

Distr. GENERAL

02 September 2013

WP 22

ENGLISH ONLY

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on Statistical Data Collection

(Geneva, Switzerland, 25-27 September 2013)

Topic (v): Integration and management of new data sources

**MIXED-MODE DATA COLLECTION – RECENT DEVELOPMENTS AT
STATISTICS NETHERLANDS**

Working Paper

Prepared by Barry Schouten¹ and Gabriël van Dam², Statistics Netherlands

Summary: Over the last five years, Statistics Netherlands has redesigned most of its large national social surveys to mixed-mode survey designs including web. The main motivation is the reduction of survey costs. However, in all cases total mode effects with respect to previous designs need to be as small as possible and need to be understood and justified. For this reason various parallel runs and pilot studies have been conducted. In 2011, a large-scale mixed-mode study was linked to the Crime Victimisation Survey in order to disentangle mode effects into mode-specific selection and measurement effects. In 2010 and 2012, parallel runs were applied to the Labour Force Survey.

In this paper, we will focus on web and contrast web to the other survey modes paper, telephone and face-to-face. We will discuss the implications mixed-mode design in general and of including web as a survey mode in mixed-mode designs in particular.

I. Introduction

1. The emergence of web as a survey mode has renewed the discussion about mixed-mode survey designs. Web surveys show similarities to mailed questionnaires, but they are computer-assisted and do not require processing of paper questionnaires. Consequently, and since the web survey mode is not interviewer-assisted, it is the cheapest survey mode; it may be up to 25 times cheaper than face-to-face interviews per sample unit. Response rates to web are, however, generally low, so that web has a very strong quality-cost differential. It is this differential that brought new impetus to the discussion on mixed-mode survey designs. Market research organisations have recognized the strong quality-cost differential of web right from the start and have been including web in their data collection designs for several years. National statistical institutes are more hesitant as they are less subject to economic competition. Budget cuts and decreasing response rates of the traditional survey modes are forcing also these institutes to seek cheaper means of data collection.

¹ Statistics Netherlands, jg.schouten@cbs.nl

² Statistics Netherlands, gdam@cbs.nl

2. Statistics Netherlands was one of the first institutes to design new social surveys and redesign ongoing social surveys using multiple modes including web. It has done so mostly in a pragmatic way; attempting to reduce costs while preserving accuracy and comparability in time. However, for most new designs parallel runs were conducted and in 2011 a large-scale experimental study was performed in order to disentangle mode effects for key statistics from the Dutch Crime Victimization Survey and the Dutch Labour Force Survey. In this paper, we will summarize the findings from the various (re)designs and experimental studies. We will focus on web as a new survey mode in mixed-mode data collection.
3. In section II, we discuss the new paradigm of mixed-mode designs for national statistical institutes. In section III, we describe the redesign of social surveys at Statistics Netherlands and the experimental studies linked to the redesign. We end with a discussion of remaining issues.

II. The mixed-mode paradigm

4. The introduction of the web survey mode has led to a strong interest in mixed-mode data collection in the literature and at survey methodology conferences. We will first start with a short general discussion and then move to different viewpoints on how to make quality-cost trade-offs.
- A. Quality and costs**
5. There is a vast and growing amount of literature on mixed-mode survey designs, and more specifically, on differences between web and the survey modes mail/paper, telephone and face-to-face. Literature overviews and discussions can be found in De Leeuw (2005), Dillman et al (2009), Jäckle et al (2010), Dex and Gumi (2011) and Klausch et al (2012).
 6. The literature conjectures and confirms the hypothesis that web has an impact on all survey errors: Web shows undercoverage, it leads to different response rates and it leads to mode-specific measurement effects. As a consequence, and since the various errors may interact, any discussion about web as a survey mode in mixed-mode designs is complex to a degree that prohibits strong conclusions or easy decisions. The introduction of web at Statistics Netherlands has affected a broad area of survey methodology, ranging from data collection strategies and questionnaire design to nonresponse adjustment.
 7. In the mixed-mode paradigm, there are two viewpoints: preserving quality while reducing costs and maintaining survey costs while improving quality. The first viewpoint is generally taken by market research organisations and national statistical institutes, while the second viewpoint is more present in academic literature. We will briefly discuss them in the following subsections. The difference in viewpoint can also be found in the subtle distinctions between the quality dimensions accuracy, comparability in time and comparability in domains. Accuracy is the traditional statistical quality dimension, relating statistical estimators to population parameters of interest. Statistics are accurate when they have a small mean square error, i.e. when they have small bias and high precision. Comparability in time and in domain are the same as stable accuracy over time and stable accuracy over relevant subpopulations. Statistical institutes are usually willing to sacrifice some accuracy to costs as long as the accuracy is stable, while academic research more often focuses on the relation between absolute accuracy and survey mode.
 8. Linked to the accuracy of statistics, the mixed-mode paradigm raises (again) the complicated discussion on true scores or true values of statistics. This holds especially true for non-platonic measures like attitudes or values. What survey mode tells the truth, when all modes lead to different answers? Statistical institutes, like Statistics Netherlands, have again adopted a pragmatic view on this issue and have usually accepted changes in accuracy as long as these level shifts can be assessed and explained to some extent using parallel runs.

