Component of Statistics Canada Catalogue no. 11-522-X
Statistics Canada's International Symposium Series: Proceedings

Article

Symposium 2008:
Data Collection: Challenges, Achievements and New Directions

Effects of Number of Call Attempts on Nonresponse Rates and Nonresponse Bias – Result from Some Case Studies at Statistics Sweden

by Jan Hörngren, Peter Lundquist and Sara Westling

2009

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2008
Data Collection: Challenges, Achievements and New Directions

Effects of Number of Call Attempts on Nonresponse Rates and Nonresponse Bias – Result from Some Case Studies at Statistics Sweden

Jan Hörngren, Peter Lundquist and Sara Westling[1]

Abstract

The choice of number of call attempts in a telephone survey is an important decision. A large number of call attempts makes the data collection costly and time-consuming; and a small number of attempts decreases the response set from which conclusions are drawn and increases the variance. The decision can also have an effect on the nonresponse bias. In this paper we study the effects of number of call attempts on the nonresponse rate and the nonresponse bias in two surveys conducted by Statistics Sweden: The Labour Force Survey (LFS) and Household Finances (HF).

By use of paradata we calculate the response rate as a function of the number of call attempts. To estimate the nonresponse bias we use estimates of some register variables, where observations are available for both respondents and nonrespondents. We also calculate estimates of some real survey parameters as functions of varying number of call attempts. The results indicate that it is possible to reduce the current number of call attempts without getting an increased nonresponse bias.

Key Words: Call attempts (WinDATI-events), Nonresponse rate, Nonresponse bias.

1. Introduction

Why is it important with the choice of number of call attempts in a (telephone) survey?

From one point of view: If you have a large number of call attempts it is very time-consuming and then very costly - Time is money! From another point of view: A small number of call attempts will decrease the response set which results in an increase of the variance, that is larger sample errors; larger confidence intervals, and probably more serious is that we will increase the risk that nonresponse bias occur.

In this paper we use paradata to calculate the response rate as a function of the number of call attempts. As an indication of the nonresponse bias we use estimates of different register variables, where observations are available for both respondents and nonrespondents. Also in this case we calculate estimates for different number of call attempts. This paper is mainly based on two different studies conducted at Statistics Sweden: Lundquist et al. (2007) and Westling (2008). An earlier extensive study in this area is Japec and Hörngren (2005). But the results in Japec and Hörngren are not comparable with the result in this paper, due to different definitions of a call attempt.

2. Possibilities and constraints

In this section we discuss the possibilities and constraints for analyzing and developing call attempt strategies at Statistics Sweden. The base in the analysis is paradata from WinDATI that is Statistics Sweden's computer assisted interview (CAI) system. When an interviewer uses WinDATI each action result in a so called *event* (WinDATI-event). Following WinDATI-events has been used in the analysis:

[1]Jan Hörngren, Senior Methodologist, Statistics Sweden, 702 33 Örebro, Sweden (jan.horngren@scb.se); Peter Lundquist, Senior Methodologist, Statistics Sweden, Sweden, Box 24300 104 51 Stockholm, Sara Westling, Senior Methodologist, Statistics Sweden, 702 33 Örebro, Sweden

- completed interview
- terminated interview → new appointment
- terminated interview → final code
- appointment code
- busy telephone number
- no answer (non contact)
- further call attempts during the day terminated
- final coding (nonresponse coding)
- close/terminate sample unit

Each event has a unique code and in this case each event (above) corresponds with a call attempt[2].

Another important object concerning possibilities and constraints is that we have two interviewer teams, field interviewers and centralized interviewers (CATI-interviewers):

- Field interviewers are located across the country and working from home. A field interviewer conducts face-to-face **and** telephone interviews. The field interviewer receives periodically a number of cases per survey, that is a personal workload.

- CATI interviewer is a centralized facility who conducts only telephone interviews. They receive cases from a common database.

A field interviewer and a CATI-interviewer also have different contracts.

