Journal of Official Statistics, Vol. 20, No. 2, 2004, pp. 341–355

Pre-printing Effects in Official Statistics: An Experimental Study

Anders Holmberg[1]

In surveys where respondents are contacted repeatedly, information from previous data collections may be used in the subsequent data collections. The responses then become dependent on the presentation and the quality of that information. Normally, the presented information is "historical" data concerning earlier reference periods, though relevant to a current reference period, and the respondents can verify and (if necessary) change it. The reasons for presenting historical data are: 1) it increases the efficiency of the data collection; 2) it can correct previous errors; 3) it reduces the response burden and sometimes, it is believed it reduces measurement errors and spurious response variability. The possible drawbacks are that it can conserve errors rather than correct them and it might lead to underreporting of changes from one period to another. Here, we focus on methodological issues of pre-printing "historical" values on self-administered (electronic or paper) questionnaires for business establishments. A planned experiment conducted in an ongoing large-scale survey indicates that on several aspects of data quality, questionnaires with pre-printed historical values outperform questionnaires without them. We present the main results of the experiment, as well as a general discussion of pre-printing experiences at Statistics Sweden.

Key words: Use of historical data; experimental design; measurement errors; response variability.

1. Introduction

When we collect data for official statistics, pre-printing (or its cousin dependent interviewing) is a frequently applied technique. The pre-printed information is "historical" data from earlier reference periods. Besides providing data for a new (up-to-date) reference period, respondents sometimes are asked to verify and (if necessary) change the pre-printed historical data. The type of data that is pre-printed can be administrative variables such as addresses and names, but here we focus on the case where more or less complex survey variables are pre-printed. The typical situation is data collections repeated over time involving common objects (e.g., surveys with panel designs or repeated business surveys with take-all strata). In this article we present the main results of a pre-printing experiment conducted within an on-going large-scale survey. The experiment was

[1] Department of Research and Development, Statistics Sweden, SE-70189 Örebro, Sweden. Email: anders.holmberg@scb.se

Acknowledgments: We gratefully acknowledge the work of Lars Lindam-Olsson and Kerstin Fredriksson at Statistics Sweden and of Marie Melkersson at the Employers' Association of Property Owners. Without their contribution to the data collections, this study had not been possible to conduct. Moreover, we wish to thank Joanne Pascale and three anonymous referees for their valuable advice and comments on an earlier draft.

designed to study certain effects of pre-printed self-administered questionnaires and it was a part of a larger project undertaken at Statistics Sweden.

When the requested information is repeated for the same object and when other survey circumstances are similar, most methodological papers focus on response variability (variation between repeated measurements) or response correlations. With few exceptions, these papers do not study the effects of using historical data, pre-printed on self-administered questionnaires or put forward in dependent interviews. However, the use of historical data in surveys is not new. Hansen, Hurwitz, and Pritzker (1964) discussed and questioned the effects of using historical data in surveys and since then various case studies have been conducted. References that provide a good subject overview are Mathiowetz and McGonagle (2000), who give a review of case studies made in household surveys, and Corti and Campanelli (1992), who summarize experiences of feeding forward earlier data in different panel surveys. Judging from these two papers and the references therein, most studies have been done on dependent interviewing and household surveys. Another paper that covers that area is that of O'Muircheartaigh (1986). He studied the U.S. Current Population Survey (CPS) and found smaller estimates of response variance (trial-to-trial variability) when historical data were used in interviews than when they were not.

However, the study presented in this article focuses on establishment populations and in that respect it has more similarities with the papers by Pafford (1986, 1988) and Stanley and Safer (1997). They did interesting embedded experiments with dependent interviewing in different surveys of farm populations. In the present case the data collection is performed with a pre-printed self-administered questionnaire.

The conclusions to be drawn from the experiment in this article can have a bearing on surveys where pre-printing of continuous variables is considered. A typical case is a repeated survey of businesses involving for the respondent neutral topics, e.g., bookkeeping or production data. However, the review at Statistics Sweden revealed that the majority of applications of pre-printing were on categorical variables. At the end we offer some short comments and a discussion of some experiences of pre-printing categorical variables.

2. Reasons for Using or Not Using Historical Data

The reasons for pre-printing historical data or not vary in accordance with the type of survey and the type of variables. On the basis of a review of the pre-printing practices regarding different statistical products at Statistics Sweden these reasons can be summarized as follows (Holmberg 2002):

- Respondent support: Pre-printing can support the respondent in a variety of ways, some examples of which are:
  - Reducing the response burden: frequently asking the respondent to fill in the same (hardly ever changing) statistics can seem unnecessary from a respondent perspective. Since respondents simply can verify that the pre-printed values still are valid or update them if necessary, pre-printing available data avoids this.
  - Questionnaire guidance, memory support, and anchoring: pre-printing can help the respondent to recognise what information is requested, where to put it in the

questionnaire, and whether the intended response is reasonable given the pre-printed information. Pre-printing can also help to keep track of events in time (e.g., whether an event has already been reported or not).