9. The trade-off between quality and costs can be made more mathematical and transparent by employing adaptive and responsive survey designs. These designs differentiate and tailor efforts to sample units. The survey mode is one of the most influential features in such designs.

B. A focus on costs

10. The main motive of Statistics Netherlands for doing mixed-mode surveys and including web as a survey mode is cost reduction. However, even if costs would not be the primary motive, the strong difference in costs to traditional survey modes raises the legitimate question whether traditional budgets are really needed. Statistics Netherlands started experimenting with web surveys around seven years ago. It asked itself the following questions:
 - a. What about response rates of web and mixed-mode designs?
 - b. Can we do mixed-mode surveys including web?
 - c. What are consequences for the lengths of questionnaires, and can we transform complicated survey items like educational level or profession to web?
 - d. What about level shifts and stability of statistics?
 - e. What about undercoverage of web?

11. The last question received less attention than the other questions. It also lacks a good definition as access to web is not restricted to personal computers or laptops. In section III.A, we will return to these questions.

C. A focus on quality

12. Academic research has a stronger focus on mode-specific survey errors than on costs or comparability. It is assumed that survey modes perform differently with respect to the various survey error components: coverage, nonresponse and measurement. A mixed-mode design including web may overcome the weaknesses of one mode by the strengths of another mode. Subpopulations may respond better to some modes than to others, may have access to some modes only or may be less prone to undesirable measurement response styles, like satisficing or social desirable answering, in some modes as compared to others. Of the survey error components, mode-specific measurement receives the most attention. It is also the most complex component as it is conditional on the other two components.
13. The research into mode-specific accuracy has two important prerequisites: 1) The mode-specific survey error must be decomposed into mode-specific coverage, nonresponse and measurement differences, and 2) a benchmark or true population parameter is needed. Without a decomposition, it is unclear how modes can optimally be combined. Without a population parameter or true value, there is no optimum. Usually, the difference between modes, termed the mode effect, is adjusted for selection due to undercoverage and nonresponse, and is then treated as the measurement effect or pure mode effect. The adjustment for mode-specific selection can be based only on information external to the survey and therefore there remains a risk of entangled survey error components. In section III.B, we will discuss an experimental study in which Statistics Netherlands has attempted to identify the relative differences between coverage, nonresponse and measurement for a number of key statistics.

D. Adaptive survey designs

14. Over the last six years, responsive survey designs (Groves and Heeringa 2006) and adaptive survey designs (Wagner 2008, Schouten et al 2011) have emerged as means to systematically make trade-offs between quality and costs. Different sample units may receive different survey design features based on known auxiliary information about these units or based on process data that is collected during the survey. As mentioned in section II.A, these designs may be used to tailor the survey mode

to the sample unit, i.e. different sample units would be offered different (sets of) survey modes. However, adaptive and responsive survey designs are still in their infancy, and, to date, have focussed mostly on nonresponse and on single mode surveys. Some recent attempts have been made to extend frameworks to measurement, see Calinescu et al (2012).

III. Developments at Statistics Netherlands

15. In 2005, Statistics Netherlands conducted its first web survey pilot studies. In 2006 the first mixed-mode survey including web was launched, the Crime Victimization Survey (CVS). All early studies were linked to surveys that were done under commission. In 2007, Statistics Netherlands initiated a large project called Redesign of Social Surveys (in Dutch Herontwerp Persoonsenquêtes or HPE) aiming at a redesign of its own main surveys: the Labour Force Survey (LFS) and the Health Survey (HS). Within the HPE project several parallel runs were done in order to assess level shifts.
16. The Crime Victimization Survey showed large mode effects, mostly between interviewer and non-interviewer survey modes. These differences formed the basis for a research project Mode Effects in Social Surveys (in Dutch Mode effecten in persoonsstatistieken or MEPS) that started in 2010 and that has recently been finished.
17. Based on the results from the HPE and MEPS projects, an investigation was started of adaptive survey designs for the CVS and LFS. The HPE and MEPS projects will form input to a new European ESSnet project on mixed-mode designs for social surveys (Blanke and Luiten 2012).
18. In the following subsections the findings of the various projects will be discussed.