An ambition at Statistics Sweden is that all interviewers should receive cases (in telephone surveys) from a common database. A common database for all interviewers is probably a condition to succeed with a call attempt strategy.

3. Empirical studies of the effects of number of call attempts

3.1 Survey data and analyses

The studies are based on two different surveys:

- The Labour Force Survey (LFS) is a monthly survey with a sample size of 21,000 individuals every month. The sample frame is the Swedish Register of the Total Population. The data collection is conducted by CATI and field interviewers. The response rate is 83 percent (average value).

- Household Finances (HF) is an annual survey with a sample size of 17, 000 households. Also in this survey the sample frame is the Swedish Register of the Total Population (network sample). The data collection is conducted by field interviewers. The response rate is 70 percent.

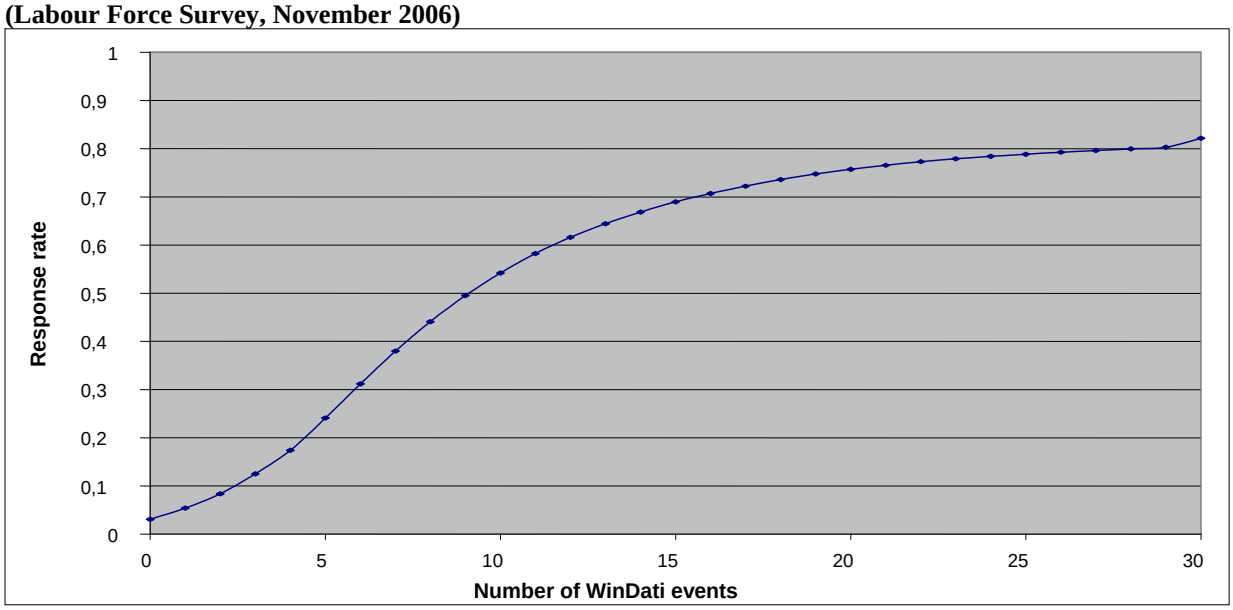
The studies of the effects of number of call attempts have the focus of nonresponse rate and nonresponse bias. Nonresponse rate is basic indicator of survey quality. Nonresponse bias is a very important indicator when we are dealing with call attempt strategies. It can really give us information about what happens with the estimates of interest.