– Feedback purposes: sometimes pre-printing is a feedback appreciated by the respondents. For instance, in Statistics Sweden's monthly electricity statistics, an electronic questionnaire is used where the responding businesses are provided with the pre-printed values and time series graphics of their previous reports. When the pre-printing was introduced the respondent reactions to the feedback data (phone calls, mails etc) clearly indicated an increased interest in the collected statistics.

- Improved efficiency in the data collection: depending on the shape of the statistical production process, pre-printing can improve certain factors – it can offer easier communication in call-back interviews, a decreased risk regarding coding errors and less editing because of early-stage involvement of respondents.
- Reduction of measurement errors and improved data quality: pre-printing makes it possible for the respondent to check, react to, and correct the previously collected information. The respondent support (see above) also reduces measurement errors, since unwanted variation decreases with pre-printing. The unwanted variation can come from spurious reports of changes or large inconsistencies in data over time.

The main reasons for **not** pre-printing historical data are:

- Risk of bias due to underreporting of changes and conservation of errors: edits, follow-ups, and evaluations have detected phenomena such as respondents copying a pre-printed value without examining whether there has been a change and respondents who are unwilling to deviate too much from previously reported pre-printed values. For variables whose values with certainty change between reference times, an error most certainly occurs when a response is an exact repetition of a pre-printed value (which normally refers to a previous reference time). The respondent may find the pre-printed information exact or close enough and is freed of some of the response burden by using it. (Krosnick (1991) used the term "satisficing" when a respondent chooses a cognitively easier way to respond.) Unclear or skipped instructions are other reasons why pre-printed data incorrectly survives from one reference period to the next. The extent and total effects of these errors are often unknown, and especially for categorical variables that seldom change value over time they can be difficult to detect.
- Loss of confidence and goodwill: if the pre-printed data is of poor quality or if it is unrecognised by the respondent, there is a risk that the co-operation of the respondent will be negatively affected.
- Disclosure risk: even with good safety routines, there is always a risk that the pre-printed data will be disclosed to outsiders.

In addition to the reasons above, economic considerations play an important role. Whether pre-printing is seen as more expensive or not depends on a lot of factors, e.g., the type of survey, its size, the environment and tools used in the survey process, the actual effects of

the pre-printing and of course which quality aspects and cost functions are used for evaluations. Economic aspects of pre-printing are not treated in this article.

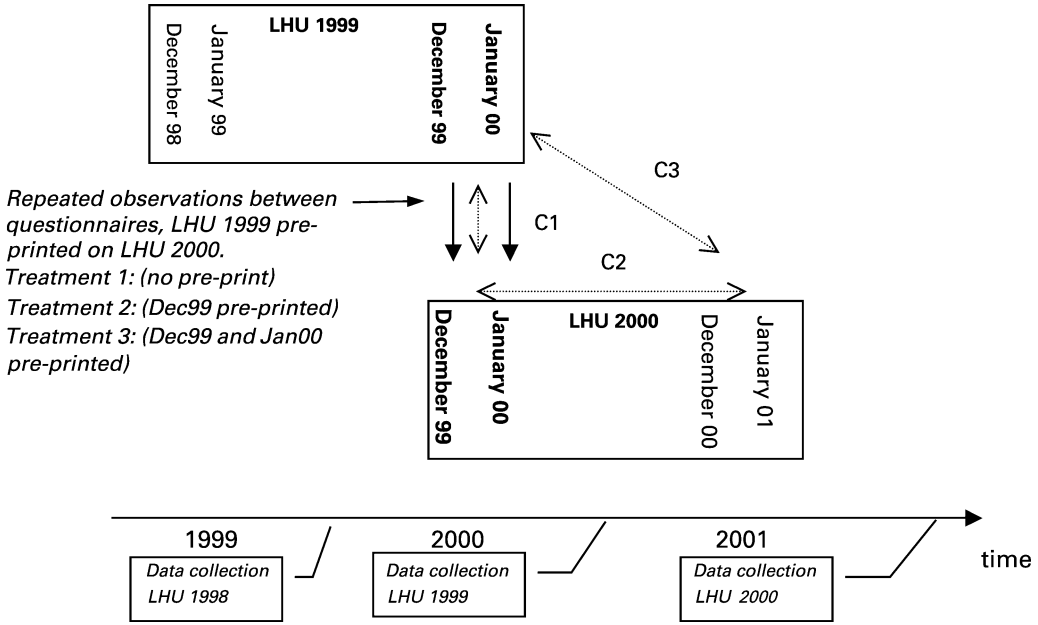
3. Background of the Pre-printing Experiment

