American Journal of Theoretical and Applied Statistics
2022; 11(4): 114-121
http://www.sciencepublishinggroup.com/j/ajtas
doi: 10.11648/j.ajtas.20221104.12
ISSN: 2326-8999 (Print); ISSN: 2326-9006 (Online)

Adaptive Survey Design for the Dutch Labour Force Survey

Kees van Berkel

Statistics Netherlands, Division of Data Services, Research and Innovation, Heerlen, the Netherlands

Email address:
cam.vanberkel@cbs.nl

To cite this article:
Kees van Berkel. Adaptive Survey Design for the Dutch Labour Force Survey. *American Journal of Theoretical and Applied Statistics*. Vol. 11, No. 4, 2022, pp. 114-121. doi: 10.11648/j.ajtas.20221104.12

Received: July 19, 2022; **Accepted**: August 23, 2022; **Published**: August 31, 2022

Abstract: A challenge for the National Statistical Institutes is to produce reliable statistics with a limited budget for data collection. During the past years, many surveys at Statistics Netherlands were redesigned to reduce costs and to increase or maintain response rates. From 2018 onwards, adaptive survey design has been applied in several social surveys to produce more accurate statistics within the same budget. In previous years, research has been done on the impact on quality and costs of reducing the use of interviewers in mixed-mode surveys that start with Internet observation, followed by telephone or face-to-face observation of Internet nonrespondents. Reducing follow-ups can be done in different ways. By using stratified selection of people eligible for follow-up, nonresponse bias may be reduced. The main decisions to be made are how to divide the population into strata and how to compute the allocation probabilities for face-to-face and telephone observation in the different strata. For this purpose, a methodology has been developed in this paper. The methodology is applied in the development of an adaptive survey design for the Dutch Labour Force Survey. Attention is paid to the survey design, in particular the sampling design, the data collection constraints, the choice of the strata for the adaptive design, the calculation of follow-up fractions by mode of observation and stratum, the practical implementation of the adaptive design, and some response and survey results.

Keywords: Balanced Response, Nonresponse Bias, Accuracy, Data Collection

1. Introduction

Adaptive survey design aims to get a better balanced response by putting different effort in different groups of the population. It is deployed in improving survey results, or reducing survey costs. The terms responsive and adaptive survey design have been introduced and described by Groves and Heeringa [9] and Wagner [21]. The designs have attracted a lot of interest in recent years due to budgetary constraints and declining response rates [5-7, 14, 15]. Adaptive survey design is based on the premise that the optimal approach strategy is not the same for every person. For example, the use of incentives may increase the tendency to respond in some people, but not in others [8]. In a telephone survey study, the days and times of day in which to call sample units are tailored, using auxiliary information [22]. In another telephone survey, census data were used to assign sampled people with low response propensity to more experienced interviewers [13]. In the United States, case studies have been conducted with propensity-based assignment to interviewers, propensity-based stopping of cases, performance-based phase duration, use of incentives, estimating response propensity for the next contact attempt and stratifying the sample for the next phase of the survey. [16].

Adaptive survey designs have four main elements: quality and cost objectives, design features, stratification of target population, and an optimisation and implementation strategy [16, 20]. Recent research is focusing on quality criteria and their uncertainty by learning from historical data [18, 19], and on decision rules for optimisation [11-13].

This paper is mainly about the latter aspect, answering how to optimise the adaptive survey design within the context of an official survey with strong cost-quality differences in the design features. Here the quality indicator is the coefficient of variation of response probabilities, the design features are the follow-up modes, stratification is done with auxiliary variables from relevant registers and the quality indicator is minimised under specified restrictions. Methodology for surveys with a simple random sampling design and one follow-up mode with an application to the Dutch Health Survey was developed by Van Berkel et al. [2]. In this paper, the methodology is extended to surveys with a sampling design with unequal selection probabilities and two follow-up modes. The developments are applied to the Dutch Labour Force Survey.

As of 2018, adaptive survey design has been introduced step by step in redesigns of social surveys at Statistics Netherlands. In 2018 it has been implemented in the Health Survey and the Public Opinion Survey, in 2019 in the Life Style Monitor and the Leisure Omnibus, in 2021 in the Labour Force Survey, and in 2022 in the Social Coherence Survey.

The paper reads as follows. Section 2 contains the methodological aspects of adaptive survey design. Section 3 deals with the development and implementation of the adaptive survey design for the Labour Force Survey. Section 4 contains the conclusion and Section 5 ends with a discussion.