[2] In this paper we use the concept WinDATI-event to avoid invalid comparisons with other call attempt concepts

3.2 Effect on the nonresponse rate

In this section we will study the sense of the field work efforts (in terms of WinDATI events) on the response rate. In Figure 3.2-1 shows the accumulated response rate in the LFS with respect of number of WinDATI events. (In this figure and following figures the valid value for maximum number of WinDATI events (=30) on the X-axis is ≥30. Cases over 30 events are extremely rare occasions)

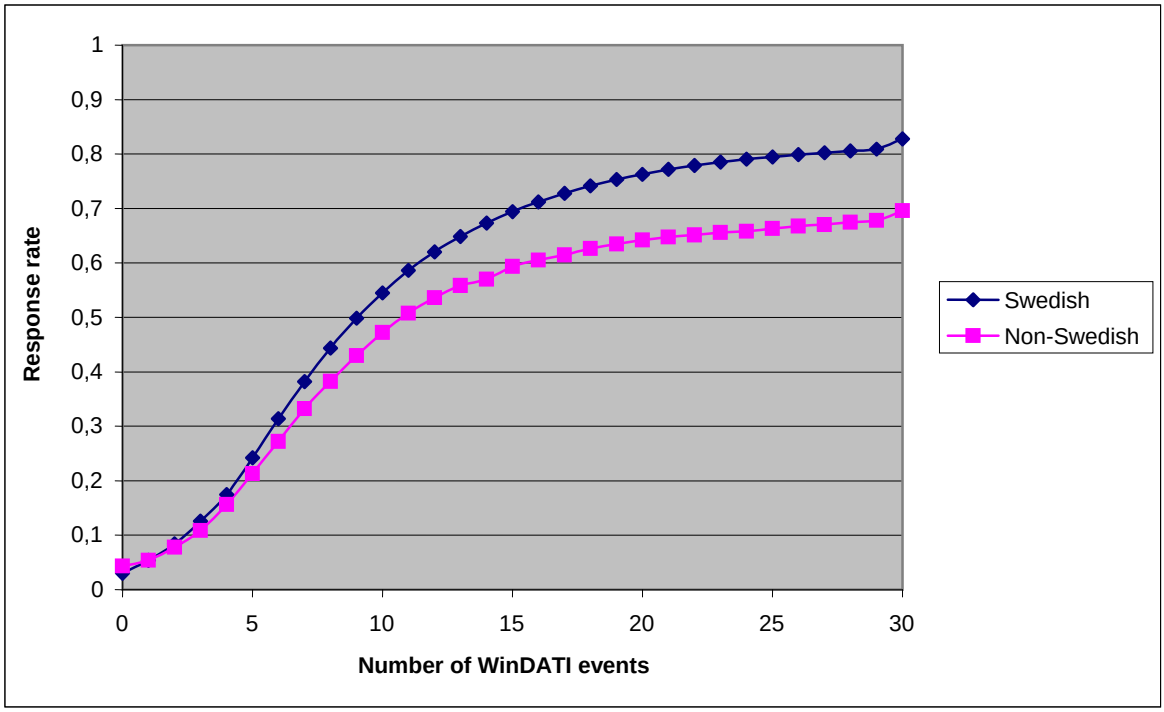
Figure 3.2-1
Accumulated response rate (overall) with respect of number of WinDATI-events (Labour Force Survey, November 2006)



We can see that the response rate seems to be stable after 20 WinDATI events and the "response level" reach about 70% after 15 events. Figure 3.2-1 is valid for the entire sample. Do we have a different picture for different domains?

