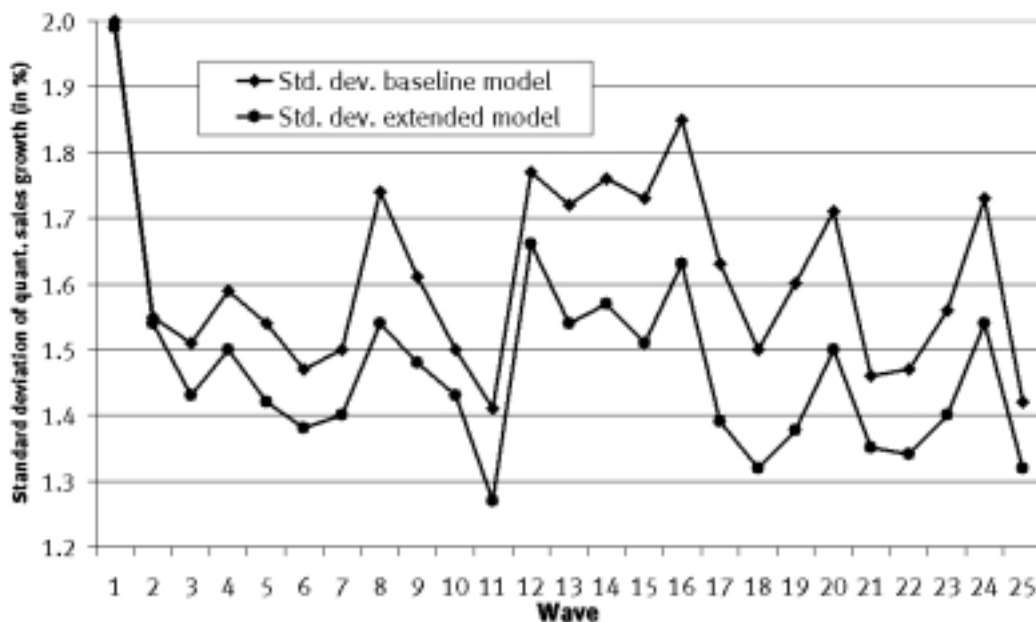


FIGURE 2: Comparison of Standard Errors of Quantified Sales Growth Rates

5 Specification Tests

As noted above, heteroscedasticity and non-normality of the standard error of the disturbance term ϵ_i lead to inconsistent parameter estimates of the ordered probit model. Tests for heteroscedasticity and non-normality can easily be implemented in applied empirical work by initially calculated generalized residuals (Chesher and Irish, 1987) and by then calculating the appropriate test statistics. The generalized residuals of a q -categorical ordered probit model are given by:

$$\hat{\epsilon}_i^{G,q} = \sigma_i \frac{\phi\left(\frac{\mu_q - x_i' \beta}{\sigma_i}\right) - \phi\left(\frac{\mu_{q+1} - x_i' \beta}{\sigma_i}\right)}{\Phi\left(\frac{\mu_{q+1} - x_i' \beta}{\sigma_i}\right) - \Phi\left(\frac{\mu_q - x_i' \beta}{\sigma_i}\right)}. \quad (5)$$

Let z_i denote the vector of variables suspected of causing heteroscedasticity. The LM test statistic for heteroscedasticity can then be obtained by linearly regressing the interaction terms $\hat{\epsilon}_i^G(x_i \beta)$ and $(\hat{\epsilon}_i^G(x_i \beta))z_i$ upon a vector of ones. The LM test statistic is N times the uncentered R^2 of this auxiliary regression and is χ^2 distributed with degrees of freedom equal to the number of variables potentially causing heteroscedasticity.

It is straightforward to apply this type of test to our baseline model from Table 2 assuming that firm size, sector and regional affiliation may cause heteroscedasticity. Since our control variables for unobserved firm heterogeneity include dummy variables only, we just obtain 37 different generalized residuals (25 wave dummies, 9 sector dummies, 2 size class dummies and 1 dummy for Eastern Germany) so that this type of test does not make

much sense. If additional information such as the number of employees in absolute term is available, a test for heteroskedasticity as sketched above can simply be calculated. An alternative test for heteroscedasticity is readily available by comparing the log-likelihood value of the baseline model with the model including the firm heterogeneity variables in the variance (but not in the mean) function. A simple Likelihood ratio test can then be performed. It turns out that the firm heterogeneity variables are jointly highly significantly different from zero in the variance function, which implies that these variables cause unobserved heteroscedasticity.