2. Methodology

In this section the four main elements of adaptive survey design are discussed: quality indicators, design features, clustering the population, and optimisation.

2.1. Quality Indicators

Consider a finite population of N people, labelled by $k = 1, 2, \ldots, N$. For the y, a probability sample with size n is drawn from the population, such that each person k has a positive inclusion probability π_k. Let a_k be the inclusion indicator for person k. This means that a_k is equal to 1 if person k is selected in the sample, and 0 if not. The expected value of a_k is equal to the probability that person k is selected in the sample, $E(a_k) = \pi_k$.

A random response model is adopted, where each person k in the target population is assumed to have a response probability ρ_k, which is only known to person k. If person k is selected in the sample, this person is subjected to a Bernoulli experiment that results in response with probability ρ_k and in nonresponse with probability $1 - \rho_k$.

Let r_k be the response indicator for person k. So r_k is equal to 1 if person k responds and 0 if person k does not respond. The expected value of r_k is equal to the probability that person k responds, $E(r_k) = \rho_k$. The number of respondents r in the survey is a random variable $r = \sum_{k=1}^{N} a_k r_k$ with expected value $E(r) = \sum_{k=1}^{N} \pi_k \rho_k$. Note that in case of simple random sampling $\pi_k = n/N$ for all k and then $E(r) = n \cdot \bar{\rho}$ with $\bar{\rho} = \frac{1}{N} \sum_{k=1}^{N} \rho_k$ the population response mean.

The aim of the survey is the estimation of population means for several target variables. An estimator of the population mean $\bar{Y}$ of variable Y is the modified Horvitz-Thompson estimator,

$$\bar{Y}_{mHT} = \left(\sum_{k=1}^{N} a_k r_k Y_k / \pi_k\right) / \left(\sum_{k=1}^{N} a_k r_k / \pi_k\right). \quad (1)$$

Note that in case of simple random sampling, $\bar{Y}_{mHT}$ is equal to the mean of the observed values for the target variable among respondents. In general the estimator $\bar{Y}_{mHT}$ is biased, unless $\rho_k = \bar{\rho}$ for all k. Bethlehem [3] shows that

$$B(\bar{Y}_{mHT}) = E(\bar{Y}_{mHT}) - \bar{Y} \approx \frac{1}{\bar{\rho} N} \sum_{k=1}^{N} (\rho_k - \bar{\rho}) Y_k = \frac{1}{\bar{\rho}} cov(\rho, Y) = \frac{R(\rho, Y)\, S_\rho\, S_Y}{\bar{\rho}}. \quad (2)$$

Here $cov(\rho, Y)$ is the population covariance between the response probabilities and the values of the target variable, $R(\rho, Y)$ is Pearson's correlation coefficient, S_ρ is the standard deviation of the response probabilities and S_Y is the standard deviation of the values of the target variable.

From this expression it follows that there is no bias if there is no correlation between response propensity and target variable. The smaller the variation in response probabilities or in the values of the target variable, the smaller the bias. And the higher the mean population response rate, the smaller the bias.

Since Pearson's correlation coefficient does not exceed 1 in absolute value, an upper limit for the bias is given by:

$$|B(\bar{Y}_{mHT})| \leq \frac{S_\rho\, S_Y}{\bar{\rho}} = CV(\rho)\, S_Y. \quad (3)$$

Here $CV(\rho) = S_\rho / \bar{\rho}$ is the coefficient of variation of the response probabilities. Since S_Y is a population parameter that cannot be influenced, in the remainder of this paper it is attempted to minimise this upper limit for the bias by minimising $CV(\rho)$ by intervening in the data collection process.

2.2. Design Features

Consider the mixed-mode strategy CAWI → CATI/CAPI with Computer-Assisted Web Interviewing (CAWI) as starting mode, and follow-up of CAWI nonresponse by a combination of Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Personal Interviewing (CAPI). The design features to adapt are the CAPI and CATI follow-ups.

In this mixed-mode strategy, all sampled people are first asked by letter to complete a questionnaire on the Internet. People who have not responded to this request after no more than two reminders, are contacted by telephone if a telephone number is known at the office, otherwise they are visited at home to conduct an interview. In ASD the entire sample starts with CAWI and the observation strategy of the follow-ups is adjusted as follows. To reduce the variation of response rates, more CATI/CAPI is used for groups that are less likely to respond via CAWI, and less CATI/CAPI is used for groups that are more likely to respond via CAWI. The identification of these so-called target groups is carried out using cluster analysis.