The Survey of Building Rental Units (in Swedish Lokalhyresundersökningen (LHU)) collects information about total monthly rental incomes and rented area in square metres for commercial rentals of premises. Industrial premises, shops, and offices are the main categorization of the premises and the data is used to produce a Fisher price index for building rentals, which is used in the National Accounts statistics. LHU is conducted once a year and independent probability samples of building rental units are drawn every year. However, the number of building units that are the same in two successive samples is considerable owing to the sampling design (stratified sampling with x-optimal allocation (Särndal, Swensson, and Wretman 1991, p.107)). For these same building units it is possible to pre-print the reported data from the previous year. In the survey of 2001 (henceforth named LHU 2000 after the reference year) a planned experiment was conducted to study the effects of pre-printing.

The data collected in one LHU questionnaire stretches over a period of 14 months. The first two of these months overlap with the last two months of the LHU questionnaire from the year before. Hence, for the overlapping months there is a repeated measurement for the building units that are included in two successive years. Figure 1 gives a schematic idea of how this works for a building unit included in both LHU 1999 and LHU 2000.

As illustrated in Figure 1, the data collection for LHU 2000 was done at the beginning of 2001. The interesting reference months are December 99, January 00, December 00 and January 01, which are used for the Fisher index. If a building unit in LHU 2000 sample reported to LHU 1999, data of the overlapping reference months (i.e., December 99 and January 00) are available for pre-printing. If the unit then responds to LHU 2000, we will have repeated observations over two surveys and it is possible to estimate the response variability. Clearly, if no measurement errors are present, data for December 99 and January 00 should be identical in LHU 1999 and LHU 2000. The first comparison in the present study (indicated by the C1 arrow in Figure 1) considers the response variability with or without pre-printing.

To provide an insight into why pre-printing was tested for the LHU, there needs to be mentioned an analysis made of earlier LHU data (the 1997 and 1998 data matched on the building unit level). A simple comparison of responses indicated that there were problems of high response variability between two LHU surveys for the overlapping reference months. Furthermore, there was a within-questionnaire tendency for data for the two overlapping reference months to be similar to data of the latest two reference months (for instance, a large number of building units reported no changes in rental price between January and January). Since one of the objectives of the LHU is to measure levels of change, this was unacceptable. We suspected that, at the time of the data collection, it is easier for the respondent to give accurate data for the latest two reference months and that the two overlapping (earliest) reference months are too far back in time (almost 1.5 years when the data collection is done).



Note: Comparison 1 (C1): Compares the pre-printing treatments with respect to response variability. The comparison is made on repeated measurements between questionnaires.
Comparison 2 (C2): Compares the pre-printing treatments, effect on other reference periods, within the LHU 2000 questionnaire.
Comparison 3 (C3): Compares the pre-printing treatments' (indirect) effect on other reference periods; comparisons are made between questionnaires.

Fig. 1. Two successive LHU surveys and the comparisons of the experiment

Because of this, and because of a more general as well as vague underlying response burden argument, we considered pre-printing data from LHU 1999 in the questionnaires of LHU 2000. With LHU 1999 values printed nearby, the respondents were asked to fill in the corresponding reference months in LHU 2000. (Note that the values for December 99 and January 00 of LHU 2000 are needed for the Fisher index, which is estimated using data from one survey only. Old values for the units not responding or not included in LHU 1999 are not available.) Among the respondents who got pre-printed questionnaires there were surprisingly few who gave blank answers in the LHU 2000. These answers are handled as nonresponse in this study. However, in the editing work of the survey most of them turned out to be verifiers of the pre-printed values, and some respondents thought they did not have to fill in new values since they agreed with the pre-printed values.

The effect we hoped to achieve with pre-printing was to decrease the response variability of the overlapping reference months and thereby decrease the measurement errors and increase the between-questionnaire consistency of the units taking part in two consecutive LHUs. However, since the pre-printing of previously reported values also may affect the data quality of the latest reference months, the experiment was designed to also study pre-printing effects on (i) the size and frequency of outlier values before editing, (ii) the number of pre-printing recurrences and changed values between reference months, and (iii) estimates of yearly changes. To investigate these aspects the following experimental plan was used.

