

Etter, Richard

Working Paper

Do response rates and responses differ between mail and internet answers in a mixed mode survey? A comparison of qualitative business survey data collected by regular mail and internet questionnaires

KOF Working Papers, No. 68

Provided in Cooperation with:

KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Etter, Richard (2002) : Do response rates and responses differ between mail and internet answers in a mixed mode survey? A comparison of qualitative business survey data collected by regular mail and internet questionnaires, KOF Working Papers, No. 68, ETH Zurich, KOF Swiss Economic Institute, Zurich

This Version is available at:

<https://hdl.handle.net/10419/50850>

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.

Arbeitspapiere/ Working Papers

Richard Etter

Do Response Rates and Responses Differ between Mail and Internet Answers in a Mixed Mode Survey?

A Comparison of Qualitative Business Survey Data
Collected by Regular Mail and Internet Questionnaires

Do Response Rates and Responses Differ between Mail and Internet Answers in a Mixed Mode Survey? *

A Comparison of Qualitative Business Survey Data Collected by Regular Mail and Internet Questionnaires

Richard Etter, KOF, Zurich

Abstract

For many years, Business Tendency Surveys on the basis of mailing suffer from an erosion of the response rate. To counter this problem, there are traditional methods as limitation of the number of questions, improvement of the design of the questionnaire, intensified recalls by mail or telephone, more attractive outputs and a time-reduction of the delivery of the results. Another way to stop the reduction of the response rate is to offer different modes of surveys. Therefore, the KOF introduced an online survey in addition to the mail survey.

Up to now, there are only few research projects which analysed the effects of the new technique of the online survey on the response rate and the quality of the responses. First studies indicate that mail surveys produce higher response rates than internet surveys. Our results show a big difference in the unit non-response rate too. Differentiating according to language regions, branches, size of firms and by the percentage of exports on sales does not change the picture.

If there is a different behaviour of firms between the two modes this could be reflected not only on the response rate but also on the selection of the response category. This would have a significant draw-back for the survey results and their interpretation. The results show at least in the case of "Total manufacturing industry" that the differences between the two modes can be neglected.

The results of business tendency surveys are rarely used on a micro level. The internationally common aggregation technique is the balance method. We therefore analysed the effect of the differences of the two survey modes on the balance indicators. The examination had to be restricted to the monthly survey in manufacturing industry. Thirteen out of fifteen questions have equal means in the two survey modes.

Key Words: survey mode, unit-non-response, item-non-response, ordinal measures of association, business tendency survey

JEL Classification: C42; D21

* Prepared for the 26th CIRET Conference in Taipei, Taiwan
16th - 19th October 2002

Contents

<i>1</i>	<i>Introduction</i>	<i>2</i>
<i>2</i>	<i>The Data</i>	<i>3</i>
<i>3</i>	<i>Non-Response Rates in Different Survey Modes</i>	<i>4</i>
3.1	Unit (Non-)Response	5
3.2	Item Non-Response	10
<i>4</i>	<i>Differences in the Choice of the Response Category</i>	<i>13</i>
<i>5</i>	<i>Ordinal Measures of Association</i>	<i>17</i>
<i>6</i>	<i>Differences on Aggregated Data</i>	<i>19</i>
<i>7</i>	<i>Conclusions</i>	<i>23</i>
	<i>References</i>	<i>26</i>

1 Introduction

Business tendency surveys (BTS) collect quickly available data from enterprises about their past and future development and their judgement of the current situation. This information is used by the participating enterprises and in economic policy in the decision process, and by researchers for business cycle analysis.

For many years, BTS using regular mailings suffered from an erosion of the response rate. To counter this problem, there are traditional methods such as a limitation of the number of questions, improvement of the design of the questionnaire, intensified recalls by mail or telephone, more attractive outputs and a time-reduction of the delivery in the results. Another possible way to stop the reduction in the response rate is to offer different modes of surveys. Therefore, an increasing number of survey institutions are introducing an Internet survey in addition to the regular mail survey.

There is a fast growing literature on the use and experience of Internet surveys. Vehovar V. et al. (2001) analysed the research literature in this field and found that response rates of Web surveys fall below those in mail surveys. But up to now, only a few research projects have analysed the effects of the new technique of the Internet survey on the response rate and the quality of the responses in a mixed mode. First studies indicate that mail surveys produce higher response rates than Internet surveys (e.g. D.A. Dillman et al., 2001a). Kaiser (2001) found that small firms participate less and firms in the software sector more by the Internet and that item non-response is more widespread among Internet survey participants.

In this paper we will show the differences between the mail and the Internet mode within a mixed mode survey in the KOF BTS. In a first step we analyse the difference in unit non-response. The mail respondent is confronted with questions which are presented on a single sheet which can be filled out anywhere and anytime. The Internet respondent is forced to fill out the questions on his or her computer. Therefore the danger of forgetting the response seems to be higher on the Internet than in mail surveys. Taking into account that the sample structure between these two modes could be quite different, the sample has to be divided according to structural elements. In particular, the calculation of the response rate has to be controlled for branches and cultural effects (division into three language regions).

Apart from the problem of unit non-response, there is the problem of item non-response which is not to be neglected. But there is no a priori hypothesis about the differences between the mail and the Internet survey mode.

If there is difference in the behaviour of firms between the two modes this could be reflected not only by the response rates but also by the selection of one of the three offered response categories. In the next step, we analyse the differences of the response behaviour between these two modes.

The results of business tendency surveys are rarely used on a micro level. The aggregation technique which is commonly used in international practice is the balance method. In the last step we will therefore analyse the effect of the differences of the two survey modes on the balance indicators.

2 The Data

The Swiss Institute for Business Cycle Research (KOF), Zurich, has conducted BTS as a panel survey for many years in different economic sectors of Switzerland. The BTS are addressed to firms and ask them – based on their units in Switzerland - about the economic activities, the judgement of the current business situation and short term expectations.

Until recently, KOF business tendency surveys were based on a paper questionnaire which was sent by regular mail. On the background of the demands of firms and the obvious advantages of an Internet survey (lower data collecting costs, no errors at digitising the responses, no delay in delivery), KOF successfully established in 1999 a pilot Internet survey in the sector of architects and engineers. During 2000, the final version of the Internet survey¹ was successively introduced in all monthly and quarterly surveys.

This study includes the manufacturing industry and retail trade conducted on a monthly, and construction, hotel/catering and architects/engineers on a quarterly frequency. The largest monthly data set on the Internet sample belongs to the manufacturing industry and includes responses from October 2000 to December 2001. Therefore, the analysis is limited to this time span.