Observe that the observation strategies CAWI → CAPI and CAWI → CATI are special cases of the CAWI →

CATI/CAPI strategy where, in the first case the CATI-follow-ups have been set to zero, and in the second no CAPI-follow-ups are conducted.

2.3. Clustering the Population

Determining target groups is also called segmentation or clustering the population. The target groups are composed by means of response propensities of people per mode. This may mean that two target groups have approximately the same response rate at CAWI, but that their CATI or CAPI response rates differ. It is also possible that the total response rates of two target groups are approximately the same, but that their response rates differ per mode.

Clustering is performed in two steps. First a classification tree algorithm is applied, dividing people into groups based on personal characteristics. The algorithm uses the characteristics that explain the response most, and it divides each selected characteristic into categories. Second k-means clustering is applied with the selected characteristics and corresponding categories yielded by the classification tree. This is a method that divides data into groups which are homogeneous according to response probabilities, where outliers can be detected. The advantage of this method is that small groups with extremely high or low response rates can be identified as target groups. These target groups can be assigned a separate approach strategy. A disadvantage of the k-means method may be that the target groups are less homogeneous according to the characteristics used.

2.4. Optimisation

Consider a survey with the mixed-mode strategy CAWI → CATI/CAPI. Let G be the partition of the population used to determine the target groups, where in each group $g \epsilon G$ the response probabilities do not differ much from one person to another. Each target group is the union of one or more groups from G. For each $g \in G$, let $N(g)$ denote the population size of group g and $n(g)$ the sample size of group g. It is assumed that within each group g all people have the same inclusion probability $\pi(g)$.

The response mean $\bar{\rho}$ is estimated by the Horvitz-Thompson estimator

$$\bar{\rho}_{HT} = \frac{1}{N}\sum_{k=1}^{N} a_k r_k/\pi_k = \frac{1}{N}\sum_{g\in G} r(g)/\pi(g), \qquad (4)$$

where $r(g) = \sum_{k=1}^{N(g)} a_k r_k$ is the estimator for the number of responses in group g.

For each group $g \in G$, let $p_w(g)$ be the CAWI-response probability, $p_e(g)$ the probability of sampled people of being eligible for follow-up, $p_t(g)$ the CATI-response probability and $p_p(g)$ the CAPI-response probability. Let $f_t(g)$ and $f_p(g)$ be the fractions of CAWI-nonrespondents eligible for follow-up to be approached by telephone and face-to-face respectively in group g. So the response probability in group g equals

$$p(g) = p_w(g) + p_e(g)\left(f_t(g)p_t(g) + f_p(g)p_p(g)\right). \qquad (5)$$

Since $r(g) = n(g)p(g)$, this allows the mean response probability and the population variance of the response probabilities to be estimated:

$$\bar{\rho} = \frac{1}{N}\sum_{g\in G} n(g)\, p(g)/\pi(g), \qquad (6)$$

$$S_\rho^2 = \frac{1}{N}\sum_{g\in G} N(g)\,(p(g) - \bar{\rho})^2. \qquad (7)$$

The following problem needs to be solved.

Minimise $CV(\rho) = S_\rho/\bar{\rho}$ under specified constraints. (8)

Different types of constraints can be used. For instance constraints on budget, interviewer capacity, response numbers, response rates, or ratio of modes in response.

One CATI sampling fraction and one CAPI sampling fraction is used per target group. The decision variables for which the minimum can be found under the specified constraints, are the CAWI sample size n, the inclusion probabilities $\pi(g)$ and the CATI and CAPI sampling fractions $f_t(d)$ and $f_p(d)$ per target group d. Determining the partition G is a crucial part in designing the survey, and it is the starting point for clustering the population.

The minimisation problem requires a search for the numbers of people to be approached by target group and observation mode. The lower the CAWI response propensity of a target group, the more telephone and/or face-to-face observation is applied. This may lead to a smaller variation of response rates, and the ratio of the target groups in the response may be more similar to the ratio of the target groups in the population. This may be at the expense of the overall response rate.

The minimisation problem is solved with the Auglag function of the Alabama R package [1]. This R package uses the "Augmented Lagrangian Adaptive Barrier Minimisation Algorithm for optimising smooth nonlinear objective functions with constraints".

3. Adaptive Survey Design for the Dutch Labour Force Survey

3.1. Observation Strategy and Sampling Design

The Labour Force Survey (LFS) aims to provide statistics about participation of the Dutch population in the labour market. Core indicators are unemployment rate, participation rate, and job characteristics.