A. Project Redesign of social surveys

19. Project Redesign of Social Surveys mainly dealt with the five questions formulated in section II.B. Here, we will run by the five questions. A detailed summary of results for the LFS can be found in Cuppen et al (2011) and Mars et al (2012).

Response rates

20. The first question that was asked is whether response rates of web surveys are close to those of the traditional survey modes, and if not, if they can approximate these response rates by adding the more expensive interviewer modes in sequence. Later, this question was detailed by looking at subgroup response rates and representativeness for a fixed set of socio-demographical variables (age, ethnicity, job status, house value at zip code, household composition, urbanisation level of area of residence).
21. The response rates to web turned out to be much lower than those of any other single survey mode for all subgroups. Overall web response rates range from as low as 15% to a maximum of 35% over various surveys conducted at Statistics Netherlands. For this reason, it was quickly decided that web cannot be used as a single mode in cross-sectional surveys and that interviewer modes need to be added in a sequential design. The default mode strategy is web → telephone + face-to-face; web nonrespondents are approached by telephone if they have a registered phone number and by face-to-face otherwise. Most surveys have subtle deviations from this default strategy, either by adding mail or by restricting the set of modes for some subpopulations. In the LFS large households are never approached by telephone, for instance. The response rates of the sequential mixed-mode designs are in virtually all cases similar to those of the traditional single mode designs.
22. Various pilots have been fielded to determine the optimal number of reminders and to construct optimal panel recruitment questions in web for the LFS rotating panel. Research into optimal advance letters is still on-going. Currently, the web strategy consists of two reminders with one week time lags and an advance letter that announces a follow-up using telephone and face-to-face. The recruitment rate for the LFS panel still is much lower than the traditional rates achieved by interviewers in telephone and face-to-face.

Logistics and timeliness

23. Directly following the question of response rates, was the question whether mixed-mode surveys using web can be done in practice.
24. Web is a non-interviewer mode and must be handled differently in the survey administration systems. Still, the administration systems for web, telephone and face-to-face must be able to communicate in order to reallocate and monitor sample cases. This turned out to be a complex, logistical operation and required additional loops and bypasses to the existing systems. However, after the first few years of mixed-mode data collection, these systems are now running satisfactorily. Monitoring of data collection has, however, become more complicated and labour intensive.
25. An important conclusion was that the total duration of the data collection needed to be extended to three months: a first month of web data collection, a second month of telephone data collection for web nonrespondents with a registered phone and a third month of face-to-face data collection for the remaining web nonrespondents. The additional month for face-to-face is inevitable to reallocate nonrespondents to interviewers.

Questionnaire design

26. Two main questions were asked: what is the maximal length of questionnaires in web and can complex survey items like educational level be implemented in web. These questions essentially relate to break-off, item nonresponse and measurement errors in web.
27. The maximal average length of web questionnaires was set at the same length as for telephone surveys: at 20 minutes. The restriction of the interview duration had a major impact on the questionnaire design. Both the Labour Force Survey and Health Survey were too long to transform to mixed-mode and had to be bisected. For the LFS parts of the survey were moved to the panel follow-up waves. For the HS an additional wave was introduced.
28. Complex survey items like educational level, type of profession and type of industry require automated coding due to the large amount of possible answer categories. In interviewer-assisted modes this is done using a database and probing by the interviewer. In web such an approach cannot easily be performed as the communication with the database may be slow and prone to technical errors. For this reason the questions were either simplified and shortened or removed and replaced by registry data. A pilot study was conducted in 2011 and led to a positive evaluation of the new questions for educational level; see Mars et al (2012). The questions relating to type of profession are now taken from tax board registry data. The questions relating to type of business are shortened. For all statistics, small level shifts are expected.

