Distr. GENERAL
02 September 2013
WP 23
ENGLISH ONLY

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on Statistical Data Collection
(Geneva, Switzerland, 25-27 September 2013)

Topic (iv): Multiple modes of data collection

THE PHILOSOPHY OF FRENCH EXPERIMENTS ON INTERNET AND MIXED-MODE DATA COLLECTION

Working paper

Prepared by Tiaray Razafindranovona (Insee, France), Gaël de Peretti (Insee, France) and Alexis Barrau (ENS Cachan, France).

I. Introduction

1. The Insee, like most of the national statistics institutes, has to face a request of surveys always more demanding in quality terms, quality in the broad sense, with dimensions of accuracy, relevance, comparability, coherence, clarity and timeliness, in a general context of budgetary restriction (de Peretti and Razafindranovona, 2013). Besides, the complexity of surveys is also increasing to meet the social demand: particular populations (surveys on homeless persons, surveys on handicap and dependence) and specific protocols (cohorts, panel data, linked surveys on establishment/worker or helper/helped).

2. One of the envisaged solutions to face some of these challenges is the use of the internet as privileged or complementary mode of data collection. The objective would be to maximise the survey quality (Lyberg, 2012) in all the dimensions expressed previously, under constraints of budget. In particular, the reduction of data collection costs (budget and time) associated with such a mode is a strong incentive to engage into a global re-thinking of the households surveys' statistical production process.

3. But before generalizing the use of the internet to collect data on households, probably in a mixed modes way, some issues need a bit of investigation. Coverage, selection and measurement effects relevant to this mode have to be clearly understood to guarantee, in the future, results of high quality. For all of these reasons, Insee has decided to launch a series of experimental CAWI surveys.

II. Principles of French experiments on internet and mixed modes surveys

A. Experiments for each new households survey

4. The research literature on differences between modes is already quite significant and continues to expand. Some studies focus on the outcomes, taking into account the whole process of data collection, other studies focus on particular sources of error, like coverage or measurement errors

(satisficing, social desirability). But results on mode effects are quite difficult to generalize as they depend on the population being studied, the sampling frames, the questions asked or the data collection protocols.

5. Institutes can partially rely on the general results or recommendations of this rich literature to implement internet or mixed modes data collection in their processes. But if the institutes want to re-design existing surveys, by introducing new modes of data collection, these general guidelines could not solve all the issues: specific tests should be carried out. Therefore, Insee has decided that internet and mixed modes experiments would be made, survey by survey, in parallel of almost each new CAPI households survey to take into account specificities.

B. Standard protocol

6. In parallel of almost each households survey, an experimental survey with mixed modes data collection (internet and paper) is carried out with the following standard protocol.

1. Sample frame

7. People are drawn from a sample frame, with all elements of name and address available for mass mailing. Those elements are available in taxpayers files provided by tax services.

2. Questionnaire

8. The web questionnaire usually consists in a fitted subset of the CAPI questionnaire (of the main survey): its completion length should not be more than 30 minutes. The paper questionnaire can be even less complex as filters should be used cautiously. But questionnaires should allow comparisons on some key variables between these 3 modes.

3. First contact, notification letter

9. First, postal mails are sent to the selected households or individuals including a notification letter that gives the URL of the web questionnaire with username and password to log in. The mail also gives a presentation of Insee, the aim of the survey and explains how it is crucial to respond. There is no mention of the experimental dimension of the survey and response is said to be mandatory.

4. First reminder with paper questionnaire

10. After three weeks, if the household did not answer, another mail is sent, as a reminder. This mail also includes a paper version of the questionnaire and a courtesy reply envelope. So, from this date, the person can respond either by mail or on the internet.

5. Second reminder

11. Three weeks later, if there is still no answer, a reminder is sent but this time, without the paper questionnaire.

C. No interference with the main survey

12. As the nowadays top priority is the highest quality of the traditional face-to-face survey, experiments must not interfere with this: workload and sampling frames should take this order of priority into account.

13. Concerns about workloads mean that if units are involved in both the main traditional survey and the experiment, the priority is first to achieve quality results on time for the traditional survey.

14. One main issue is to minimise the overlap between the face-to-face and the experimental surveys. Draws have to be made in order that a household selected in the face-to-face survey sample (drawn in

the French Master Sample based on Census data) must not be selected in the experimental survey sample and vice versa.