At the time when the Internet survey was introduced the firms could choose to continue with the traditional paper questionnaire or switch to the Internet mode; thus the surveys are currently all conducted in a mixed mode. In 2001, 28% of the sample size participates by Internet and this percentage is slowly increasing.

The questions in BTS are mainly of a qualitative nature. Within this study, only the ordered qualitative questions consisting of trichotomous responses (e.g. higher, unchanged, lower) are analysed. The items are well known and used world-wide by many other survey conducting institutions.

Most analysis of BTS data is based on aggregated data because micro-data are normally not accessible. As the data collecting institution we dispose of these micro-data. This offers a direct way to measure the reaction and behaviour of firms, thus opening up an insight which is otherwise not possible.

¹ The application can be seen in: <http://survey.kof.ethz.ch/>

3 Non-Response Rates in Different Survey Modes

In panel surveys such as the BTS there are several sources of errors: coverage error, sampling error, measurement error and non-response error (Dillman, 2001b). The KOF BTS panels are structured according to the universe. But the subsamples for mail and for Internet do not fulfil this condition because firms could decide voluntarily to switch to the Internet mode. A comparison of the two modes therefore has to cope with errors of the total panel, differences because of unequal sample structure and differences because of different response behaviour in the two subsamples.

For our purpose we are interested only in differences due to unequal response behaviour. Fortunately, these differences are somewhat limited because both modes are self-administered, in contrast to interviewer-administered (Couper, 2001). Different response behaviour in the two modes can principally show its effects in different response rates and in different choices of response categories by the respondents. Unfortunately, not only unequal response behaviour influences the response rate in the two survey modes differently. There are effects arising from the unequal sample structure too. The errors in total panel are either equal in both samples or go with the unequal sample structure.

The response rate is probably the most important single factor to ensure a high quality of the BTS results. The central point is whether the unit and item response rate will suffer or will even profit from the introduction of the possibility for the participants to respond online to the BTS. Before analysing the behaviour of the firms participating in the Internet survey in comparison to those participating in the mail survey we look at the possible factors influencing the response rate. The factors not depending on the survey mode are the following:

- Number of questions
- Sensitivity of questions
- Simple or difficult questions
- Principal willingness to participate in BTS
- Reputation of the institution conducting the survey
- Actual business situation of the firm

We proceed under the assumption that these determinants of the response rate will be constant and that they do not influence the two modes differently. Nevertheless, the structure of the sample of the two survey modes are not identical. Therefore, the reaction of the participants to these elements could give two different results on the aggregated level.

The most important factors influencing the response rate, depending on the survey mode are

a) for regular mail

- The clear layout of the questionnaire
- A pre-paid return envelope
- The system of recalls of a mail survey
- The specific advantages and disadvantages of the mail mode itself (e.g. hard-copy as a reminder, delivery delay)

b) for Internet

- A simple and clear layout²
- An easy and quick access to the Internet survey
- A secure Internet message exchange
- The system of recalls for the Internet survey
- The specific advantages and disadvantages of the Internet mode itself (e.g. e-mails get lost in the huge number of incoming e-mails, quick transmission, problems with the computer systems, computer frustrations on the part of the respondent)

Changing one of the factors determining the response rate of only one mode has a direct effect on the comparison of the mode response rates. But under the period of observation there was no change of any of the mentioned elements on the side of the survey conducting institution.

Obviously, the conditions for a firm to answer the questions of the BTS on paper or on Internet are quite different and can distort the results of a mixed mode survey. The ideal methodology to analyse this problem would be to send both, paper and Internet questionnaires, to the Internet participants and examine if there is a different reaction to the questions read on paper or read in the Internet. But such an experiment is not feasible because we would risk unacceptable unit non-response in both modes. Therefore we cannot prove a difference or non-difference between the two survey modes. In the following chapters we try to give an impression of the magnitude and the relevance of such differences.

3.1 Unit (Non-)Response

To offer more than one mode in a survey is costly. An institution conducting surveys therefore introduces an additional survey mode only if it can expect an

² The designers of Web questionnaires face an unusual challenge. Instead of designing at the cutting edge of their evolving science, there is a need to hold back on the incorporation of advanced features (Dillman 2000, p. 375).

increase of the unit response rate, a lowering of the costs or saving time collecting the data. The latter can be assumed for sure and costs will be reduced at least after the introduction phase. Therefore a small reduction in the unit response rate will be offset by the other two advantages. But this reduction has to be limited, otherwise the survey results have a quality problem.

Table 1 gives an overview and a first impression about the unit non-response of BTS in different sectors. It provides quite a clear picture of how firms respond to the different KOF BTS and how they react on the Internet survey even if they selected this mode voluntarily.

The mail surveys of architects/engineers, construction and manufacturing industry reached the highest response rates with values between 67.9% and 77.1%. In the case of the mail survey of retail trade and hotel/catering the response rates range around 60%. One reason of this result could be that in the first group the respondents consist in general of persons who are more used to administrative work and to statistical instruments than in the second group. In addition, they might interpret the results in our reports more easily.

In the Internet survey the response rate is generally considerably lower than in the mail survey. The highest value on the Internet is achieved in the manufacturing industry (60.6%) but it is 7.3%-points below the response rate of the mail survey. Retail trade, hotel/catering and architects/engineers indicate a response rate in the Internet survey between 52.2% and 56.5%. The rate is strongly reduced for the survey of architects/engineers, it differs to the mail survey by 23.2%-points. But the biggest difference shows up in the construction survey which only reaches a response rate of 41.5% in the Internet mode.

An aspect not to be neglected: cultural differences can lead to different unit response rates. In Switzerland, the cultural differences can be identified by the three main languages – German, French, Italian. But it is not a priori clear if this element is independent from the survey mode or if on the contrary, preferences of a particular mode differ between the different language speaking regions.

The results in Table 1 indicate that the response rates in the three language regions of Switzerland do not differ systematically. In the manufacturing industry and in the case of architects/engineers the firms of the German speaking part of Switzerland are more reliable in responding regularly to our surveys. On the other hand, in construction it is the Italian questionnaire and in hotel/catering the German and French questionnaire which get the highest response rates. Additionally, in retail trade we do not see any significant difference between the language regions.

Taking the mail and the Internet survey together, hotel/catering gets the lowest response rate with 60.6%, followed by the manufacturing industry, construction and the retail trade; in the case of architects/engineers an overall value of 70.6% was reached over the period the Internet survey was disposable.