Level shifts and stability

29. The fourth question that was asked concerned comparability in time: what is the size of level shifts and does the new design produce a stable accuracy. For all redesigned surveys, mode effects are expected after nonresponse adjustment, i.e. are expected to be a mix of measurement effects and unadjusted selection effects.
30. Two measures were taken. First, parallel runs were conducted for the redesigned surveys. For the LFS the parallel run had a length of six months. For the HS the parallel run had a length of one year. The level shifts for the LFS are estimated on a monthly basis and will be estimated and published by the end of 2013. For the HS the level shift is estimated for the annual statistic for one year only. For this reason, a structural time series model, as described in Van den Brakel and Roels (2009), was applied to borrow statistical strength from the long HS time series. After the introduction of web, the HS level shifts were small for almost all variables.
31. Second, a detailed quality report was made for various years of the surveys before and after redesigns. The quality report consists of a large number of quality indicators, clustered under quality dimensions accuracy, coherence, relevance, timeliness, comparability and accessibility. Accuracy is

further detailed to precision, unit-nonresponse, item-nonresponse, undercoverage, linkage and measurement. The quality reports are used to understand and explain possible level shifts due to the redesign. Again for the HS relatively few changes in quality levels were found.

32. An important finding during the parallel runs is the instability of response rates for web. The web response rates fluctuate more strongly from one month to the other than must be expected from sampling variation. As yet, it is unclear what causes this variation. It is conjectured to be the result of seasonal influences, the timing of weekends and the occurrence of special events. This variation leads to varying shares of survey modes to the total response. Since mode effects are expected, varying shares lead to varying mixtures of mode effects, and, hence, to method effects. A pragmatic solution to this problem is found in weighting the response to fixed shares of modes; see Buelens and Van den Brakel (2012). When the web response rate is larger than the annual mean, then web respondents receive a smaller weight.

Coverage of web

33. The last question that was asked, related to the coverage of the population from households that have access to the Internet at home. Any selection effects due to undercoverage are partially removed by weighting. However, since surveys of national statistical institutes are considered democratic instruments, undercoverage is also a political issue. By adding interviewer modes in mixed-mode designs, in sequence to web, this problem is mostly resolved. Still, it is considered important to know what persons and households drop out because of undercoverage, as these households have a higher probability of being a final nonrespondent and may be annoyed by the sequential design.
34. Using follow-up surveys in face-to-face, the household web coverage rate in the Netherlands is estimated at 88% in 2009 and at 90% in 2011. The web coverage rate seems to converge to a rate around 90%. Underrepresentation is found for the older households, for lower incomes and for dwellings with a lower real estate value. These types of households are now considered as candidate subgroups for tailoring in adaptive survey designs for the LFS.

B. Project Mode effects in social surveys

35. Although large investments have been made in systems and parallel runs, project Redesign of Social Surveys is a relatively pragmatic project aiming at comparable and stable statistics using survey designs that include web. Project Mode Effects in Social Surveys is complementary to this project and has a focus on the accuracy of statistics as a function of survey mode. It was initiated to decompose mode effects for a number of key statistics. The results from the decomposition are input to changes in mode strategy, questionnaire design and adjustment methods.
36. In 2009, time series of key statistics of the “oldest” mixed-mode survey, the Crime Victimation Survey (CVS), showed unrealistic changes. These changes were attributed to the large variation in the shares of survey modes to the total response. The CVS is special in the sense that municipalities can buy additional samples on a yearly basis. These samples are surveyed by market research organisations using mostly paper and web questionnaires, but are combined with the Statistics Netherlands regular CVS to form national statistics. These combined samples show large variations in the shares of web and paper. For this reason, the CVS was selected as the instrument to perform an experimental study.
37. The experiment consisted of two waves. In the first wave, 8800 sample persons were randomly assigned to one of the four survey modes, face-to-face (CAPI), telephone (CATI), web or paper. The regular data collection strategy was applied for each mode, e.g. length of data collection period and number of visits/calls/reminders. The first wave of the experiment was CVS with two modifications. Part of the modules at the end of the CVS survey questionnaire were replaced by the LFS module for employment status and by two scales from the European Social Survey (ESS). The full sample were approached once more in a second wave using face-to-face. This second wave of the experiment employed a new questionnaire, consisting of: a repetition of the key statistics from the CVS, general attitudes towards safety and politics, general attitudes towards surveys, an evaluation of survey

participation in wave 1, an evaluation of the CVS questionnaire (wave 1 respondents only) and access to web and mode preferences. Table 3.1 contains the response rates to wave 1 and wave 2 given the allocated wave 1 mode.