The survey applies a rotating panel design with five waves at three-monthly intervals. The observation strategy for the first wave is CAWI → CATI/CAPI, with different CATI- and CAPI-sampling fractions per target group. The observation strategy for the subsequent waves is CAWI → CATI.

The sampling design for the first wave is a stratified two-stage sample of people aged 14-89 years with unequal probabilities. The first stage is a stratified systematic sample of municipalities selected with probabilities proportional to their number of inhabitants. The stratification is COROP area. The desired number of people to be selected in the second stage per municipality is one. For self-selecting municipalities, this number is adjusted to the product of the sampling fraction and the population size of the municipality concerned. The second stage is a simple random sample of people aged 14-89 years in the selected municipalities with numbers as determined in the first stage. Initially, the design is such that every person in the target population has the same probability of being selected in the sample, and the numbers of sampled people per COROP area are proportional to the population sizes per COROP area. Thereafter, in order to improve the precision of unemployment figures, people registered at the Netherlands Employees Insurance Agency (EIA) as a job seeker are overrepresented. Non-western migrants and 15 to 24 year olds are overrepresented. People aged 65 or over and 14 year olds are underrepresented. The relative inclusion probabilities are included in Table 1.

Table 1. *Relative inclusion probabilities, first wave.*

EIA	ethnicity	age				
		14	**15-24**	**25-64**	**65-74**	**75-89**
no	non-western	1/4	3/4	3/4	1/4	1/8
no	western	1/4	3/4	1/2	1/4	1/8
yes	all	1	1	1	1	1

3.2. Development of Adaptive Survey Design for First Wave

Developing the adaptive design starts with identifying the target groups. The characteristics used in the sampling design are strongly associated with both response propensity and the target variable unemployment. Therefore, the target groups have been defined according to the characteristics used in the sampling design: registered at the EIA as a job seeker (no, yes), ethnicity (non-western, western), and age (14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-89). With these characteristics, the population can be split up into 32 categories. After merging some near-empty ones, 24 remain.

Using 2019 LFS-pilot data, response probabilities per mode of observation and likelihoods of availability of a telephone number can be estimated for each of the 24 categories. By means of k-means clustering, the eight target groups in table 2 were composed.

Table 2. *Target groups LFS.*

EIA	ethnicity	age							
		14	**15-24**	**25-34**	**35-44**	**45-54**	**55-64**	**65-74**	**75-89**
no	non-western	8	3	1	2	2	2	2	6
no	western	8	5	4	6	7	8	8	6
yes	non-western	3	3	1	3	2	2	7	7
yes	western	3	3	4	4	6	7	7	7

With these target groups, the coefficient of variation of response probabilities was minimised under the constraints:

1) A maximum of 50 percent of all CAWI nonrespondents may be followed up via CATI or CAPI.
2) A maximum of 40 percent of all follow-ups may be assigned to CAPI.

For this purpose, the Alabama solver in R was applied with different random initial values for the CATI and CAPI follow-up fractions per target group, since the algorithm can stop at a local minimum. The optimal solution is the solution with the lowest coefficient of variation. Table 3 shows the follow-up fractions of the optimal solution, yielding a coefficient of variation of response probabilities of 0.06 and mean response rate of 44.3%.

If half of the CAWI nonrespondents eligible for follow-up were randomly selected for follow-up, of which randomly three fifth assigned to CATI and two fifth to CAPI, the coefficient of variation of response probabilities would be 0.186 and the mean response rate 45.0%.

In Table 3, the columns p CAWI, p CATI, p CAPI, and p tot show the expected response rates for CAWI, CATI, CAPI, and the total adaptive strategy. The columns f CATI, f CAPI and f tot represent the CATI, CAPI and total follow-up fractions as a percentage of the CAWI-nonrespondents eligible for follow-up.

Table 3. *Response probabilities and selection fractions per target group LFS.*

group	p CAWI	p CATI	p CAPI	f CATI	f CAPI	f tot	p tot
1	14.4	11.9	34.3	0	96.9	96.9	42.5
2	21.2	19.3	32.6	51.3	48.7	100.0	41.3
3	21.5	20.6	40.6	0	68.1	68.1	42.9
4	28.4	25.1	42.9	31.6	31.5	63.1	43.5
5	32.1	31.1	45.3	56.9	0	56.9	43.9
6	34.1	31.1	44.8	47.3	0	47.3	43.6
7	39.9	31.8	37.6	22.9	0	22.9	44.2
8	47.9	36.0	40.4	0	0	0	47.9
tot	34.0	28.1	37.4	30.0	20.0	50.0	44.3

Table 3 shows that in target groups 1 and 2, with the lowest CAWI response rates, almost everyone is followed up. In target group 3, also with a low CAWI response rate, 68.1% of CAWI-respondents eligible for follow-up are selected for CAPI-follow-up because of the relatively high CAPI response rate in this group. In target groups 5, 6, and 7, no CAPI-follow-up is used because a sufficiently high response can be obtained by CAWI and CATI. In group 8, no follow-up is needed at all due to relatively high CAWI response rate in this group.