3.1. Experimental plan and data collection

The plan to obtain necessary data was as follows. From LHU 1999, 1,590 building units were identified as being the same as in the sample drawn for the LHU 2000 data collection. 1,115 of these 1,590 units had reported both rental price and rental area and for at least one of the categories shops, offices or industrial premises. The number of building unit owners (often, but not always, identical to the number of respondents) is less than this since several building units have the same owner. For policy reasons it was inappropriate to let one and the same unit owner get both pre-printed and not pre-printed questionnaires. To handle this, and to control possible variation due to owner size, three blocks were constructed. Block A comprised building units owned by owners with a large number of units and a high total rental income, Block B was a middle-sized group and Block C comprised building units owned by owners with only one building unit. In order not to get too unbalanced a design the units in each block were ordered into three relatively homogeneous groups with respect to owner size. Three different treatments were then randomly assigned among the units within each block. The three treatments were (1) no pre-printing, (2) pre-printing of rental price and rental area for December 99 only, and (3) pre-printing of rental price and rental area for both December 99 and January 00.

Treatment (2) (pre-printing December 99 only) was included in the experiment for two reasons: firstly, since December and January represent two calendar years, changes in rents and contracts might often occur between those two months; secondly, by pre-printing only one of the overlapping reference months it was possible to make additional comparisons of the pre-printing properties.

Table 1. Number of building units with office premises by block and treatment. n = Number of units in the experimental design plan (chosen from LHU 1999). n = Number of responding units matching between LHU 1999 and LHU 2000*

	Treatment								
	(1) No pre-printing			(2) Dec 99 pre-printed			(3) Dec 99 Jan 00 pre-printed		
	*n**	*n*	%	*n**	*n*	%	*n**	*n*	%
Block A	31	48	65	41	52	79	49	54	91
Block B	43	52	83	33	39	85	28	36	78
Block C	126	170	74	142	194	73	146	187	78
Total	200	270	74	216	285	76	223	277	81

In questionnaires with Treatments (2) and (3), data was pre-printed for all types of premises although only shops, offices and industrial premises were of interest in this study. Analyses were done for these three types of premises. Regardless of type of premises the conclusions of the analyses are similar, and for reasons of space this article mainly concentrates on the results for office premises. Offices are the most common type of premises, and only 283 of the 1,115 building units did not have office premises (666 building units did not have shops and 778 did not have industrial premises). Table 1 shows the number of building units in each treatment cell for offices, where n is the number in the design and n^* is the number that was available for the analysis at the data collection deadline.

As can be seen in Table 1, the proportion of available units varies between 65% and 91% of the units in the design plan. The main reason for missing data is unit nonresponse but changes in type of premises (from offices in LHU 1999 to some other type of premises in LHU 2000) is another common reason. (There was no clear evidence that pre-printing affected the unit nonresponse.)

3.2. Study variables and statistical models

The main variables in LHU are total monthly rental income and rented area in square metres. For every type of premises we use the ratio of total monthly income and total rented area as our basic variable, i.e., let Z_{tq_k} be the total monthly income per square meter for building unit k; t is a reference month indicator, where $t = 1$ if the reference month is January 00 and $t = 0$ if the reference month is December 99; q is an indicator of the questionnaire (the LHU from which data origins) with $q = 1$ if data come from LHU 2000 and $q = 0$ when data come from LHU 1999.

To measure discrepancies between repeated measurements and response variability (see arrow C1 in Figure 1) we compute the ratio $Y_{t_k} = Z_{t1_k}/Z_{t0_k}$ and the logarithm $Y^*_{t_k} = \ln Y_{t_k}$, for every unit k and the two overlapping reference months. By means of an ANOVA model (given in Equation 1) we can investigate if pre-printing has any effect on response variability.

$$Y^*_{ijk} = \mu + B_i + T_j + BT_{ij} + \varepsilon_{k(i,j)} \tag{1}$$

where

Y^*_{ijk} = Deviation in rental income/m^2 between LHU 2000 and LHU 1999 for the k:th building unit in the i:th block and the j:th treatment. (log-scale)

B_i = Effect of the i:th block. ($i = A, B, C$) (fixed effect)

T_j = Effect of the j:th treatment ($j = 1, 2, 3$) (fixed effect)

BT_{ij} = Interaction effect of the j:th treatment and the i:th block.

$\varepsilon_{k(i,j)}$ = Random effect in the (i,j):th cell

The randomised block model in Equation 1 is a basic model that was used for all types of premises and for both the December 99 and January 00 reference months separately. Note that Treatment 2 is a pre-printing treatment when the response variable refers to December 99, but is not a pre-printing treatment when the response variable refers to January 00. However, in the latter case one might guess that the pre-printed values of December 99 were helpful to the respondent when he/she filled in the values of January 00.

To study pre-printing effects on the properties of outlier values, pre-printing recurrences and estimates of yearly change traditional t-tests and chi square tests were used.