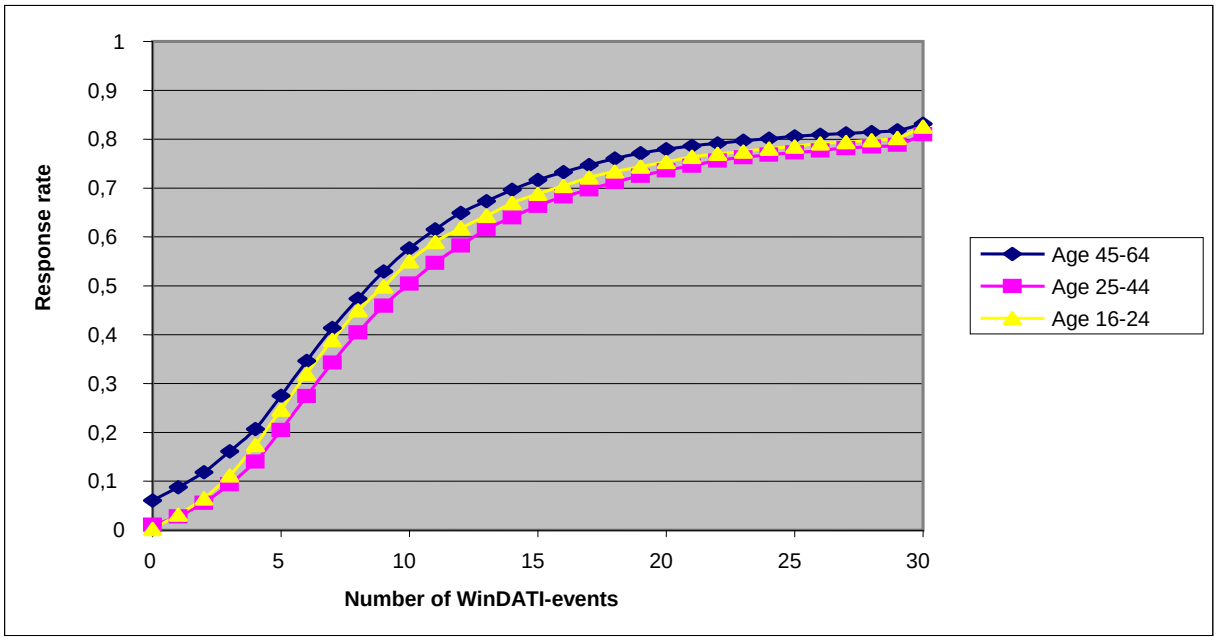
Figure 3.2-2 gives an illustration of the accumulated response rate divided by citizenship and Figure 3.2-3 by classes of age.

Figure 3.2-2
Accumulated response rate with respect of number of WinDATI-events by citizenship (Labour Force Survey, November 2006)



The pattern is similar with Figure 3.2-1 (overall). The accumulated response rate by citizenship does not tell us much more than that we have a larger nonresponse rate among non-Swedish citizens, that is already a well-known fact.

Figure 3.2-3
Accumulated response rate with respect of number of WinDATI-events by age (Labour Force Survey, November 2006)



The pattern repeats itself. Just as in the two figures above the response rate stabilize after 20 WinDATI events and the level reach 70% after 15 events.

3.3 Effect on the nonresponse bias

The approach to estimate nonresponse bias is to use register variables. When we match a register variable with the entire sample we have observations on an arbitrary register variable for both respondents and nonrespondents. In that case we can estimate the nonresponse bias as follows:

Calculate estimates based on respondents, $\hat{Y}_r$

Calculate estimates based on respondents + nonrespondents, $\hat{Y}_s$

And,

Estimated relative bias (%) = $\frac{\left(\hat{Y}_r - \hat{Y}_s\right)}{\hat{Y}_s} \cdot 100$ according to LFS

In a matter of fact we are estimating the bias of the register variable and not the real survey variable. But we can find register variables that are related to the survey variable of interest. For example in the LFS ”register employed” (based on statement of income) is closely related to employed according to LFS. But in this case we use Register employed as auxiliary information in the estimation procedure in a calibration estimator. So you could say that the register variable is already busy. It should be an infected analysis if we use a register variable that is already in use in the estimation procedure.

In the following example Annual Salary according to the Swedish Tax Register is our $\hat{Y}$ in the formula above. In a post-survey procedure we have match HF 2006 (the entire sample) with data from the Swedish Tax Register and paradata (WinDATI) on a sample unit level.