D. Each experimental survey must test something new

15. Carrying out these experiments will increase knowledge in the field of internet data collection, for example about the design of the questionnaire, its optimal length, technological issues or about reachable response rates and representativeness. The accumulated experience will also be re-usable to optimize and standardize processes. So, these CAWI experiments will provide a better global understanding on mode effects and give some clues for each survey on how mixed modes could be (or not) implemented in the future.

16. Another part of the philosophy is that each experiment should test some new specific points not already handled. It could concern for example, the protocol of the survey and its respect or specific mode effects; some of these tests will be developed further.

III. Experiments of internet and mixed modes data collection

E. Internet in production process

17. This document focuses on experimental surveys with internet as a part of mixed modes data collection. But, in fact, internet is already used as a mode of data collection in the processes of some households surveys.

1. Survey for non respondents to LFS (from 2008)

18. This survey concerns households that have not responded to LFS or for whom contact has failed. These households receive a notification letter that invite them to answer by mail or on the internet. About 1,000 households respond by this way each trimester.

2. Drop off experiences replacing SILC complementary questionnaires (from 2011)

19. The notification letter is given to the household by the interviewer that comes for SILC. For example, in 2013, the topic of the questionnaire is about well-being.

3. Quality of life (2011)

20. This survey that mixes internet and paper data collection is an example of easy and cheap way to have a first look on a topic of interest.

4. ICT survey

21. Three sub-samples are used in the new protocol of this survey, where mixed modes help to face coverage problems:

    a. Subsample 1: CATI for selected households that have been found in the phone register

    b. Subsample 2: CAWI (+paper) for selected households not found in the phone register

    c. Subsample 3: CAWI (+paper) for selected households that have been found in the phone register. This subsample is a methodological sample that allows to adjust the two other samples.

F. Past experiments

1. Housing 2010

22. This experimental CAWI (+paper) survey has been compared to the Housing face-to-face survey of 2006. The main result is that even after controlling for many socio-demographical variables (age, sex, diploma, household size, professional situation and income) by calibration techniques, people that respond on the internet are less satisfied about their housing conditions (Amiel and Denoyelle, 2012).

Table 1. Comparison between internet and face-to-face on Housing surveys (after calibration)

	Housing 2010 (internet experiment)	**Housing 2006** (face-to-face)
Suffer from cold in winter	21%	15%
Well exposed accommodation	73%	82%
Good neighbourhood relationships	63%	77%
Safety in the district	60%	74%
Public transport accessibility	46%	55%
Store nearness	53%	63%

Amiel and Denoyelle (2012)

G. Current experiments

1. Safety Survey (SASU) 2013

23. This is the first French internet (+paper) experimental survey on victimization and issues exist on how people are reluctant or not to respond to such a questionnaire on the internet, as it includes quite sensitive topics like sexual violence (outside the household).

24. Another major aim of this experiment is to measure the potential bias of an internet survey on victimization. The problem with such surveys is that we can fear strong issues of self-selection where potential victims belonging to the same household of selected individuals respond instead of these targeted individuals. This could happen as it is not really possible to control the process of selection like in face-to-face surveys. Therefore, some European experiences show that victimization levels are higher among those using CAWI than other modes. This experiment allows to test some ways to control this problem of selection.

25. Two protocols were carried out to control the problem of selection:

    a. A first sample of individuals receive a notification letter with a unique couple of login and password. The questionnaire includes a table where the respondent can list all the individuals of the household and the victimizations for each of them (robberies with violence, robberies without violence, physical violences and threats). By introducing these questions on all the members of the household, the aim is to maximise chances that the selected individual responds since all potential victims could declare their victimizations in the table.

    b. A second sample of individuals receive a notification letter with two logins and two passwords. The selected individual is asked to answer the questionnaire with the first couple of login/password and he can give the second couple of login/password to another person in the household. The aim here is also to maximise chances that the selected individual responds since another potential victim in the household could also answer to the questionnaire.

26. Concerning these two samples, responses rates are about 42% with about half of these responses by internet, the other half on paper.

2. Working Conditions Survey 2013, see IV

H. Future experiments

1. Housing 2014

27. One key financial variable in the Housing survey is the rent paid by the household. In face-to-face, the interviewer can strongly insist for the individual to consult external paper documents in order to give the accurate numbers. In self-administered questionnaires there is no such an incentive to search for the external documents. The main objective of this experimental survey on the internet (+paper) is to test whether it is possible or not to make correct estimates on financial variables like rental amount when questionnaires are self-administered.

28. As some concepts of the questionnaire are quite complex for the respondent without the help of an interviewer, a special attention will be given in the internet version of the questionnaire to the instructions and the dynamic tooltips.