4. Statistical Analysis and Results

4.1. Data considerations

To fulfil the objectives of the study we chose to do all analysis on unedited data. A drawback then is that certain model assumptions in the analysis may become less valid. Preliminary screenings and model diagnostics revealed the presence of outliers and nonnormality. The severity of these violations of assumptions (with respect to changing the general test conclusions) was checked by also performing alternative nonparametric analysis. Since the conclusions from nonparametric and parametric analysis coincided, the results would appear to be fairly robust against these violations.

4.2. Pre-printing effects on response variability

Table 2 shows the ANOVA results (based on sum of squares of type III in SAS 8.1) when we fit Model (1) to the data for December 1999 and January 2000, respectively.

Table 2. ANOVA results for office premises: December 99 and January 00 ($n^* = 639$).

Source	Degrees of freedom	Reference month	
		December 99 p-values	January 00
Model	8	0.02*	0.07
Block	2	0.55	0.74
Treatment	2	0.03*	0.01*
Interaction	4	0.06	0.15
Error	630		

* Significant on $\alpha = .05$ level

As can be seen, there is no benefit from any block effect but there is a treatment effect for both December and January. (The nonsignificant block effect means that the blocking did not fulfil its purpose of reducing systematic variation due to owner size. For the analysis, however, the effect is only that degrees of freedom are lost in the model test and that the precision of our treatment means is unimproved. In Sections 4.3 and 4.4, the blocking is disregarded in the presentation.) The comparison of treatment means showed that the biggest differences exist between Treatment 1 and Treatment 3 (p-values of mean difference tests were 0.016 for December and 0.009 for January). For December data there was also a large difference between the means for Treatment 1 and Treatment 2 (p-value 0.069). For both December and January the means of the pre-printing treatments were lower than the means of Treatment 1.

Without pre-printing the data quality is in fact very poor. The (geometric) mean of Treatment 1 was significantly >1 both for December (1.06) and January (1.07), and 95% confidence limits were (1.02; 1.09) for December and (1.02; 1.11) for January. (If the mean discussed here equals 1, it implies that the respondents on average give the same response in the two LHUs, i.e., there is less variability in repeated measurement. The mean computed is the antilogarithm of the treatment mean, e.g., for December and Treatment 1, $\bar{Y}_1^*$, we have $Exp(\bar{Y}_1^*) = 1.06$.) This indicates that the responses on average were 2%-11% higher in LHU 2000 than in LHU 1999. Obviously, the response variation between repeated measurements is very high without pre-printing, and this is unacceptable in a survey whose aim is to monitor changes of a few percent. The response variability was lower for Treatment 2 (with a mean of 1.00 for December and 1.03 for January) and Treatment 3 (with a mean of 0.99 for December and 0.98 for January). All confidence intervals of the means for Treatments 2 and 3 covered 1.0 (although the mean of Treatment 2 was suspiciously high for the January data).

4.3. Pre-printing effects on the presence and size of outliers

Smaller response variability is one indicator that pre-printing decreases the measurement errors. Another indicator is the presence and properties of outliers. If we compare data from our treatments, we note that pre-printing gives us both fewer and smaller extreme values. Treatment by treatment, Table 3 shows the minimum deviation, the maximum deviation and the proportion of units where the response in LHU 2000 deviates by more than 25% from the response in LHU 1999.

The difference between pre-printing and not pre-printing is clear in Table 3. (Note that it contains December values, i.e., Treatment 2 is a pre-printing treatment.) For all types of premises, Treatment 1 always has the largest absolute extreme values. The worst was observed for an industrial unit, with a 1,016% higher value in LHU 2000 than in LHU 1999. From a cognitive viewpoint the presence of outliers should perhaps be interpreted a bit differently for the treatments. When an outlier value turns up for the pre-printed treatments, it is put there although the respondent is aware of the previously reported value. It can therefore be a correction of the previous value. For the not pre-printing treatment the respondent can be unaware that the value he/she reports deviates markedly from the reported value of the previous year.

Table 3. Max. and Min. values for $100(Y_{(t=0)k}-1)$, and the proportion of building units with $|100(Y_{(t=0)_k}-1)| > 25$, by treatment and type of premises. (December 99 values)

Type of Premises		Treatment		
		1	2	3
Shops ($n^* = 329$)	Minimum	−87	−29	−22
	Maximum	497	77	292
	Proportion of units with a deviating value >25%	14.3	4.2	2.0
Industries ($n^* = 254$)	Minimum	−92	−67	−45
	Maximum	1,016	31	223
	Proportion of units with a deviating value >25%	10.6	4.8	9.4
Offices ($n^* = 639$)	Minimum	−48	−75	−71
	Maximum	783	119	275
	Proportion of units with a deviating value >25%	14.0	5.5	6.7

The proportion of units with highly deviating values is also higher for Treatment 1 (over 10% of the units for all the premises). A table similar to Table 3 for January data had the same pattern.