Figure 3.3-1
Annual salary – Estimates of relative bias (%) with respect of number of WinDATI events by age (Household Finances 2006)

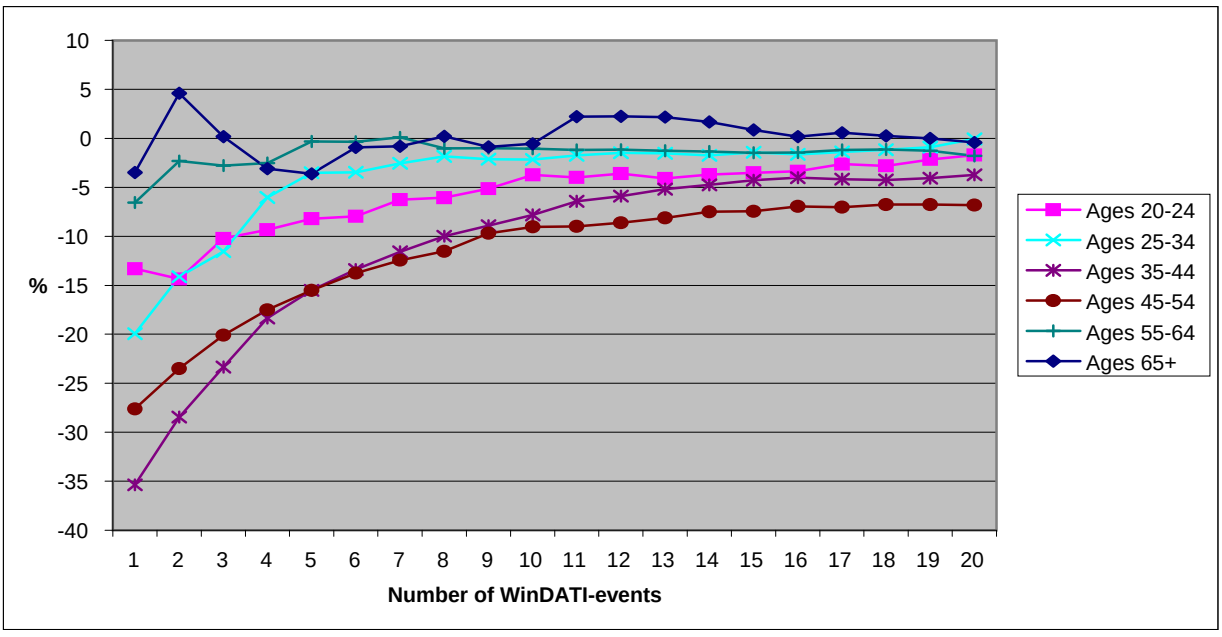


Figure 3.3-1 illustrates the relative bias for different groups of age. The blue line (ages ≥65) has almost a neglect bias all the way. Individuals over ages 65 also have none or low salary, which probably is the factual explanation.

We have a similar pattern for ages 55-64 (green line). Concerning the domains of ages we can see a large negative bias in the beginning of the data collection and the bias steady reduces with respect to number of WinDATI events.

The gap between different domains seems to be negligible after 15 WinDATI events. There is another study who confirms these results. Isaksson, Lundquist and Thorburn (2008) estimate the relative bias of annual salary in the LFS. The results show that the relative bias stabilizes after 15 WinDATI events.

3.4 Effect of number of WinDATI events on a real survey parameter

As mentioned in 3.3 the estimates of the bias of the register variable. Even if it is an indicator of the bias of survey variables it is of interest to complete the analysis with study the effect of Number of WinDATI events on a real survey parameter. In Figure 3.4-1 we illustrate the estimates of the unemployment ratio (LFS) according to the number of WinDATI events.