3.3. Practical Implementation of Adaptive Survey Design

To determine the CAWI sample size for the first wave, it is required that the fifth wave yields 4068 respondents every month. The conversion of this response target into a sample size for the first wave depends on a large number of parameters. The parameters for waves 1 and 2 were estimated based on the LFS pilot 2019-2020. For the subsequent waves, probabilities were estimated using empirical data from EU-SILC and the regular LFS. With these estimates, it was calculated that a sample of 3154 people should be approached weekly via CAWI.

In practice, there are some complications. Firstly, estimated response rates will differ from realised response rates; secondly, fewer telephone numbers may be available than expected; thirdly, sample sizes for CATI and CAPI are fixed per month, due to scheduling of interviewer capacity. With estimates according to the adaptive design of the previous section, follow-up sample sizes are set at 616 and 410 people per week for CATI and CAPI respectively. For CATI and CAPI, each month consists of follow-ups of four or five CAWI week samples. This means that in a month with follow-ups of four CAWI week samples, the CATI and CAPI samples contain 2460 and 1640 people, and in a month with follow-ups of five CAWI week samples these numbers are 3075 and 2050.

Next it is explained how the CATI and CAPI samples are drawn each month. CAWI nonrespondents within the same target group are selected for follow-up with as close to equal probabilities as possible, sticking to the agreed upon CATI and CAPI sample sizes. A two-stage procedure is followed. In the first stage selections for CATI and CAPI by target group are made in accordance with the computed adapted survey design. In the second stage the selections are fitted to the agreed upon sample sizes per mode.

Stage 1, per target group: Split the eligible CAWI nonrespondents into CATI-eligible and CAPI-eligible people, where the groups are as close to proportional in size to the required selection fractions as possible. Nonrespondents with 'best' available phone numbers are marked as CATI-eligible, the rest is CAPI-eligible. Select people from both groups with equal probabilities. If there is a lack of available telephone numbers, the CAWI-nonrespondents cannot be divided into two groups proportional to the CATI- and CAPI- follow-up fractions of the adaptive design. In this case, all CAWI- nonrespondents with telephone number are considered CATI-eligible, and selection fractions from both pools of eligible people are adjusted such that 1) the total follow-up fraction of the CAWI-nonrespondents is equal to the total follow-up fraction as included in the adaptive survey design, and if possible 2) the CATI-follow-up as a percentage of all nonrespondents is as close to the one given in the design. As a consequence, the CATI- and CAPI-selection fractions differ in this way.

Stage 2: Merge all selections for CATI follow-up, and merge all selections for CAPI follow-up. If both merged selections are smaller or greater than the agreed sizes, then both selections are amended by simple random sampling of the remaining eligible CAWI nonrespondents, or randomly removed. If exactly one merged selection is larger than the agreed size and the other smaller than the agreed size, then an attempt is made to transfer randomly selected elements from one to the other. After that, the previous case may occur and the same solution procedure is applied.

3.4. Starting up the Adaptive Survey Design

The introduction of adaptive survey design was part of a major redesign of the Labour Force Survey. In order to be able to estimate a shift in results due to this redesign, both the new and regular designs were to run in parallel in the fourth quarter of 2020. As CATI and CAPI are follow-up-modes, a start was made in July 2020 with CAWI. To this end, 24 weekly samples of 3154 people each were selected to complete the LFS questionnaire via CAWI. Figure 1 shows the weekly CAWI response rates for the separate target groups, and for the total sample.