The size of the differences between mail and Internet can be influenced by different measures mainly with respect to mode-specific elements. But the difference is such that the surveys will suffer from a reduction of the response rates with the introduction of a supplementary Internet survey mode – at least in the short term.

Table 1: Sample size and response

(pooled data)

Surveys	Mail survey			Internet survey		
	Sample size	Response	Quota	Sample size	Response	Quota
Manufacturing industry Oct. 2000 to Dec. 2001	17666	11993	67.9	7973	4802	60.2
German questionnaire	10811	7770	71.9	6213	3717	59.8
French questionnaire	5474	3395	62.0	1393	825	59.2
Italian questionnaire	1381	828	60.0	367	260	70.8
Construction 2.Q.2001 to 4.Q.2001	3294	2455	74.5	1032	428	41.5
German questionnaire	1767	1365	77.2	788	324	41.1
French questionnaire	926	606	65.4	186	74	39.8
Italian questionnaire	601	484	80.5	58	30	51.7
Retail trade Jan. 2001 to Dec. 2001	6272	3722	59.3	2162	1129	52.2
German questionnaire	4515	2668	59.1	1749	917	52.4
French questionnaire	1648	988	60.0	341	157	46.0
Italian questionnaire	109	66	60.6	72	55	76.4
Hotel+catering 2.Q.2001 to 4.Q.2001	2899	1787	61.6	773	437	56.5
German questionnaire	1602	1019	63.6	572	326	57.0
French questionnaire	815	509	62.5	154	78	50.6
Italian questionnaire	482	259	53.7	47	33	70.2
Architects+engineers 3.Q.2000 to 4.Q.2001	3850	2969	77.1	2274	1226	53.9
German questionnaire	2997	2356	78.6	1993	1088	54.6
French questionnaire	823	591	71.8	245	113	46.1
Italian questionnaire	30	22	73.3	36	25	69.4

The overall picture can probably disguise greater problems in some subgroups because of unequal sample structure. In particular, there could be quite a different behaviour of the firms between branches, firm size and the export categories (percentage of exports in total sales of a firm). Going into further detail, we shall from now on limit our analysis to the manufacturing industry only. In this survey we have the highest number of waves (15) in the period under consideration and the heterogeneity is at least as big as in other sectors. The results of the manufacturing industry will give

therefore some hints of the disaggregated response situation of BTS in other sectors too.

Table 2 takes a closer look at the above-mentioned disaggregations in the manufacturing industry. Differentiating among nine branches, the results show big differences in the unit response rate. In regular mailing, "Pulp, paper, printing" reached a response rate of only 53.6% but "Textile, leather", "Wood, other non-metallic products" and "Other manufacturing industries" one of 74% or more. This unequal structure does not coincide with the results of the Internet survey. Here, "Other non-metallic products" gets by far the lowest response rate (39.8%), "Textile, leather" one as high as 54.2%. The response rates of the other branches are all quite close to the overall average of 60%.

Looking at the behaviour of the firms according to their size, in the mail survey the small firms (65.2%) have a somewhat lower response rate than the medium and big firms. In the Internet survey, this difference between the firm size is much more pronounced: The small ones are clearly below (53.5%), the big ones (71.0%) clearly above average. This confirms the experiences of Kaiser (2001).

In the third type of disaggregation, by the percentage of exports on sales, there is no systematic and relevant difference in response rates in the mail survey. But in the Internet survey firms with no export activities are much less reliable in answering by Internet (response rate of 55.9%). The other export-categories show a higher response behaviour.

Table 2: Sample size and response in manufacturing industry

Oct. 2000 to Dec. 2001	Mail survey			Internet Survey		
	Sample size	Response	Quota	Sample size	Response	Quota
Manufacturing industry						
Total	17666	11993	67.9	7973	4802	60.2
By branches						
Food, beverages, tobacco	1203	811	67.4	598	342	57.2
Textile, leather	882	654	74.1	271	147	54.2
Wood, other non-metallic prod.	2386	1780	74.6	748	454	60.7
Pulp, paper, printing	2657	1425	53.6	1184	669	56.5
Chemicals, plastic products	1737	1188	68.4	1065	671	63.0
Metal	3412	2402	70.4	1378	902	65.5
Machinery, transport	2230	1622	72.7	1196	760	63.5
Electrical and optical equipm.	2282	1460	64.0	1091	681	62.4
Other manuf. industries	877	651	74.2	442	176	39.8
By firm size						
1-49 employees	10177	6636	65.2	4100	2195	53.5
50-249 employees	5346	3844	71.9	2422	1558	64.3
250 and more employees	2143	1509	70.4	1451	1030	71.0
By export-categories						
Exports 0-4%	8733	5860	67.1	3466	1938	55.9
Exports 5-33%	3154	2165	68.6	1334	824	61.8
Exports 34-66%	2114	1541	72.9	1054	672	63.8
Exports 67-100%	3668	2403	65.5	2119	1349	63.7

In total, we notice a rather higher non-response rate for firms participating in the new Internet survey than for those still answering with the mail questionnaire in all of our surveys. The behaviour in the language regions of Switzerland has no clear pattern but shows an unexpected difference within the two survey modes. Within the firm size, there is a clear tendency for medium and big firms to respond more regularly on the Internet than small ones in both modes. Differentiated according to the percentage of exports on sales, the home-market oriented firms only register a below average response rate on Internet, the others do not differ significantly .

These results are in accordance with international experiences which have to cope with significantly lower unit non-response on mail than on Internet surveys too³. It shows clearly that there is a big need for research in the field of response/non-response with respect to Internet surveys. One of the main topics will be to find ways how the firms can be reminded to respond to the questionnaire. From contacts to participating firms we know that the respondents forget the Internet more easily than the mail survey. Therefore some measures have to be found for a more adequate recall system, e.g. a combination of e-mail and telephone recalls.

³ Bosnjak M. (2002), page 47.

3.2 Item Non-Response

Apart from the problem of unit non-response, which seems to be obvious, there could be a problem of item non-response too. The items and the wordings are identical in the mail and Internet survey modes, but the questions are presented in a quite different surrounding and even the layout of the questions differs quite considerably. Concerning the layout, we had a trade-off problem between a simple layout which suits for different operating systems and currently used releases of almost all Internet-browsers on one hand and a more sophisticated application with frames which makes a visually identical layout of the mail questionnaire possible on the other hand. We preferred the technically clear, simple and fast method which is in line with the proposals of Dillman (2000).