2. Wealth 2014

29. In the face-to-face version of this survey, an important part of the questionnaire focuses on the value of real estate property. In particular, if the household owns its accommodation, the respondent is asked to give an estimation of its value. Our hypothesis is that the presence of the interviewer in the accommodation limits huge deviations from the market value as the interviewer can see the property, which the respondent is aware of.

30. When the response is made on the internet, there is no such a social control that prevents from giving a false estimation. So, the main objective of this experiment is to test if such a bias exists and if so, to measure its magnitude.

31. As the Wealth survey is a longitudinal survey, the internet will also be used as a tool of the production process in the follow-up of households. The question is about how the internet could be used to maintain a contact with an already surveyed household to prevent from attrition in future waves of the survey.

3. Census

32. A dedicated project called ‘Homere’ works on a re-design of the protocol where internet would be one of the options to answer.

4. Labour Force Survey

33. Also a dedicated project whose aim will be to think about how internet could be introduced as one of the modes of data collection in this recurrent and longitudinal survey.

IV. Focus: Working Conditions experimental survey

I. Presentation of Working Conditions survey

34. The face-to-face survey addresses many topics surrounding working conditions: working hours and organization of working time, risks, work hardness and its prevention, psychosocial constraints, violence at work, relations with the public. The auto-administered part of the face-to-face interview (audio-CAPI) allows to quantify the exposure to psychosocial risks.

35. The experimental parallel Working Conditions survey, called ‘Quality of life at work’, tackles the same topics: it is more or less (some questions are completely new) a shorter version of the questionnaire adapted to the internet.

36. The sample of this experimental survey is drawn in a base built on taxpayers files; only people that have declared earnings from work can be selected. If this solution is probably the best one available to draw individuals of our target population, some drawbacks exist. Most of issues are due to the time lag between the date of the frame (tax returns made in 2011 on earnings of 2010) and the date of the survey (2012/2013). Individuals that were not working in 2010, but work at the date of the survey are not covered: unemployed finding a job, young people entering the active population and other inactives (housewives, disabled persons) with changes in their status. Another issue is that individuals could be hard to reach when they have moved after 2010: it could partially explain the large rate of owners of their accommodation (who move less than tenants) in the population of respondents (table 5).

J. Aims of the experimentation and protocols

37. This experimental survey allows to compare:

    a. Response rates between a single mode (internet) data collection protocol and a mixed modes (Internet + paper) data collection protocol

    b. Quality and content of responses between the traditional survey and the experimental survey, and more broadly what is the magnitude of mode effects

    c. Responses between two protocols involving different kinds of notification letters

38. Another aim of the experimentation is to test some questions of the future survey on psychosocial risks (2015). There are also some open questions included in the questionnaire that could raise topics for this future survey.

39. Two models (A & B) of notification letters differ by the way they focus or not on psychosocial risks. The model A letter informs that the survey allows to obtain a concrete description of work and its organization, on different dimensions: hours, rhythms, atmosphere at work. The model B letter informs that the survey allows to obtain a description of work, its organization and psychosocial risks that are sometimes associated with it (stress, lack of recognition, insecurity …). Half of the sample receive the model A letter, the other half, the model B letter.

40. For half of the sample, the protocol is the standard one described earlier (II B.): first a notification letter, then a reminder with a paper questionnaire and then another reminder without the paper questionnaire. For the other half of the sample, only internet is available as a mode of response: the paper version of the questionnaire is never sent.

41. 40,000 individuals were drawn and 4 subsamples of 10,000 individuals were made:

    a. Subsample 1 - model A letter, internet and paper available as modes of response

    b. Subsample 2 - model B letter, internet and paper available as modes of response

    c. Subsample 3 - model A letter and only internet available as mode of response

    d. Subsample 4 - model B letter and only internet available as mode of response.

K. Response rates

42. Before the 1st reminder (3 weeks after the beginning of the data collection)

Table 2. Responses before the 1st reminder

	Internet response	Paper response	Moved (*)	No response	Total
Subsample 1	1,678	2	606	7,700	10,000
Subsample 2	1,627	1	576	7,785	10,000
Subsample 3	1,651		637	7,678	10,000
Subsample 4	1,612		625	7,730	10,000
Total	**6,568**	**3**	**2444**	**30,893**	**40,000**

Insee - Quality of life at work (2013)