4.4. Pre-printing recurrences and effects on other reference months

One concern using pre-printing was a high frequency of recurrences and that changes would be underestimated. In this application it is actually the other way around. The pre-printing treatments yield more changed values from one reference period to the next. In LHU 2000, we compared the responses of rental income for the reference months covering a 12-month period (refer to arrow C2 in Figure 1). Table 4 shows the numbers of units that had lower, the same and higher values of rental income, respectively, in December 00 compared to December 99.

Table 4. Number of units with office premises reporting different values of rental income in December 00 as compared to December 99, by treatment and direction of change. (%)

Direction of change	Treatment			Total number of units
	1	2	3	
Decrease	38	62	35	198
lower value reported for Dec-00	(19.2)	(28.7)	(15.8)	
Unchanged	60	38	51	132
Same value reported	(30.3)	(17.6)	(23.1)	
Increase	100	116	135	368
higher value reported for Dec-00	(50.5)	(53.7)	(61.1)	
Total	198	216	221	635
	(100)	(100)	(100)	

Note: Observe that the comparison now is between December 1999 and December 2000 and that 4 building units that had offices in December 1999 had changed to a different type of premises in December 2000.

The difference between the treatments is statistically significant (Chi square statistic is 18.5 with four d.f.) Treatment 1 has the largest percentage units reporting the same value for December 99 and December 00. This does not support the concern that pre-printing would mean more unchanged values. On the contrary there are more reporting of the same values without pre-printing. However, we guess that the large number of "same-value reporting" without pre-printing is due to respondent's practical problems of giving correct information for the oldest reference period.

The comparison above just counts the number of units giving different values for two reference periods. It does not tell us whether pre-printing would affect our estimates of yearly change. Therefore, with the null hypothesis of no differences between treatments, we tested the differences in the mean yearly rate of change with the following statistic:

$$\hat{D} = \left(\frac{1}{n_j^*} \sum\nolimits_{k=1}^{n_j^*} \frac{Z_{t'_{1jk}}}{Z_{t_{1jk}}} - \frac{1}{n_{j'}^*} \sum\nolimits_{k=1}^{n_{j'}^*} \frac{Z_{t'_{1j'k}}}{Z_{t_{1j'k}}} \right) \tag{2}$$

Here t and t' are either reference months December 99 and December 00 or reference months January 00 and January 01, and j and j' are pre-printed and not pre-printed treatments, respectively. Hence, positive values of $\hat{D}$ indicate a higher mean rate of change for pre-printing treatments.

None of the tests were statistically significant. However, the sign of $\hat{D}$ was positive for every pairwise comparison. This indicates that pre-printing would lead to higher estimates of yearly change than no pre-printing would. Together with results from analyses such as the one in Table 4, this observation suggests that the risk of underestimation of change because of pre-printing is low in this survey, at least compared to the alternative of not using pre-printing.

In practice, values from the foregoing survey are not used in the estimation in the LHU. The survey needs values for the overlapping months for all units (not only the units whose questionnaires can be pre-printed). Nevertheless, from a pre-printing perspective comparing the treatment means of yearly rates of change across questionnaires can be of interest (see arrow C3 in Figure 1).

If we in the denominators of $\hat{D}$ use the values from LHU 1999, i.e., replace $Z_{t_{1jk}}$ and $Z_{t_{1j'k}}$ by $Z_{t_{0jk}}$ and $Z_{t_{0j'k}}$ in Equation (2), then the signs of the $\hat{D}$ statistic become negative. Hence, if we measure the first reference month with values taken from the previous survey, it seems as if pre-printing has a slight dampening effect on the mean of the yearly rate of change.

In sum: when we compute the yearly rate of change within LHU 2000 (C2), the pre-printing treatments give higher average values, but when we compute across questionnaires (C3, using the LHU 1999 values for the overlapping months), these pre-printing treatments give smaller estimates. These patterns were the same across all types of premises and regardless of whether it was a question of December to December changes or January to January changes. A possible explanation of this is that pre-printing affects both the overlapping (pre-printed) reference months and the latest reference months. The effects may be in the same direction but with different strength. Without pre-printing, the measurement of the overlapping months tends to be biased towards the values of the latest reference months (which probably are easier for the respondents to give).

With pre-printing, this bias is smaller since the respondents are helped by pre-printed values from the previous survey. (In the previous years' surveys those values were probably easier to give (see data collection time in Figure 1).) However, the rate of change comparisons using LHU 1999 data suggests that pre-printing also may affect the measurements of the latest reference months towards the pre-printed values. Whether this is good or bad is hard to tell, since we do not have the true values and it depends on the quality of the pre-print. Careful evaluation studies have to be performed to compare the magnitudes of these two possible effects.