Our hypothesis is that the differences of the modes affect the item non-response only marginally because these two survey modes both get into contact to the respondent in an indirect way (in contrast to a telephone or face-to-face interview). In Table 3 we compare the item non-response over all 15 questions and all 15 periods. The general impression is that item non-response is indeed not a big problem in the KOF manufacturing industry BTS, but it isn't to be neglected either. The Internet survey realised on the level of total manufacturing industry an even slightly lower item non-response rate (13.6%) than the mail survey (16.1%).

But this average disguises considerable differences in the branches. There are some branches with a very low item non-response rate in both modes: A rate of 10% or lower results for "Textile, leather", "Chemicals, plastic products", "Machinery, transport" and "Electrical and optical equipment". "Wood, other non-metallic products" and "Pulp, paper, printing" achieved the highest values with more than 20% of item non-response; these values are shown in both modes.

These differences astonish and will hardly come from a different behaviour in different branches. The main reason could be the standardised questionnaire. Not all questions suit all firms. There are firms producing only according to orders and thus responding only based on order books but not on stocks of finished products. Other firms produce stocks in advance and sell the products when orders come in; they do not respond on the basis of the order books but only on incoming orders and on stocks of finished products.

Table 3: Item non-response**Manufacturing industry**

All Questions, all periods

Branches	Mail survey			Internet survey		
	All Items	Item Non-Response	Quota	All Items	Item Non-Response	Quota
Food, beverages, tobacco	11415	2033	17.8	5835	1129	19.3
Textile, leather	8310	877	10.6	3735	388	10.4
Wood, other non-metallic prod.	23475	4948	21.1	10065	2294	22.8
Pulp, paper, printing	18165	4742	26.1	12480	2835	22.7
Chemicals, plastic products	15915	1366	8.6	11895	972	8.2
Metal	33510	6075	18.1	16005	2145	13.4
Machinery, transport	21165	2235	10.6	14385	1248	8.7
Electrical and optical equipm.	19425	2023	10.4	12045	800	6.6
Other manuf. industries	9090	1516	16.7	3315	438	13.2
Total	160470	25815	16.1	89760	12249	13.6

Apart from this firm-specific effect, there is a question-specific effect too. BTS consist of three types of questions – past changes, judgements and expected changes. Each of them will show a different response behaviour by the firms: Past changes are known data and therefore easy to report and we expect a low non-response rate. This information could also be collected with quantitative statistics. Judgements are more of a subjective type and need some interpretation by the respondent. A somewhat higher non-response rate is expected. The third type, expected changes, belongs to the subjective type too, but has to take into consideration future developments. Changes in the market situation are in some branches very fast and in others rather slow. In the latter case a forecasting statement is easier to note than in the former. In general, the item non-response rates in the case of expectations are normally higher than on past developments.

Table 4 shows that this difference in item non-response based on the past compared to future-oriented questions is small in the BTS conducted by KOF. In mail questionnaires, incoming orders compared to the previous month, production compared to the previous month and compared to previous year and expected demand have almost no missing responses (non-response rate below 2%). Even the question on midterm expectations has a non-response rate of only 3%, which is surprising. On the other hand, the judgement with regard to order books for export reaches a non-response rate of almost 50% and the questions on stocks – changes and judgements – get values of over 30%. These are exactly the questions to which some firms cannot reply because of the reasons mentioned above.

The item non-response rate on Internet questionnaires gives a very similar picture. The questions about production and demand show the lowest rates again. But in the case of order books for exports and the different questions about stocks the non-response rates lie considerably below the rates of the mail survey.

Now we compare the two survey modes and look whether there is a statistically significant different behaviour on item non-response. We have in manufacturing industry 15 periods of parallel responses on mail and on the Internet. Therefore we conducted an unpaired Student's t-test on the mean values of the 15 quotas to be equal on each of the 15 questions.

As the last column of Table 4 indicates, the item non-response rate differs significantly between mail and the Internet in most questions. A p-value lower than 0.05 is taken as evidence to reject the null hypothesis of a zero coefficient and is marked by (*). In only two questions the t-test does not allow to reject the possibility of identical means of the item non-response quotas. Both questions are related to demand: Incoming orders compared to the previous month and to the same month of the previous year.

As a conclusion, questions suitable for all firms have a very low item non-response rate. But for these questions, the rate differs to a statistically significant extent between the two survey modes. This can result from the unequal sample structure or from an unequal response behaviour. In our view, the absolute value of the non-response rate in both survey modes is so low that even statistically significant differences are not of great economic relevance. In the other questions different response rates come mainly from the unequal sample structure which disguises a possible effect of a different behaviour between mail and Internet surveys.

Table 4: Item non-response, by questions

All periods, manufacturing industry

Questions	Mail survey			Internet survey			T-test on Quotas
	Items Asked	Item Non- Responses	Quota	Items Asked	Item Non- Responses	Quota	
Incoming orders, PP	10698	179	1.67	5984	74	1.24	1.90
Incoming orders, PY	10698	240	2.24	5984	108	1.80	1.76
Order books, PP	10698	1304	12.19	5984	613	10.24	4.99 *
Order books, J	10698	1428	13.35	5984	622	10.39	6.61 *
Order books for export, J	10698	5174	48.36	5984	2396	40.04	8.61 *
Production, PP	10698	3	0.03	5984	172	2.87	-12.80 *
Production, PY	10698	145	1.36	5984	142	2.37	-4.11 *
Stocks of interm. products, PP	10698	3495	32.67	5984	1554	25.97	11.90 *
Stocks of interm. products, J	10698	3645	34.07	5984	1590	26.57	13.10 *
Stocks of finished products, PP	10698	4370	40.85	5984	2022	33.79	11.90 *
Stocks of finished products, J	10698	4454	41.63	5984	2055	34.34	13.10 *
Expected incoming orders	10698	162	1.51	5984	127	2.12	-2.12 *
Planned production	10698	148	1.38	5984	131	2.19	-3.46 *
Planned purch. of interm. Products	10698	745	6.96	5984	520	8.69	-4.89 *
Expected midterm development	10698	323	3.02	5984	123	2.06	4.14 *

PP = compared to previous period

PY = compared to previous year

J = judgement

* probability of equal item non-response quota between internet and paper questionnaires smaller than 0.05

4 Differences in the Choice of the Response Category

A possible reaction of firms to alternative survey modes is not limited to unit or item non-response but could affect the choice of the response category too. This would have a significant draw-back for the survey results and its interpretation. Such behaviour is not expected but the importance of such a hypothesis makes it necessary to control for this factor.