Another main source of criticism of the Carlson and Parkin (1975) method is their assumption of normally distributed price expectations — or, equivalently, non-normal error terms — which Carlson (1977) himself found to be non-normal. In this context, it seems advisable to test for the distribution of respondents' sales assessments. This test can be performed as well by using an auxiliary regression of the interaction terms $\hat{\epsilon}_i^G \mathbf{x}_i$, $\hat{\epsilon}_i^G (\mathbf{x}_i \boldsymbol{\beta})^2$ and $\hat{\epsilon}_i^G (\mathbf{x}_i \boldsymbol{\beta})^3$ on a vector of ones. The corresponding LM test statistic is χ^2 distributed with two degrees of freedom. The coefficient related to the term $\hat{\epsilon}_i^G (\mathbf{x}_i \boldsymbol{\beta})^2$ corresponds to skewness, the term $\hat{\epsilon}_i^G (\mathbf{x}_i \boldsymbol{\beta})^3$ corresponds to kurtosis.

Unfortunately, such normality tests are infeasible if heteroscedasticity is present as indicated by simulation results by Davidson and MacKinnon (1992).

This reveals issues for future research, e.g. quantification in a non-parametric setting where the distribution is based on a kernel density estimation. Based on earlier findings of non-normal error terms, such as that by Carlson (1977), it seems likely that normality has to be rejected quite often in practice. It therefore seems advisable to non-parametrically estimate the link function F . This issue, however, has to be left to further research.

6 Conclusion

This paper reviews the probably most important technique to quantify qualitative survey data: the quantification method proposed by Carlson and Parkin (1975). We interpret their methodology in an ordered probit context and show that respondent-specific variables can be easily implemented in this type of estimation approach. The ordered probit model is particularly simple to apply since it is included in standard econometric software packages such as LIMDEP and STATA.

Using data taken from a quarterly business survey in the German business-related services sector, we demonstrate that the inclusion of such firm-specific variables such as regional and sectoral affiliation or firm size may substantially reduce the inaccuracy of the standard error of the quantified variables.

Quantification by means of an ordered probit model also enables the analyst to test for significant effects of firm size for example, on survey responses and on their variability. Moreover, tests for misspecification such as non-normality of the error term or heteroscedasticity which lead to inconsistent parameter estimates can be implemented using standard econometric software packages.

Although the pace of the development of quantification techniques has slowed down remarkably within recent years, there still are avenues for further research. An important

aspect in this context is to non-parametrically estimate the distribution function, linking individual survey responses to the quantified value. This issue will be discussed in our further research.

References

- BATCHELOR, R.A. (1986), The Psychophysics of Inflation. *Journal of Economic Psychology* **7**, 269–290.
- CARLSON, J.A. (1977), Are Price Expectations Normally Distributed? *Journal of American Statistical Association* **70**, 749–754.
- CARLSON, J.A., PARKIN, M. (1975), Inflation Expectations. *Economica* **42**, 123–138.
- CHAMBERLAIN, G. (1984). Panel Data. In *Handbook of Econometrics, Vol. II.* (Griliches, Z., Intriligator, M., Eds.) North-Holland, Amsterdam.
- CHESHER, A., IRISH, M. (1987). Residual Analysis in the Grouped and Censored Normal Linear Model. *Journal of Econometrics* **34**, 33–61.
- CIRET (1998). *International Business, Investment and Consumer Surveys: A Synoptic Table.* CIRET office, Munich.
- DASGUPTA, S., LAHIRI, K. (1992). A Comparative Study of Alternative Methods of Quantifying Qualitative Survey Responses Using NAPM Data. *Journal of Business & Economic Statistics* **10/4**, 391–400.
- DAVIDSON, R., MACKINNON, J.G. (1992). A New Form of the Information Matrix Test. *Econometrica* **60**, 145–157.
- FISHE, R.P.H., IDSON, T.L. (1989). Information-Induced Heteroscedasticity in Price Expectation Data. *The Review of Economics and Statistics*, 304–312.
- GEIL, P., ZIMMERMANN, K.F. (1996). Quantifizierung qualitativer Daten. In *Konjunkturindikatoren* (Oppenländer, K.H., Ed.). Vahlen, München.
- GREENE, W.H. (1997), *Econometric Analysis*, Prentice-Hall, Upper Saddle River.
- HAMERLE, A., RONNING, G. (1995). Panel Analysis for Qualitative Variables. In *Handbook for Statistical Modelling for the Social and Behavioural Sciences* (Arminger, G., Clogg, C.C., Sobel, M.E., Eds), Plenum Press.
- HASS, H.-J. (1995). Industrienähe Dienstleistungen: ökonomische Bedeutung und politische Herausforderung. *Beiträge zur Wirtschafts- und Sozialpolitik* **223**, 3–39.
- HAX, H. (1998). Anforderung an die Statistik zur Konjunkturbeobachtung durch den Sachverständigenrat. *Allgemeines Statistisches Archiv* **82**, 15–24.
- JANZ, N., EBLING, G., GOTTSCHALK, S., NIGGEMANN, H., PETERS, B. (2000). The Mannheim Innovation Panel (MIP): Survey Methodology and Empirical Research. *ZEW mimeo*, presented at the 9th International Conference on Panel Data at Geneva, June 22–23, 2000.
- KAISER, U. (1998). Die Wirkung wirtschaftspolitischer Ereignisse auf Erwartungen über den Starttermin der Europäischen Währungsunion, *Ifo-Studien* **2/1998**, 103–120.