4.5. Further comments and results

There were no resources for doing a systematic follow-up among the respondents. Therefore the only knowledge we have of respondents' reactions to pre-printing comes from communication during the data collection and editing process. In that communication there were only favourable opinions regarding the pre-printed questionnaires, some respondents mentioning their satisfaction with the possibility of correcting their own previously reported pre-printed values.

The major differences were between Treatment 1 and Treatment 3. Treatment 2 behaved the same as Treatment 3 for December data. For January data, when it is a non-pre-printing treatment, it performed better than Treatment 1, i.e., with less response variability and fewer and smaller outlier values. There were indications that the respondents to some extent used the pre-printed December values to fill in the January values. However, they did not just copy the December values on to January. The correlation of rental income/m^2 for December 99 and January 00 was 0.81 for Treatment 2, whereas the corresponding correlations for Treatment 1 and Treatment 3 were higher. However, since there was no clear evidence that Treatment 2 was better than Treatment 3, and since Treatment 3 was more practical to implement, the LHU now pre-prints both December and January values for the units where possible.

5. Discussion

5.1. Some conclusions from the experiment

The conclusions from the experiment were in favour of pre-printing. There are clear indications that pre-printing improves the data quality. The response variation of the overlapping reference months is significantly decreased with pre-printing. There are indications that there will be smaller measurement errors and that less editing work will be required with pre-printing. (In a sense, pre-printing moves some editing work from the data collector to the respondents and they are in many cases in a better position to do it.) None of the possible drawbacks were discovered. The same value for two different months was actually more common without pre-printing. The fear that pre-printing would lead to underestimation of change was not justified, and last but not least pre-printing helped the respondent to fill in the questionnaire. It should be recalled that the overlapping reference months are almost 1.5 years back at the time of the data collection. If that data is hard to get, seeing the previously reported values (collected and reported closer to the reference month) should be helpful.

It should be mentioned that to fully assess the effects of pre-printing on measurement error, one would have to look at the responses from a cognitive perspective. A multi-disciplinary study design also involving studies of the behaviour of the respondents as well as on-site area measurements of the building units and evaluations of the records that the respondents use when they complete the questionnaire would have been ideal. With such a study we would also be able to assess the bias. Respondent policies, the production time of regular statistics and budgetary constraints at the statistical office are some reasons why this was not feasible in the LHU case.

However, the measurement situation in the LHU is not the common situation for pre-printing practices. The LHU includes both a repeated and a new measurement (the overlapping months and the December 00 and January 01 reference months, respectively). Usually when pre-printing is applied it is only the new measurement that is interesting. Nevertheless, if a repeated measurement is important, as in the LHU, and if a previous measurement was done at a more favourable point in time, it seems natural to use the first measurement to assure good quality in the next. Therefore it is not surprising that pre-printing works well for repeated measurements. The results concerning pre-printing effects on the quality of the new measurements are more vague. No statistically significant effects were found other than that it was more common that the same values were reported without pre-printing. A tendency for lower rates of change when comparisions were made with data from LHU 1999 (C3), and higher rates of change when comparisons were made within LHU 2000 (C2). None of these results can with certainty be interpreted as positive or negative for pre-printing, which of course is interesting since it does not rule out pre-printing as being a bad practice.

5.2. Other issues and experiences of pre-printing

In Sweden it is more common to pre-print categorical variables. The study variables in the LHU experiment are continuous and the discovered positive properties of pre-printing may be related to that. With continuous variables it is easier to practically study response variation and analyse effects. For a hardly ever-changing categorical variable this can be more difficult. Often there is a methodological problem in the analysis of pre-printed data. Without pre-printing we do not want too many changed values over time, since that is an indication that many of the changes are spurious. With pre-printing, on the other hand, we do not want too many unchanged values, since that may be a sign of respondents choosing a simple way to respond or not noticing that the values should be changed.

The study in the LHU was done as a part of a bigger review project at Statistics Sweden. The study is of a rare kind, since it is seldom possible to do large-scale experiments within an ongoing (traditional) survey. Although pre-printing is frequently used in data collections, there are few methodological studies of its properties. Nevertheless, within the project we also performed an experiment on categorical variables. This was done in co-operation with The Employers' Association of Property Owners and made on their annual wage survey, which uses an electronic questionnaire (for details of the experiment see Holmberg 2002). That experiment was designed to study the effects of ending the pre-printing of some categorical variables. It was thought that the pre-printed variables were not updated frequently enough and that this led to a lot of recalculations and editing,

since they were connected to other wage-related variables. We performed an analysis of the change and update frequencies, which showed that pre-printed questionnaires were updated just as often as those without pre-printing. Hence that study did not support the idea that pre-printing conserves outdated values to a larger extent.