The three response categories of the BTS were coded with 2, 3 or 4 respectively (number 1 leaving for some questions with a category “not existing”, e.g. at inventories or order books). The quantitative codes of the response categories offer the possibility to perform a t-test on the means of the two survey modes. A t-statistics value ≥ 0.05 indicates that there is a significant difference between responses of the participants in the Internet and in paper questionnaires.

Table 5 shows the results over all periods and all questions. The t-test on total manufacturing industry between the two samples does not reject the hypothesis of different respective means. Differentiated by branches, the picture is not uniform. In three branches – “Wood, other non-metallic products”, “Pulp, paper, printing” and “Chemicals, plastic products” – a significant difference in the respective means is rejected. In these branches, there is in general a comparable behaviour of the firms in

mail and Internet surveys. In all the other branches the results indicate a statistically significant difference in the response behaviour.

There are two ways to interpret these results: there is indeed a different behaviour of the firms in the choice of the response category between mail and Internet surveys or the unequal sample structure produces by the aggregation of all questions a result which disguises the identical behaviour. To eliminate the second possibility we calculate the t-test of the questions separately.

Table 5: Response differences between the two modes

Manufacturing industry

All questions, all periods

Branches	Mean		T-test	Prob .
	Mail	Internet		
Food, beverages, tobacco	2.988	2.9223	3.21	0.0013
Textile, leather	3.1077	3.0484	3.77	0.0002
Wood, other non-metallic prod.	3.1033	3.076	1.82	0.0694 *
Pulp, paper, printing	3.0355	3.0097	1.43	0.1538 *
Chemicals, plastic products	3.0429	3.0273	0.51	0.6131 *
Metal	3.0717	3.0361	3.60	0.0003
Machinery, transport	3.0133	3.0512	-5.12	0.0001
Electrical and optical equipm.	3.0668	3.0377	3.51	0.0005
Other manuf. industries	3.1055	3.0115	3.73	0.0002
Total	3.0587	3.0313	5.4	0.0001

* no significant difference on a 5% level between internet and paper questionnaires

Table 6 shows the results for three questions, each representing one type of BTS questions: Incoming orders compared to previous period, judgement of stocks of finished products and planned production. And indeed, the results of the t-tests change considerably if we restrict the sample to one question. In the case of the question on incoming orders compared to the previous period, only “Total manufacturing industry” and “Machinery, transport” prove unequal means on this question. For all other branches a significant difference of the respective means can be rejected.

The question on the judgement of stocks of finished products indicate for “Total manufacturing industry” a difference of the means in the two modes. The same is true for four branches, including again “Machinery, transport”. In the case of the other branches a significant difference of the respective means can be rejected. It is a clear cut between the two groups of branches. In branches producing more on stocks than on orders, a significant difference between the means can be rejected. But for branches which are more oriented to produce on orders the question about stocks of finished products is not so relevant and a difference between the two survey modes seems to come more from the sample structure than from different behaviour.

The question about planned production gives a similar picture as on the question about past incoming orders. Only for “Metal”, “Other manufacturing industries” and “Total manufacturing industries” can a significant difference of the respective means not be rejected. In all other branches, the t-test signals no significant difference in the means of the two survey modes.

In conclusion, we find clear indications that in the branches “Food, beverages, tobacco”, “Textile, leather”, “Wood, other non-metallic products” and “Pulp, paper, printing” the means of the quantified response categories do not differ. In the other branches the picture is not so clear. But particularly with regard to questions about stocks, differences are mainly between branches producing on orders rather than on stocks. This effect can result from an asymmetric sample size regarding the questions about stocks and not from different behaviour of the firms in selecting the response category.

Table 6: Response differences between the two modes, by questions

Manufacturing industry:

Question on incoming orders compared to previous period

Branches	Mean		T-test	Prob .
	Mail	Internet		
Food, beverages, tobacco	2.9153	2.849	1.26	0.2091 *
Textile, leather	3.0804	3.061	0.14	0.8865 *
Wood, other non-metallic prod.	3.1086	3.0318	1.82	0.0690 *
Pulp, paper, printing	2.9938	2.8878	0.82	0.4140 *
Chemicals, plastic products	2.9577	2.9727	1.25	0.2097 *
Metal	3.0866	2.9734	0.19	0.8465 *
Machinery, transport	3.0401	2.9689	-2.25	0.0246
Electrical and optical equipm.	3.1045	3.0038	1.66	0.0968 *
Other manuf. industries	3.0556	2.9862	0.92	0.3560 *
Total	3.1032	3.0785	1.97	0.0486

Question on stocks of finished products, judgement

Branches	Mean		T-test	Prob .
	Mail	Internet		
Food, beverages, tobacco	2.9888	2.9767	0.58	0.5587 *
Textile, leather	2.803	2.7904	0.29	0.7734 *
Wood, other non-metallic prod.	2.8867	2.9249	-1.16	0.2457 *
Pulp, paper, printing	2.8312	2.8197	0.32	0.7484 *
Chemicals, plastic products	2.8174	2.9037	-3.38	0.0007
Metal	2.9357	2.791	5.68	0.0001
Machinery, transport	2.8634	2.9307	-2.56	0.0107
Electrical and optical equipm.	2.9556	2.7713	6.31	0.0001
Other manuf. industries	2.8867	2.8112	1.49	0.1378 *
Total	2.8927	2.8621	3.01	0.0026

Question on planned production

Branches	Mean		T-test	Prob .
	Mail	Internet		
Food, beverages, tobacco	2.8779	2.8145	1.49	0.1358 *
Textile, leather	3.0584	2.9834	1.75	0.0806 *
Wood, other non-metallic prod.	3.0528	3.0091	1.46	0.1433 *
Pulp, paper, printing	2.927	2.9037	0.83	0.4075 *
Chemicals, plastic products	2.9156	2.9355	-0.66	0.5095 *
Metal	3.0519	2.9753	3.33	0.0009
Machinery, transport	2.9412	2.9607	-0.73	0.4657 *
Electrical and optical equipm.	3.0126	2.9722	1.38	0.1676 *
Other manuf. industries	2.9916	3.0995	-2.07	0.0387
Total	2.9899	0.0345	3.34	0.0008

5 Ordinal Measures of Association

Additional insight into the behaviour of the firms participating in the manufacturing survey by mail or by Internet give the measures of the degrees of association between two variables (Cliff, 1996). Four questions are particularly suited for this purpose: Incoming orders and production are asked not only with respect to the past but also to the future development. These questions do not only indicate the strength of the association but also the capability of the firms to predict the future development, differentiated by mode.