Figure 3.4-1
Estimates of the unemployment ratio (LFS 2006) with respect to the number of WinDATI events

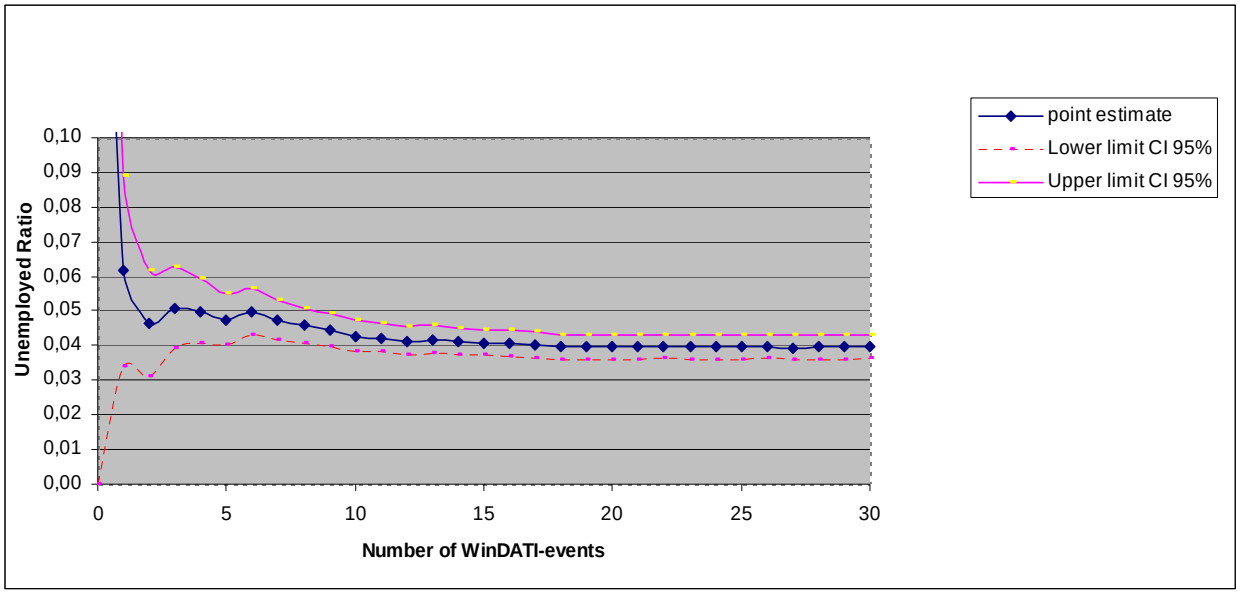


Figure 3.4-1 shows estimates of the unemployment ratio with the limits of a confidence interval (95%). Estimates based on few number of WinDATI events are hardly valid estimates. These estimates are based on, comparatively, very few observations. But also when we are study a real survey estimate it looks stable after 15 WinDATI events.

4. An experiment with the purpose to reduce the number of call attempts

The sample in a typical individual or household survey at Statistics Sweden is selected from the register of total population. For all sampled individuals the interviewers have the name, postal address, age and gender. By using the personal ID-number it is also possible to match other information such as education, occupation and income. As mentioned above (section 2) Statistics Sweden has two interviewer teams: field and CATI-interviewers. In some major surveys both teams are involved in the data collection. Earlier studies have shown that field interviewers who use the auxiliary information gain a higher personal response rate (Japec and Lundquist 2000). The field interviewers have developed their own contact strategies, while the CATI-interviewers are using an automatic call scheduler that not use auxiliary information or information from earlier waves in panels surveys.

Logistic regression in Japec and Lundquist (2000) has confirmed that the effect of register variables decline after the first call attempts. Instead the process-variables dominate when it comes to explain the probability to respond. In Lundquiat et al. (2007) we have focused on factors that could explain the probability to respond in the first and second call attempt. We have investigated differences and similarities in the contact strategies (based on the observed data) for two surveys: LFS and HF. The LFS is an individual rotating panel survey (eight waves) and HF is a one-wave household survey.

In a logistic regression we found some similarities between LFS and HF: Evening (5pm-10pm) and individuals over 65 years are factors that increase the response rate. People living in big cities decrease the response rate. For the LFS did Monday-Thursday give a higher response rate, this was not found in the HF. (This finding may reflect how the interviewers are working. The same interviewers are included in both surveys.) The kind of house property (rented or owned) did have an effect in HF but not in LFS. Income did have an effect in LFS but not in HF. It was easier to find factors for the LFS than for the HF, this may depend on that the respondents are the sampled persons in LFS. In the second call attempt the general similarities were still significant, the effects in HF were however weaker. It was a difference in how the data were collected in the two surveys. One conclusion is that we need different strategies for the surveys.