Table 3.1: Sample sizes and response rates to wave 1, wave 2 and both waves.

| | CAPI | CATI | Paper | Web | Total |
|----------------------|------|------|-------|-----|-------|
| Wave 1 response | 61% | 45% | 49% | 29% | 46% |
| Wave 2 CAPI response | 49% | 47% | 50% | 49% | 49% |

38. Buelens et al (2012) gives a detailed description of the estimation strategy that was employed to disentangle mode effects and an evaluation of the experimental assumptions. Essentially, the first wave is weighted towards the second wave where all wave 2 variables are candidate weighting variables. The mode effect decomposition was made relative to face-to-face, i.e. face-to-face was treated as the benchmark. Face-to-face was selected because it does not suffer from undercoverage, it has the highest response rates and is the traditional survey mode for many Statistics Netherlands' surveys. Detailed results can be found in Schouten et al (2013).
39. For the LFS variable employment status, the total mode effect turned out to be a mix of coverage, nonresponse and measurement effects. Importantly, in the weighting model for employment status from wave 1 to wave 2, only standard registry variables were selected, i.e. none of the CVS repeated variables and none of the attitudes towards politics, safety and surveys turned out to play a role in mode-specific response or answering behaviour. The selection effect for the LFS variable employment status can be explained using register variables. This was a very important finding as it suggests that mode-specific selection can be removed for employment status.
40. For the CVS variables, all total mode effects that were significant were due to measurement differences. When not significant, the measurement bias remained the largest of the three components. The coverage effect generally contributes the least to the total. In some instances the bias terms had different signs, indicating that the effects counteracted against each other. From all CVS estimated mode effects, we made the important observation that web and telephone are the contrasting modes, and paper and face-to-face are always in between. The web group outcomes are virtually always more negative with respect to the target variables. This group shows more offences, more victimization, feeling less safe, and scoring higher for nuisance. These differences seem to be mainly caused by measurement bias. The opposite applies to the telephone group, showing a more positive picture.
41. In-depth analyses were made for subpopulations based on age, gender, household composition, ethnicity, type of income and urbanisation level of the area of residence. They showed that web has a different impact on subpopulations from telephone and face-to-face, and, thus, the choice of mode also affects associations between survey target variables and socio-demographical variables. Both paper and web showed large frequencies of “don’t know” answers, if offered, and showed more non-differentiation in rating scale questions.
42. The results confirmed the observed, unexpected changes in the CVS time series, and were used as input to a major redesign of the CVS questionnaire. The lessons learned from this experiment for CVS are:
- Weighting removes only a small part of the mode effect;
 - Mode effects must be avoided by design and/or stabilized by calibration to fixed mode totals.
 - The mode effect decomposition for the LFS indicated a decrease in unemployment rate when web and telephone are added. These effects are not significant at the 5% level. For LFS we learned that:

- Weighting removes a large part of the mode effect;
- Remaining mode effect is difference in measurement.

IV. Discussion

43. Over the last seven years several projects and studies were conducted at Statistics Netherlands aiming at the implementation of web as a survey mode for large national surveys. The results from these projects and studies must be considered in the context of official surveys in The Netherlands. It is likely that some of the findings may be different for other types of surveys and within the survey climate of other countries.
44. The most influential early conclusion was that web cannot be used as a single mode for these surveys, despite the relatively high web access rate in The Netherlands. The response rates and recruitment rates in web are simply too low. Web is a potential single mode in panel studies, but it is not for cross-sectional surveys. Remarkably, the response rates to paper are much higher than to web. This finding implies that there still may be a lot to gain in the recruitment for web surveys, as also mailed questionnaires lack the interviewer interaction with the respondent.
45. The conclusion that web cannot be used as a single mode meant the start of a large and ambitious redesign of various surveys from single mode to mixed-mode designs. After various pilot studies and after some trial and error, it was concluded that mixed-mode designs are feasible but imply a more complex statistical process and monitoring of data collection. Response rates and representativeness of response to mixed-mode surveys are comparable to those of traditional face-to-face surveys and costs are considerably lower. The gain in costs, however, requires a significant investment in administration systems, monitoring procedures and pilot studies to derive optimal data collection strategies. Furthermore, the inclusion of web as a survey mode implies a redesign of survey questionnaires and may force questionnaires to be split or to be simplified.
46. It is sometimes conjectured that mixed-mode designs may be introduced to improve survey quality. In the various analyses that were conducted at Statistics Netherlands no indication was found that this hypothesis is true. Mixed-mode surveys may certainly solve undercoverage problems of web, but they do not lead to higher or more representative response and measurement differences between modes may be considerable.
47. The mode effects between web and the interviewer modes telephone and face-to-face can be large to an extent where they lead to unstable statistics. As long as the shares of modes to the total response are relatively stable and/or mode-specific measurement biases are relatively small, mixed-mode designs may provide comparable statistics over time. When contrasts between modes are large or response rates to web are volatile, survey designs including web may suffer from method effects in the month-to-month or year-to-year changes in statistics. These method effects may be avoided by calibration of the survey response to fixed mode distributions, by adaptive survey designs that tailor the choice of modes to subgroups, or by avoiding certain combinations of modes.
48. Open issues at Statistics Netherlands related to web surveys are the height and volatility of response rates, panel recruitment rates and household data collection. Research is on-going into methods to increase and stabilize web response rates, to increase recruitment rates for subsequent waves, to balance self versus proxy reporting in household web surveys and into implementation strategies for adaptive designs.