- KAISER, U., BUSCHER, H.S. (1999). The Service Sentiment Indicator—A business Climate Indicator for the German Business—related Services Sector. *Center of Finance and Econometrics discussion paper 99-06*.
- KAISER, U., STIRBÖCK, C. (1999). The Impact of the Introduction of the Euro on Firms' Expectation Concerning Export Behaviour, Product Innovation and Foreign Competition: An Empirical Assessment of the German Business—Related Services Sector. *intereconomics* **34(3)**, 107–115.
- KAISER, U., KREUTER, M., NIGGEMANN, H. (2000). The ZEW/Creditreform Business Survey in the Business—Related Services Sector: Sampling Frame, Stratification, Expansion and Results,” *ZEW Discussion Paper 00-22*.
- KAISER, U., PFEIFFER, F. (2000). Collective Wage Agreement and the Adjustment of Workers and Hours in German Service Firms. *ZEW Discussion Paper 00-33*.
- KAISER, U., VOSS, K. (2000). On the Lead/Lag Relationship between Manufacturing and Services in the Business Cycle. *Ifo-Studien* **2/2000**, 161-176.
- KLODT, H., MAURER, R., SCHIMMELPFENNIG, A. (1997). “*Tertiärisierung in der deutschen Wirtschaft*,” Kieler Studien **283**, Tübingen.
- MADDALA, G.S. (1990). Survey Data on Expectations: What Have We Learnt? In *Issues in Contemporary Economics, Vol. 2: Macroeconomics and Econometrics* (Nerlove, M., Ed.), Macmillan.
- MILES, I. (1993). Services in the New Industrial Economy. *Futures*, 653–672.
- NEYMAN, J., SCOTT, E.L. (1948). Consistent Estimates based on Partially Consistent Observations. *Econometrica* **16**, 1–32.
- OPPENLÄNDER, K. H. (1997). Characteristics and Classification of Business Cycle Indicators. In *Business Cycle Indicators* (Oppenländer, K.H., Hrsg.), Aldershot.
- RONNING, G. (1984). Welche Informationen enthält die Antwortkategorie ‘unverändert’ in Tendenzbefragungen? *Ifo-Studien* **18**, 261–272.
- RONNING, G. (1990). The Informational Contents of Responses from Business Surveys. In *Microeconometrics: Surveys and Applications*, (Florens, P.-J., Laisney, F., Eds.), Basil Blackwell.
- SAEBETZKI, A. (1994). Die ZEW-Umfrage bei Dienstleistungsunternehmen: Panellaufbau und erste Ergebnisse. *ZEW Dokumentationen 94-05*.
- SEITZ, H. (1988). The estimation of inflation forecasts from business survey data. *Applied Economics* **20**, 427–438.
- STRAMBACH, S. (1995). *Wissensintensive unternehmensnahe Dienstleister: Netzwerke und Interaktion*. Münster (Diss.).
- Tutz, G. and W. Hennevogl (1996), Random Effects in Ordinal Regression Models, *Computational Statistics & Data Analysis* **22**, 537–557.

- WALLER, S. (1997). Struktureller und sozialer Wandel — Konsequenzen für die amtliche Statistik: Ausgewählte Beiträge von der 23. CIRET-Konferenz in Helsinki. *Ifo-Schnelldienst 29/1997*.
- ZIMMERMANN, K.F. (1985). Innovationsaktivitäten, Preisinflexibilität, Nachfragedruck und Marktstruktur. In *Industrieökonomik-Theorie und Empirie*, (Bombach, G., et al., Eds.), Tübingen.
- ZIMMERMANN, K.F. (1997). Analysis of Business Surveys. *Handbook of Applied Econometrics, Vol. II: Microeconometrics*. (Pesaran, M.H., Schmid, P., Eds.). Basil Blackwell.