Further analysis on the LFS, using operational research on the estimated response probabilities from the logistic regression, gave an estimated optimal response rate for the first call given the distribution of the interviewer resources during the day. The solution was to only use the evening shift for the first call attempt. This would increase the response rate with 10 per cent compared with the used strategy, according to the model.

During the first quarter 2008 the work with designing an embedded experiment in the LFS started. The purpose of the experiment is to investigate if it is possible to reduce the interviewer time for call attempts. A draft of the design of the experiment is produced according to the experiment manual used at Statistics Sweden. (The development of the strategy working with embedded experiments is given in Karlberg et al. 2002).

The experiment is planned for the central CATI-group. The purpose is to schedule the first two call attempts in the experiment-group (the control group will work as usual). For the new panel the two calls should be done in the evening Monday-Thursday, for wave 2 to 8 (7/8 of the sample) we already know when the previous interview took place. This information together with information on convenient time for the next interview (additional question) is used automatically by the call scheduler to book the sample to a time slot and day.

5. Summary

We have studied the effects of the fieldwork in terms of WinDATI events used in the LFS and HF on nonresponse rates and nonresponse bias. In the analysis we found that:

- Nonresponse rate seems to stabilize after 20 WinDATI events in the Labour Force Survey and after 15 WinDATI events in Household Finances.
- Nonresponse bias seems to stabilize after 15 WinDATI events in the Labour Force Survey and after 10-15 WinDATI events in Household Finances.
- Estimates (overall) of the unemployment ratio seems to stabilize after 15 WinDATI events

The results are a clear indication that it is possible to reduce the current number of WinDATI events (call attempts) without getting an increased nonresponse bias. In spite of that Statistics Sweden is currently not ready for a responsive design strategy in the data collection process. More analysis is necessary. For example, if we look at the LFS which is very important input for the National Accounts and other considerable secondary statistics. We need to go further with the analysis for domains of special interest. Essential statistics can not be put at risk.

Furthermore we need more and better paradata. Some example:

- Connecting WinDATI event and telephone number (home number / cell phone number / place of work number)

- Interviewer time by WinDATI event codes
- Separate paradata database independent of the WinDATI system
- Special WinDATI event codes for panel surveys

The next step in this work is to stage the experiment described in section 4.

References

Isaksson, A., Lundquist, P. and Thorburn, D. (2008). Optimalt antal kontaktförsök i en telefonundersökning, *Research and Development,* 2008:1, Statistics Sweden [In Swedish].

Japec, L. and Hörngren J. (2005). Effects of Field Efforts on Nonresponse Bias and Costs in the Swedish Labour Force Survey, *Quality Issues in Interview Surveys – Some Contributions,* (Doctoral Dissertation of L. Japec), Department of Statistics, Stockholm University, Paper IV.

Japec, L. and Lundquist, P. (2000). Bortfallet - påverkas det av intervjuarnas attityder och strategier?, unpublished report, Stockholm, Sweden; Statistics Sweden. [In Swedish]

Karlberg, M., Björnram, A., Boynton, I-M., Göransson, B. and Lundquist, P. (2002). Embedded Experiments in Sample Surveys at Statistics Sweden: The Development of an Experimentation Manual. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Lundquist, P. et al. (2007). Utveckling för system av kontaktstrategier i intervjuundersökningar med individer och hushåll Delrapport I, unpublished report, Stockholm, Sweden: Statistics Sweden. [In Swedish]

Westling, S. (2008). Utveckling för system av kontaktstrategier i intervjuundersökningar med individer och hushåll Delrapport II, unpublished report, Örebro, Sweden: Statistics Sweden. [In Swedish]