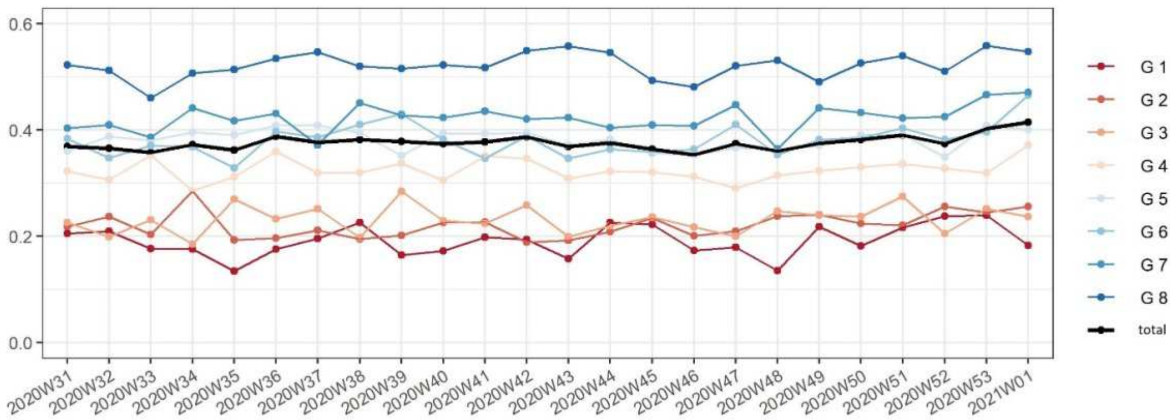

Figure 1. *CAWI response rates by target group and week.*

The CAWI response rates per target group over the entire period of 24 weeks are shown in Table 4.

Table 4. *Expected (e) and realised (r) CAWI response rates per target group.*

group	e	r	r - e	100 × (r – e)/e
1	14.4	19.2	4.8	33.3
2	21.2	22.1	0.9	4.2
3	21.5	23.2	1.7	7.9
4	28.4	32.5	4.1	14.4
5	32.1	38.2	6.1	19.0
6	34.1	37.1	4.0	11.7
7	39.9	42.2	2.3	5.8
8	47.9	52.2	4.3	9.0
tot	34.0	37.5	3.5	10.3

In each target group the realised response rate (r) is greater than the estimated response rate (e). The measures taken against the COVID-19 pandemic could be a reason for this. The difference (r - e) is greatest in group 6 with 6.1 percentage points and smallest in group 2 with 0.9 percentage points. The overall realised CAWI response rate is 3.5 percentage points greater than estimated. The relative difference (r – e)/e is largest in group 1 and smallest in group 2.

Figure 2 contains the weekly CATI- and CAPI- sampling fractions per target group. As established by design, CAPI is mainly applied in target groups with a low number and CATI in target groups with a high number. Due to a lack of telephone numbers in some weeks in group 2, shifts from CATI to CAPI can be seen there. The selection fractions for target groups 4 to 8 show an upward trend in the later portions. This is caused by the increase in CAWI response, which left capacity for random addition after selection with the fractions from the sampling design.