The Swedish experience is that the response burden argument often prevails when pre-printing is decided upon. The experimental studies show that a data quality argument for pre-printing also can be strong, especially in cases where there are problems with a large response variation such as in the LHU. Moreover, in our review we did not find any cases where the possible negative properties of pre-printing had any significant effect on macro level statistics. Even though it may be feared that pre-printing will influence the data in such a manner that estimates will be biased, we still suggest that it is worth looking at. Studies indicate that unwanted, spurious response variability is reduced with pre-printing. With carefully designed experiments it is possible to estimate and monitor the possible bias due to pre-printing. With such an estimate we can then correct the statistics and at the same time keep the advantage of increased measurement accuracy.

A task for future research is to increase the knowledge of pre-printing effects on self-administered questionnaires when the pre-printed variables are categorical. Although we could not find any bad examples in our review, knowledge is scarce concerning when and in which applications it works well and when and in which it does not. A carefully planned, preferably multidisciplinary study, would be a valuable contribution.

6. Summary

In surveys where respondents are contacted repeatedly, it is common that previously collected (historical) information is used in the data collection. To study the effects of practices such as pre-printing and dependent interviewing is important, but also often difficult. The variation between the surveys in which the methods are used is large and the results of such studies are therefore not easily generalised. In addition, certain methodological issues involved may make research on this subject expensive. Most of the work that has been done is in the form of case studies connected to interview surveys of household populations. Since the focus in this article was on establishment populations and on surveys using pre-printed self-administered questionnaires, the findings complement that earlier work as well as add to the general understanding of the possible effects of using historical data in the data collection.

In this article we have discussed the possible advantages and disadvantages of pre-printing historical data on questionnaires. These matters were studied in a large-scale experiment, and in that experiment treatments with pre-printed questionnaires outperformed treatments without pre-printing. Pre-printing led to better data quality in the sense of less spurious response variability and fewer and smaller outliers. Moreover, the study revealed no indications of the possible drawbacks, such as tendencies to underestimate change or frequent recurrences of the pre-printed values. Pre-printing is clearly recommended for the type of survey where the experiment was conducted. We also discussed Swedish experiences of pre-printing. In general these are positive. In a recent review no really bad examples of pre-printing were found. Another experiment with categorical variables in an electronic questionnaire also indicated advantages with

pre-printing. We suggest that pre-printing should be considered when there are problems with high response variability and measurement errors. With a proper testing of its effects and a proper redesign of questionnaires, we believe that pre-printing is a useful method to improve data quality, and a method that has advantages for the respondents as well as in the practical editing work.

7. References

Corti, L. and Campanelli, P. (1992). The Utility of Feeding Forward Earlier Wave Data for Panel Surveys. Survey and Statistical Computing, 109–118.

Hansen, M.H., Hurwitz, W.N., and Pritzker, L. (1964). The Estimation and Interpretation of Gross Differences and the Simple Response Variance. In Contributions to Statistics, C.R. Rao (ed.). Presented to P.C. Mahalanobis on the Occassion of his 70th birthday, Pergamon press Ltd, Calcutta, 111–136.

Holmberg, A. (2002). Rapport om förtryckningens fördelar och nackdelar: En sammanställning av erfarenheter på SCB samt resultat från två experimentella studier. (The pros and cons of pre-printing: A summary of experiences at Statistics Sweden and results from two experimental studies.) Statistics Sweden. [In Swedish]

Krosnick, J. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. Applied Cognitive Psychology, 5, 213–236.

Mathiowetz, N.A. and McGonagle, K.A. (2000). An Assessment of the Current State of Dependent Interviewing in Household Surveys. Journal of Official Statistics, 16, 401–418.

O'Muircheartaigh, C.A. (1986). Correlates of Reinterviewer Response Inconsistency in the Current Population Survey. Proceedings of the U.S. Bureau of the Census Second Annual Research Conference, March 23–26. Reston, VA.

Pafford, B.V. (1986). Studies of Response Errors in NASS Surveys: The Effect of Using Previous Survey Data. Proceedings of the American Statistical Association, Section on Survey Research Methods, 574–579.

Pafford, B.V. (1988). The Influence of Using Previous Data in the 1986 April ISO Grain Stock Survey. Washington D.C.: National Agricultural Statistical Service.

Särndal, C-E., Swensson, B., and Wretman, J. (1991). Model Assisted Survey Sampling. New York: Springer Verlag.

Stanley, J. and Safer, M. (1997). Last Time You Had 78, How Many Do You Have Now? The Effect of Providing Reports on Current Reports of Cattle Inventories. Proceedings of the American Statistical Association, Section on Survey Research Methods, 875–879.

Received February 2003
Revised January 2004