Acknowledgements: Many people were involved in the projects and studies referred to in this paper. We like to thank Menno Cuppen for his background information to the Project Herontwerp Persoonsenquêtes.

References

- Blanke, K., Luiten, A. (2012), ESSnet Project on Data Collection for Social Surveys using Multi Modes (DCSS), Paper for the UNECE Conference of European Statistics, Oct 31 – Nov 2, Geneva, Switzerland.
- Brakel, J. van den, Roels, J. (2009), Intervention analysis with state-space models to estimate discontinuities due to survey redesign, Discussion paper 2009039, Statistics Netherlands.
- Buelens, B., Van den Brakel, J. (2011), Inference in surveys with sequential mixed-mode data collection, Discussion paper 201121, Statistics Netherlands.
- Buelens, B., Laan, J. van der, Schouten, B., Brakel, J. van den, Burger, J., Klausch, T. (2012), Disentangling mode-specific selection and measurement bias in social surveys, Discussion paper 201211, Statistics Netherlands.
- Calinescu, M., Schouten, B., Bhulai, S. (2012), Adaptive survey designs that minimize nonresponse and measurement risk, Discussion paper, Statistics Netherlands.
- Cuppen, M., Laan, P. van der, Nunspeet, W. van (2011), Re-engineering Dutch social surveys: From single-purpose surveys to an integrated design, Paper presented at the ISI World Statistics Congress, Dublin, Ireland, 21-26 August.
- De Leeuw, E. (2005), To mix or not to mix? Data collection modes in surveys, *Journal of Official Statistics* 21, 1 – 23.
- Dex, S., Gumy, J. (2011), On the experience and evidence about mixing modes of data collection in large-scale surveys where the web is used as one of the modes in data collection, National Centre for Research Methods Review paper, National Centre for Research Methods, UK.
- Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., Messer, B.L. (2009), Response rate and measurement differences in mixed mode surveys using mail, telephone, interactive voice response (IVR) and the internet. *Social Science Research* 38, 1 – 18.
- Groves, R.M., Heeringa, S.G. (2006), Responsive design for household surveys: tools for actively controlling survey errors and costs, *Journal of the Royal Statistical Society: Series A*, 169, 439 – 457.
- Jäckle, A., Roberts, C., Lynn, P. (2010), Assessing the effect of data collection mode on measurement, *International Statistical Review* 78, 3 – 20.
- Klausch, T., Hox, J., Schouten, B. (2012), Assessing current beliefs about the mode-dependency of survey response and nonresponse bias, Discussion paper 2012xxx, Statistics Netherlands.
- Mars, G., Chadli, R., Janssen, B., Cuppen, M. (2012), Introducing web interviewing in the Labour Force Survey at Statistics Netherlands: a pilot, Discussion paper, Statistics Netherlands.
- Schouten, B., Brakel, J. van den, Buelens, B., Laan, J. van der, Klausch, L.T. (2013), Disentangling mode-specific selection and measurement bias in social surveys, *Social Science Research*, 42, 1555 – 1570.
- Schouten, B., Calinescu, M., Luiten, A. (2011), Optimizing quality of response through adaptive survey designs, Discussion paper 201118, Statistics Netherlands³.

³ All Statistics Netherlands' discussion papers are available at <http://www.cbs.nl/en-GB/menu/publicaties/reeksen/default.htm>

Wagner, J. (2008), Adaptive survey designs to reduce nonresponse bias, PhD thesis, University of Michigan, USA.