Unfortunately, the question about the past development compares only with the last period whereas the question about the future development asks about its movement in the next three months. We do not use an aggregation method to transform the monthly replies about the past development to a period of three months because every aggregation method has its draw-backs. To compare these two questions we selected the expectations formulated at the beginning of a quarter and put them into relation to the question about past changes two months later. e.g. the answers to the expectations questions about October 2000 were compared with the actual results realised in December 2000. To prevent statistical inferences we pooled data only on a three month step.

A direct measure of correspondence between expectations and realisations gives the Pearson correlation coefficient based on scores; it is applicable for ordinal data. The correlation is defined by

$$r = \frac{SS_{rc}}{\sqrt{SS_r SS_c}}, \text{ with}$$

$$SS_{rc} = \sum_i \sum_j n_{ij} (r_i - \bar{r})(c_j - \bar{c})$$

$$SS_r = \sum_i \sum_j n_{ij} (r_i - \bar{r})^2$$

$$SS_c = \sum_i \sum_j n_{ij} (c_j - \bar{c})^2$$

where r_i are the row scores and c_j are the column scores. The coefficient lies between +1 and -1, the value 0 indicates no relation between the two variables. The results presented in Table 7 for incoming orders show a positive relationship for "Total manufacturing industry" with a coefficient of 0.23 for the mail and 0.20 for the Internet survey. The asymptotic standard errors (ASE) with 0.019 and 0.025 lie far below the correlation coefficient and indicate a significant positive value in both modes, rejecting the hypothesis of a zero coefficient. This result holds for all the branches too, except for "Other manufacturing industries" in the Internet mode. The highest correlation coefficients were realised by "Wood, other non-metallic products" in the mail survey and by "Textile, leather" in the Internet survey.

Ordinal data of BTS typically contains information about tendencies. Therefore, measures of ordinal association are of interest which can reflect the degree of correspondence. There are different but similar measures of ordinal association. A seeming peculiarity of these measures is that they are based on pairs of scores rather than on the scores themselves (Hildebrand et al., 1977). This makes it possible to get information about the degree of the correlation and not only the direct correlation.

We choose two statistical measures of ordinal association. They differ in selecting the events which are counted as prediction errors. Goodman/Kruskal (1979) eliminate all ties before measuring the association:

$$\gamma = \frac{P(C) - P(D)}{P(C) + P(D)},$$

where $P(C)$ is the probability of concordant pairs, $P(D)$ is the probability of discordant pairs. Ties are paired observations on a variable with identical response and can therefore not give an ordinal information about the direction of the movement. The Goodman/Kruskal gamma coefficient as a measure for ordinal association states for “Total manufacturing industries” a clear positive relation for the mail (0.34) and for the Internet (0.29) mode. Again, the asymptotic standard errors are very low and therefore the coefficient is significantly different from zero. Only “Other manufacturing industries” do not have a significant gamma coefficient. This is exactly in line with the results of the Pearson correlation. The values of gamma in both modes are somewhat lower than for architects and engineers analysed earlier (Etter, 1997) which can be explained partly by the different time span for past and future questions in manufacturing industry.

The interpretation of the Goodman/Kruskal gamma coefficient is somewhat restricted because the probability is conditioned on the absence of ties. The second measure for ordinal association, Kendall’s T, includes all ties – $P(T)$ as the probability of ties – and is therefore lower or equal to the gamma coefficient:

$$\tau = \frac{P(C) - P(D)}{P(C) + P(D) + P(T)}$$

And indeed, all coefficients of Kendall’s T are lower than those of Goodman/Kruskal’s gamma. But all coefficients reject the hypothesis of no association within a 95% confidence bound too. The only exception is again “Other manufacturing industries” in the Internet mode. The poor result of this branch has to be seen on the very small bases of participating firms. Nevertheless, these overall results indicate a high capability of the firms to anticipate the future development in demand. This holds good for the mail as for the Internet survey.

Table 7: Expected and realised incoming orders

Manufacturing industry						
Branches	Mail survey		Gamma	ASE	Kendall's T	ASE
	Pearson Corr.	ASE				
Food, beverages, tobacco	0.0865	0.0715	0.1465	0.1059	0.0920	0.0669
Textile, leather	0.2441	0.0773	0.3871	0.1185	0.2234	0.0717
Wood, other non-metallic prod.	0.3381	0.0463	0.4823	0.0617	0.3146	0.0432
Pulp, paper, printing	0.1915	0.0540	0.2971	0.0813	0.1764	0.0495
Chemicals, plastic products	0.2372	0.0580	0.3591	0.0856	0.2180	0.0538
Metal	0.2186	0.0407	0.3312	0.0599	0.2011	0.0375
Machinery, transport	0.1799	0.0502	0.2703	0.0739	0.1647	0.0459
Electrical and optical equipm.	0.1947	0.0547	0.3057	0.0848	0.1762	0.0502
Other manuf. industries	0.2409	0.0813	0.3447	0.1130	0.2231	0.0754
Total	0.2291	0.0185	0.3397	0.0268	0.2093	0.0171
Internet survey						
	Internet survey		Gamma	ASE	Kendall's T	ASE
	Pearson Corr.	ASE				
Food, beverages, tobacco	0.2834	0.1004	0.4047	0.1356	0.2669	0.0932
Textile, leather	0.3327	0.1239	0.4531	0.1729	0.2920	0.1193
Wood, other non-metallic prod.	0.3171	0.0767	0.4410	0.0997	0.2962	0.0710
Pulp, paper, printing	0.2404	0.0636	0.3333	0.0870	0.2151	0.0579
Chemicals, plastic products	0.1764	0.0700	0.2503	0.1032	0.1552	0.0651
Metal	0.1802	0.0564	0.2497	0.0837	0.1510	0.0514
Machinery, transport	0.1414	0.0656	0.1989	0.0970	0.1237	0.0609
Electrical and optical equipm.	0.1379	0.0688	0.2083	0.1014	0.1274	0.0626
Other manuf. industries	0.0318	0.1448	0.0372	0.2075	0.0235	0.1314
Total	0.2049	0.0252	0.2938	0.0356	0.1857	0.0230

The second item in the BTS in manufacturing industry which is asked not only about expected and but on past development too is production. The results for this paired variables look similar to the incoming orders but the coefficients are generally higher. This is plausible because the answer on expected production is often based on existing internal plans. In contrast, the expected incoming orders depend totally on external conditions.