(*) Moved : Does not reside at this address

43. Before the 2nd reminder (3 weeks after the 1st reminder)

Table 3. Responses before the 2nd reminder

	Internet response	Paper response	Moved	No response	Total
Subsample 1	2,283	863	1,006	5,810	10,000
Subsample 2	2,209	851	955	5,943	10,000
Subsample 3	2,789		1,055	5,830	10,000
Subsample 4	2,740		981	5,953	10,000
Total	**10,021**	**1,714**	**3,997**	**23,536**	**40,000**

Insee - Quality of life at work (2013)

44. At the end of the data collection

Table 4. Responses at the end of the data collection

	Internet response	Paper response	Moved	No response	Total
Subsample 1	2,924	1,527	1,164	3,796	10,000
Subsample 2	2,756	1,638	1,125	3,863	10,000
Subsample 3	3,738	1	1,247	4,055	10,000
Subsample 4	3,751	1	1,178	4,046	10,000
Total	**13,169**	**3,167**	**4,714**	**15,760**	**40,000**

Insee - Quality of life at work (2013)

45. For this experiment, at the end of the data collection, the global response rate is 46% (moved individuals excluded, representing 12% of the sample). Just before the 1st reminder this rate is 17% and at the 2nd reminder, 33%.

46. For the half of the sample (subsamples 1 and 2) with internet and paper modes available, the response rate at 1st reminder is 18%, at 2nd reminder 34% and the final response rate, at the end of the data collection, is 50%. Internet represents about two thirds of the responses collected.

47. Response rates are weaker for the other half of the sample where internet is the only option. No difference at 1st reminder, but then the gap increases over time : 31% at 2nd reminder and 43% at the end of the data collection.

48. There are no significant effects of the model of notification letter sent on the global response rates. But investigations will have to be made to see whether this has an impact or not on the content of the responses.

49. The major feedback on data collection was that some people complain about the fact that internet was the only option to respond, which was the case for the subsamples 3 and 4 and also before the 1st

reminder (the notification letter does not mention the availability of a paper version of the questionnaire). People could not know that the survey is an experiment and some of them have been hurt to see that an official survey systematically excludes those who don't have any internet access. On very rare and special occasions, paper questionnaires were sent if they were asked even if it was out of the protocol (only 3 respondents on paper for subsamples 1 and 2 before the 1st reminder and 2 respondents for subsamples 3 and 4).

L. First results : characteristics of respondents, comparison between paper and internet

1. Comparison between sample, internet respondents and paper respondents on sample frame variables

50. Variables of sample frame allow us to compare individuals from the sample (kind of "general" population) with those who have answered on the internet or by paper. There are more aged people among paper respondents than among internet respondents or in the sample (those 2 populations have almost the same average age). The proportion of women is larger among paper respondents than among internet respondents and these proportions are themselves higher than in the sample. Owners are more numerous, in proportion, among internet respondents and paper respondents than in the sample. And earnings of paper respondents don't differentiate too much from individuals of the whole sample. On the other hand those who respond on the internet have larger earnings.

Table 5. Comparisons on sample frame variables

	Sample	Internet respondents	Paper respondents
Age (mean)	44.4	44.7	47.1
Women (proportion)	50%	52%	55%
Owners (housing occupation status, proportion)	59%	69%	67%
Annual earnings from work (median)	18,500 €	21,150 €	18,150 €

Insee - Quality of life at work (2013)

2. Occupations and education of internet and paper respondents

51. Some comparisons between internet and paper respondents on survey variables also show that these subpopulations are different. As a sum up, internet respondents are younger, richer, with a better level of education and are more frequently managers or professionals than paper respondents.

Table 6. Comparisons on survey variables

	Internet respondents	Paper respondents
Occupations (ISCO)		
Managers	8%	4%
Professionals	24%	14%
Technicians and associate professionals	23%	21%
Clerical support workers	12%	13%
Service and sales workers	13%	17%
Skilled agricultural, forestry and fishery workers	1%	2%
Craft and related trades workers	6%	9%
Plant and machine operators, and assemblers	6%	9%
Elementary occupations	6%	9%
Armed forces occupations	1%	1%
Levels of education (ISCED)		
At most lower secondary education	14%	25%
Upper secondary education	39%	46%
Tertiary education	47%	29%

Insee - Quality of life at work (2013)

52. Such differences show that if we want to study mode effects on key variables about working conditions, we would have probably first to control for selection effects before inferring conclusions from comparisons.

53. One way to control for selection effects is to implement econometric techniques of evaluation of treatment (Lugtig et al., 2011) like kernel matching on propensity score (Heckman et al., 1998). Variables used in this control are sex, age, living as a couple, diploma, socio-professional categories, having a part-time or a full-time job and earnings from activity. These variables should be insensitive to the mode of data collection but correlated both with the selection and the variable(s) of interest.