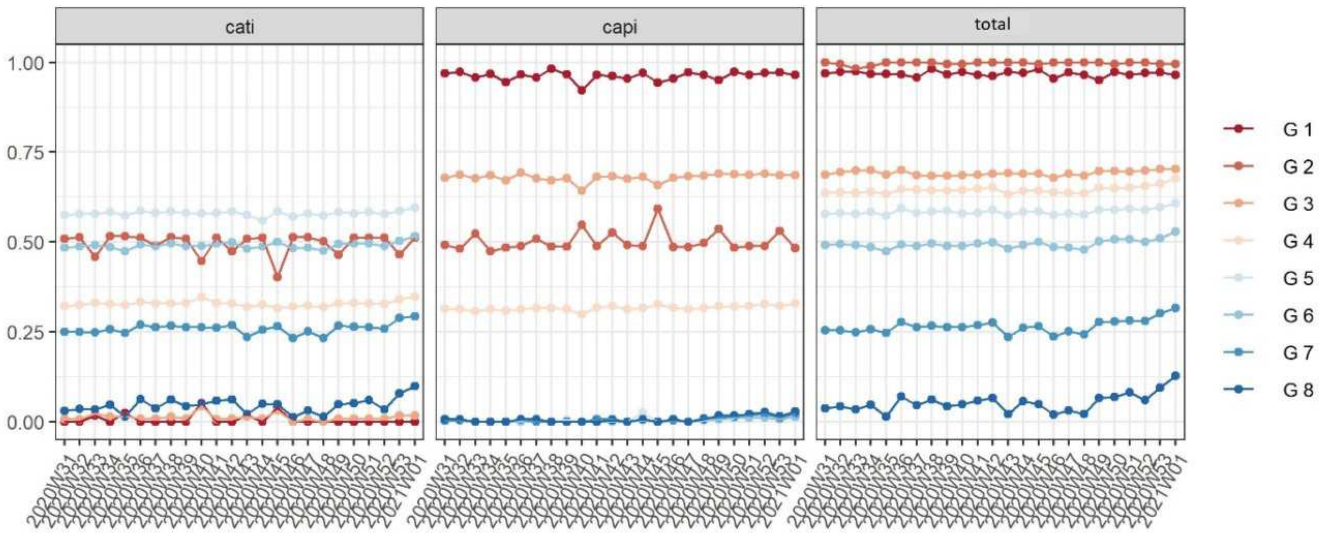

Figure 2. *CATI-, CAPI-, and total sampling fractions by target group and week.*

3.5. Impact of Adaptive Survey Design on Survey Results

Due to the COVID-19 pandemic in 2020, some CAPI samples were only partially or not at all observed. So both the CAPI response rates and the effect of the adaptive design on the survey results of the LFS could hardly be determined. Therefore it was decided that the first wave of the regular design was extended with a period of six months from January to June 2021. This resulted in a nine-month parallel run for the first wave and the transition to the adaptive design was postponed until July 2021. An earlier transition to the new design could entail the risk of inaccurate estimates of labour force figures, as turning points due to the lockdown could coincide with discontinuities caused by the adaptive design.

Unfortunately, due to the lockdown, no CAPI was conducted from January to June 2021. Nevertheless, by applying a time series model, discontinuities in labour force figures could be quantified [4]. The new design showed an increase in the unemployed labour force at the national level from about 300,000 to 400,000 people. The employed labour force increased with about 160,000 people, at a level of 9,000,000 people. Note that these estimated differences are the result of all changes in the survey design: sample of people versus addresses, adapted survey design versus fixed survey design, changes in the observation strategy, and changes in the questionnaire.

4. Conclusion

Methodology was developed to optimise adaptive survey design within the context of an official survey with strong cost-quality differences in the design features. These cost-quality differences refer to different observation modes that have a strong influence on both budget and response. An approach is proposed to minimise the coefficient of variation of response probabilities under specified conditions. The conditions may relate to different survey aspects such as costs, interviewer capacity or response rates. The design features to adapt are the sampling selection fractions by observation mode and target group.

The target groups are composed using historical data on response behaviour. Characteristics used to predict response behaviour are preferably strongly correlated with both response propensity and a key target variable of the survey. Assuming no mode-specific measurement errors, survey bias is minimised.

As an application of the methodology, the adaptive design of the Dutch Labour Force Survey was elaborated.

5. Discussion

The mixed-mode design for the first wave of the Dutch Labour Force Survey starts with Internet observation with telephone and face-to-face follow-ups. The coefficient of variation of response probabilities was taken as the objective function in optimising the design. There are a few limitations in this approach: First, it has not been taken into account that bias can be caused by non-random response. Second, the possibility of mode-specific measurement bias was ignored. Third, the allocation of follow-up is posed as a yes-no decision. An alternative is to vary the effort of interviewers by proposing different numbers of contact attempts for the different target groups. Fourth, explanation of response and strata were based on administrative variables that are used in the sampling design.

In order to separate and quantify selection- and mode-effects, Statistics Netherlands has planned an experiment with re-interviews. This experiment will be conducted at the Health survey in the second half of 2022. A sample of people will be approached via the Internet. The nonrespondents will be approached by CAPI. The CAWI- respondents will be randomly divided into two groups. The first group is re-approached via the Internet a few months later, with almost the same questionnaire. The second group is re-approached at home at the same time for a personal interview. By comparing the answers of CAWI-respondents at two points in time, memory effects and actual changes over time can be included in the analysis. By comparing answers from CAWI-respondents at the second time point with those from the CAPI-re-interviews, mode-specific measurement errors can be identified. By comparing CAPI-respondents coming from the CAWI-nonresponse to the CAPI-re-interviews, selection effects can be quantified. Such an experiment has been carried out before in the Crime Victimization Survey [10, 17]. If mode-specific measurement differences appear, then the question arises: which mode is considered the best? A different choice could be made for different variables. Experts in the field should be consulted.

A complication in conducting the survey may be that the estimated response per target group differs significantly from the actual figures. As the Labour Force Survey is a continuous survey, it is possible to monitor and adjust during the observation period, so that the response distribution among the target groups per quarter is in line with the desired distribution.

A general question is to what extent stratification should be survey-specific and to what extent a subset of general strata will always be imposed. This is especially important for regional variables as they affect interviewer workloads over multiple surveys. For the LFS more variables related to response and target variables could be explored. Good candidates are income and benefits.