6 Differences on Aggregated Data

The results of business tendency surveys are rarely used on a micro level, only within behavioural research projects. Results are in most cases aggregated and condensed. The internationally common aggregation technique is the balance method which takes the difference between the percentages of the positive and the negative

responses. The answers on “no change” are neglected. In the calculation of the balances we used no firm specific or population weights which is equivalent to the fact that the response of each firm has the same weight.

In the last step we will therefore analyse the effect of the differences of the two survey modes on the balance indicators. We restrict our analysis again to the manufacturing industry because only in this survey we do have enough waves (15) to get relevant results.

Table 8 presents correlation coefficients between the balance values of each question calculated in the sample of the mail and of the Internet survey respectively. For the majority of the questions there is a very high correlation with $r > 0.9$, reaching the maximum of 0.98 on judgements of order books. Only the change and the judgement of stocks of intermediate products and the change of stocks of finished products signalise a much lower correlation. These results confirm our findings on micro-data that questions about stocks differ significantly stronger between the two survey modes than in the case of other questions.

A high correlation coefficient indicates only a similar movement but does not give information about the level of the indicator, expressed by the mean value. Therefore we additionally need a T-test on the equality of the mail and Internet means as a complement. It is computed under the null hypothesis of independent and identical normal distribution, with equal means and variances in each subgroup.

The picture of the results of the T-test on equality corresponds in most cases with the correlation coefficients. In thirteen out of fifteen questions the equality of the means of the two survey modes have a t-value of lower than two and therefore does not reject equality on a 95% level. Only for the question about the judgement of stocks of intermediate products, which showed in addition a low correlation coefficient, and for the question about the midterm expectations is the equality of the means rejected. The midterm expectations of the two survey modes therefore correlate strongly but not on the identical mean level.

Table 8: Difference of balances between the two modes, by questions

Manufacturing industry					
Variables	Correlation coefficient	Mean		T-test on equality	Prob.
		Mail	Internet		
Incoming orders, PP	0.937	0.56	-6.14	1.339	0.192
Incoming orders, PY	0.971	-6.78	-0.03	0.930	0.361
Order books, PP	0.924	-10.73	-11.12	0.091	0.928
Order books, J	0.980	-21.87	-17.26	0.859	0.398
Order books for export, J	0.956	-24.84	-22.26	0.508	0.616
Production, PP	0.956	-3.45	-1.46	0.363	0.720
Production, PY	0.975	0.59	8.24	1.144	0.262
Stocks of interm. products, PP	0.660	0.81	-2.42	1.434	0.163
Stocks of interm. products, J	0.545	9.29	13.28	3.538	0.001
Stocks of finished products, PP	0.337	1.19	-1.83	1.307	0.202
Stocks of finished products, J	0.915	10.51	13.03	1.356	0.186
Expected Incoming orders	0.945	0.56	4.59	0.754	0.457
Planned Production	0.935	2.05	5.84	0.656	0.517
Planned purch. of interm. products	0.952	-4.90	-3.42	0.281	0.781
Expected midterm development	0.925	-4.06	4.61	2.170	0.039

PP = compared to previous period

PY = compared to previous year

J = judgement

The statistical properties of the two survey mode balance indicators of the questions are confirmed by a visual interpretation in the charts. Figure 1 with regard to past changes of the incoming orders shows a very similar movement of the two indicators in the mail and the Internet survey sample. Moreover, the level of the two indicators differs only minimally. The situation with regard to the question on the judgement of stocks of intermediate products (Figure 2) is somewhat different. A low correlation coefficient and a significantly different mean value are well reproduced in the chart. In the last figure the congruent development of the question about expected incoming orders in the mail and the Internet survey are impressive.