54. We can give two examples of questions where there are significant gross differences on responses between internet and paper respondents but no mode effects detected when we control the selection:

    a. 'At work, do you breathe smokes or dust?' 26% of paper respondents answer positively vs 21% of internet respondents but this difference is only due to selection effects.

    b. 'Would you be happy if the professional activity of one of your children is the same as yours?' 43% answer positively among internet respondents, 35% among paper respondents but here too, this is only due to selection effects.

3. Examples of mode effects

55. One concern about mode effects is satisficing (Krosnick, 1991), especially when auto-administered questionnaires are involved. We study one of the numerous forms of satisficing which is the tendency to differentiate or not on responses scale. Before comparing face-to-face with auto-administered responses (Internet + paper), first, the two auto-administered modes of data collection that are internet and paper are compared on this differentiation dimension with a control of selection effects (with kernel matching on propensity score). To quantify this dimension, we compute an index of response differentiation (McCarty and Shrum, 2000) on some blocks of questions. For example, one of these blocks consists in 11 questions about different forms of pressure at work and 4 responses were available: always, often, sometimes or never. Our conclusion is that differentiation on the scale of responses seems to be more pronounced on the internet than on paper: respondents on the internet use more of the different response options.

56. Another conclusion of our comparisons is that, even after controlling of selection effects, emotional well-being is higher in paper responses than internet ones. To measure this well-being, we use responses about the degree to which the following feelings were present: 'I have felt cheerful and in good spirits', 'I have felt calm and relaxed', 'I have felt active and vigorous', 'I woke up feeling fresh and rested' and 'My daily life has been filled with things that interest me'. These responses allow to impute the WHO-5 well-being index, a score from 0 to 100, that was developed by the World Health Organization to capture emotional well-being (de Wit et al., 2007). We find that mode effects are quite significant with a magnitude of about 3 points.

V. Conclusion

57. The use of the internet in the processes of households surveys will probably expand in the next years. Not only for budgetary reasons but also because this *new* mean of communication could help to reach some populations that are missed by *traditional* ways of surveying: institutes have to adapt to major technological changes to be in line with the population they want to describe. But these adaptations must not be forced without some prior evaluation in terms of data quality.

58. This paper gives a brief overview of the general philosophy of French experiments on mixed modes surveys and provide some examples and first results. Further results about mode effects will be given later with cautious comparisons of these experiments with parallel traditional surveys. These results on mode effects will certainly be helpful if institutes want to re-design their households surveys' statistical production process. These experiments will also give a clearer picture about which kind (length, topic) of surveys could be answered on the internet and how internet could be implemented

in the processes of data collection. In fact, for households surveys, the rationale could be a sequential mixed modes way by using first cheap modes and at the end expensive modes, but this standard scheme does not probably fit to a lot of long and complex surveys designed to cope with society requests.

References

Amiel, M. H,. and Denoyelle, T. (2012), Enquêtes en ligne : comparaison de modes de questionnement sur le thème du logement, *XIèmes Journées de Méthodologie Statistique de l'Insee.* http://jms.insee.fr/files/documents/2012/945_4-JMS2012_S26-3_AMIEL-ACTE.PDF

de Peretti, G., and Razafindranovona, T. (2013), Les enquêtes multimode : multi-solution ou multi-problème ?, *45e Journées de Statistique de la SFDS.* http://papersjds13.sfds.asso.fr/submission_192.pdf

de Wit, M., Pouwer, F., Gemke, R., Delemarre-van de Waal, H. and Snoek, F. (2007), Validation of the WHO-5 Well-Being Index in Adolescents With Type 1 Diabetes**,** *Diabetes care, 30.*

Heckman, J., Ichimura, H., and Todd, P. (1998), Matching as an Econometric Evaluation Estimator, *Review of Economic Studies,65.*

Krosnick, J. (1991), Response strategies for coping with the cognitive demands of attitude measures in surveys, *Applied Cognitive Psychology. Vol 5.*

Lugtig, P. J., Lensvelt-Mulders, G. J., Frerichs, R., and Greven, F. (2011), Estimating nonresponse bias and mode effects in a mixed mode survey, *International Journal of Market Research, 53(5).*

Lyberg L. (2012), La qualité des enquêtes, *Techniques d'enquête, 38(2).*

McCarty, J. A., and Shrum L. J. (2000), The measurement of personal values in survey research: A test of alternative rating procedures, *Public Opinion Quarterly, 64.*