Figure 1: Incoming orders, PP

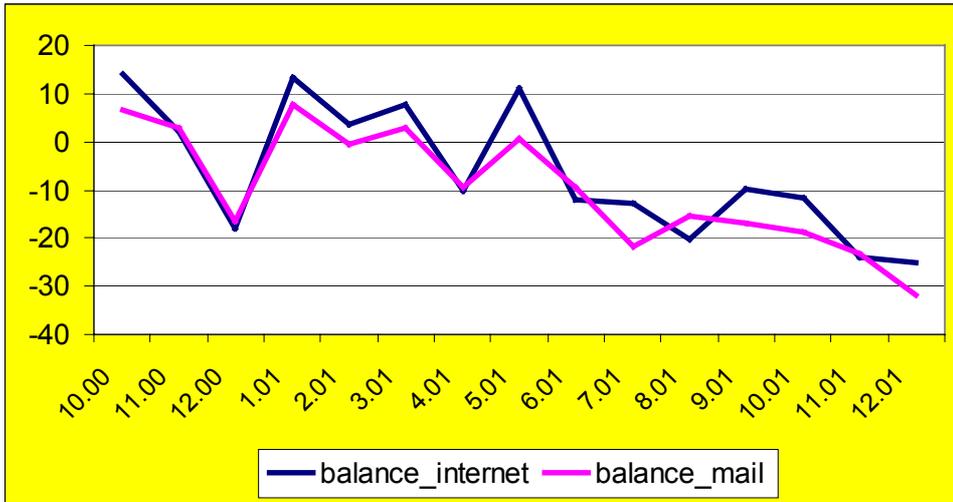


Figure 2: Stocks of intermediate products, J

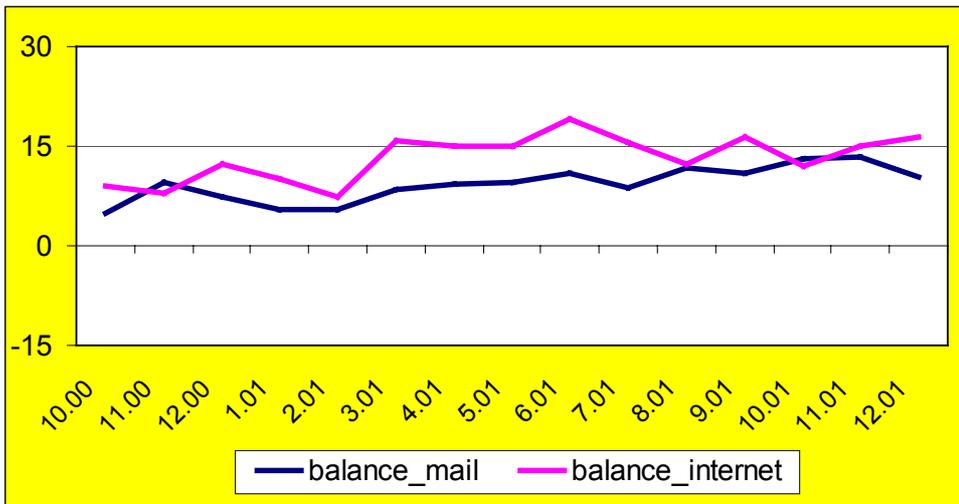
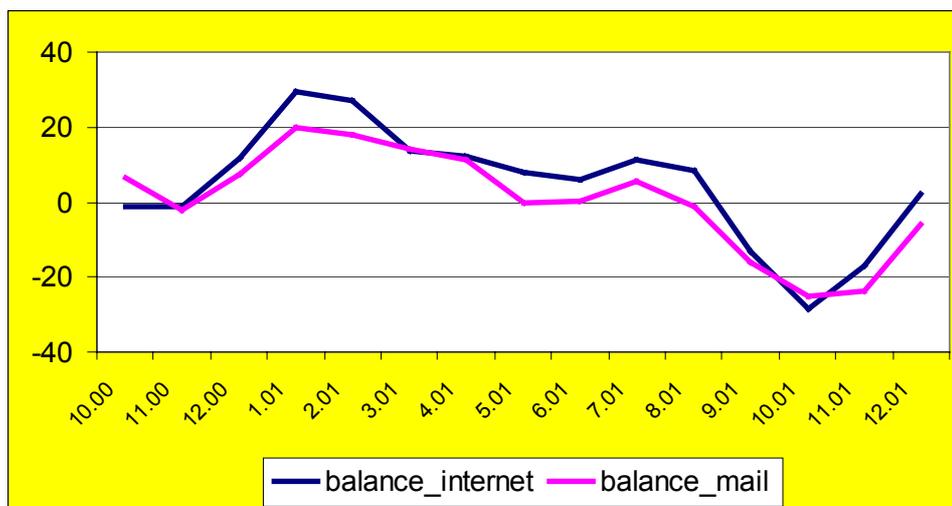


Figure 3: Expected incoming orders



7 Conclusions

The results of the comparison between the mail and the Internet survey mode in the BTS conducted by KOF show a big difference in the unit non-response rate. This holds not only for the manufacturing industry but for all BTS. Differentiating according to language regions, branches, size of firms and by the percentage of exports on sales does not change the picture. There is a strong need to improve the response situation particularly in the panel of the Internet survey mode. A central element of such measures will be an intensified recall system. Evidence exists that e-mail reminders in Internet surveys contribute up to one third of the final sample size. In addition, follow-ups contribute to a more representative sample since late respondents often differ from early respondents (Groves 2002, p. 236). Therefore the main consequence from this analysis is to intensify the efforts to increase the response rate from the Internet mode participants. Following this analysis we worked particularly on this problem. An e-mail recall and intensified telephone recalls had a significant effect on the response rate in the internet survey mode as shown in Table 9.

The item non-response rate varies between questions heavily. On the items suited for all firms (e.g. demand, production) the rate is between zero and three percentage which is very low. The differences between past and future questions of an item are minimal. The higher item non-response rates with respect to the other questions can be explained by the fact that these questions are about items not meaningful to all firms (order book, stocks). Comparing between the two survey modes, the item non-response rate is for most questions higher in the mail survey than in the Internet survey mode. This difference is in most cases significant. But these differences have to be seen in relation to the very high item response and therefore it seems not to be a major problem.

A possible reaction of firms to alternative survey modes is not limited to unit or item non-response but can affect the choice of the response categories too. We found that on disaggregated levels the means of the quantified response categories of a question do not differ in four branches. In the other branches the picture is non-uniform. In the majority of cases they do not differ. But particularly on questions to stocks, differences are mainly in branches producing on orders rather than on stocks. This effect can result from an asymmetric sample size regarding the questions on stocks and not from different behaviour on the part of the firms in selecting the response category.

The degree of association between past and future questions on identical items, differentiated according to survey mode, gives an additional hint about potentially different response behaviour. In the case of "Total manufacturing industry" the differences between the two modes can be neglected, but there are some quite considerable differences on the branch level which are difficult to interpret but seem to come from the unequal structure of the two subsamples.

The macro data in the form of the balance indicators show for most questions, except for three questions on stocks, a very high correlation between the mail and the Internet survey mode. Moreover, thirteen out of fifteen questions have equal means in the two survey modes. This is a very important result because balance indicators are extensively used in economic analysis and forecasts. The analysed period with 15 waves was very short; therefore the further development in the two modes has to be carefully observed and analysed again.

References

- Bosnjak M. (2002): (Non)Response bei Web-Befragungen – Auswahl, Erweiterung und empirische Prüfung eines handlungstheoretischen Modells zur Vorhersage und Erklärung des Partizipationsverhaltens bei Web-basierten Fragebogenuntersuchungen, Mannheim
- Cliff N. (1996): Ordinal Methods for Behavioral Data Analysis, Mahwah
- Couper M. P. (2001): Web Surveys: The Questionnaire Design Challenge, Ann Arbor
- Dillman D. A., Phelps G., Tortora R., Swift K., Kohrell J., Berck J. (2001a): Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response and the Internet. Paper presented at 2001 AAPOR Annual Conference, Montreal, Quebec, Canada, May 17-20.
- Dillman D. A., Bowker D. K. (2001b): The Web Questionnaire Challenge of Survey Methodologists, in: Reips U.-D., Bosnjak M. (eds.), Dimensions of Internet Science, Lengerich, pp. 159-178.
- Dillman D. A. (2000): Mail and Internet Surveys, The Tailored Design Method, New York.
- Etter R. (1997): Introduction and First Experiences with the Business Survey for Architects and Engineers, Working Paper KOF, Nr. 51, Zurich.
- Goodman L.E., Kruskal W.H. (1979): Measures of Association for Cross Classification, New York.
- Groves R.M., Dillman D.A., Eltinge J.L., Little R.J.A (2002): Survey Nonresponse, New York.
- Hildebrand D.K., Laing J.D., Rosenthal H. (1977): Prediction Analysis of Cross Classifications, New York.
- Kaiser U. (2001): Differences in Response Patterns in a Mixed Mode – Online/Paper & Pencil Business Survey, Discussion Paper No. 01-50, ZEW Centre of European Economic Research, Mannheim
- Vehovar V., Manfreda K. L., Batagelj Z. (2001): Errors in Web Surveys, E-Proceedings of the ISI 2001. The 53rd Session of the ISI, Seoul, 22-29, August 2001.