An advantage of the chosen stratification variables is that they are all included in the weighting model. Therefore, there is no need to adjust the weights by introducing the adaptive design, and it is likely that the variance of the weights will be reduced by the adaptive design, thus increasing the accuracy of the estimators.

References

[1] Alabama R package. Available at: https://CRAN.Rproject=alabama.

[2] Berkel, C. van, S. van der Doef, and B. Schouten. (2020). Implementing Adaptive Survey Design with an Application to the Dutch Health Survey. *Journal of Official Statistics* 36 (3): 609–629. DOI: http://dx.doi.org/10.2478/JOS-2020-0031.

[3] Bethlehem, J. G. (1988). Reduction of Nonresponse Bias Through Regression Estimation. *Journal of Official Statistics* 4 (3): 251–260. Available at: https://www.scb.se/contentassets/-ca21efb41fee47d293bbee5bf7be7fb3/reduction-of-nonresponse-bias-through-regression-estimation.pdf.

[4] Brakel, J. van den. (2021). Monthly Labour Force Figures during the 2021 Redesign of the Dutch Labour Force Survey. *Statistics Netherlands*, discussion paper, January 2022. Available at: https://www.cbs.nl/-/media/_pdf/2022/03/lfs-redesign-2021.pdf.

[5] Brick, J. M. and R. Tourangeau. (2017). Responsive Survey Designs for reducing Nonresponse Bias. *Journal of Official Statistics* 33 (3): 735–752. DOI: https://doi.org/10.1515/jos-2017-0034.

[6] Burger, J., K. Perryck and B. Schouten. (2017). Robustness of Adaptive Survey Designs to Inaccuracy of Design Parameters. *Journal of Official Statistics* 33 (3): 687–708. DOI: https://doi.org/10.1515/jos-2017-0034.

[7] Chun, A. Y., S. G. Heeringa and B. Schouten. (2018). Responsive and adaptive design for survey optimization. *Journal of Official Statistics* 34 (3): 581–597. DOI: https://doi.org/10.2478/jos-2018-0028.

[8] Groves, R. M., M. P. Couper, S. Presser, E. Singer, R. Tourangeau, G. P. Acosta, and L. Nelson. (2006). Experiments in producing nonresponse bias. *Public Opinion Quarterly* 70 (5): 720–736. Available at: https://www.jstor.org/stable/4124223#metadata_info_tab_contents.

[9] Groves, R. M. and S. G. Heeringa. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. *Journal of the Royal Statistical Society*, Series A169: 439–457. DOI: http://dx.doi.org/10.1111/j.1467-985X.2006.00423.x.

[10] Klausch, T., J. Hox, and B. Schouten. (2017). Evaluating bias of sequential mixed-mode designs against benchmark surveys. *Sociological Methods & Research* 46 (3): 456-489. DOI: https://doi.org/10.1177/0049124115585362.

[11] Paiva, T. and J. P. Reiter. (2017). Stop or Continue Data Collection: A Nonignorable Missing Data Approach for Continuous Variables. *Journal of Official Statistics* 33 (3): 579–599. DOI: https://doi.org/10.1515/jos-2017-0028.

[12] Lewis, T. (2017). Univariate Tests for Phase Capacity: Tools for Identifying When to Modify a Survey's Data Collection Protocol. *Journal of Official Statistics* 33 (3): 601–624. DOI: https://doi.org/10.1515/jos-2017-0029.

[13] Peytchev, A. (2010). Responsive design in telephone survey data collection. *International Workshop on Household Survey Nonresponse*, Nürnberg, Germany, August 30.

[14] Plewis, I., and N. Shlomo. (2017). Using Response Propensity Models to Improve the Quality of Response Data in Longitudinal Studies. *Journal of Official Statistics* 33 (3): 753–779. DOI: https://doi.org/10.1515/jos-2017-0035.

[15] Särndal, C. and P. Lundquist. (2017). Inconsistent Regression and Nonresponse Bias: Exploring Their Relationship as a Function of Response Imbalance. *Journal of Official Statistics* 33 (3): 709–734. DOI: https://doi.org/10.1515/jos-2017-0033.

[16] Schouten, B., A. Peytchev, and J. Wagner. (2017). Adaptive Survey Design. *Series on Statistics Handbooks*. Chapman and Hall/CRC. Available at: https://www.routledge.com/Adaptive-Survey-Design/Schouten-Peytchev-Wagner/p/book/9780367735982.

[17] Schouten, B., J. van den Brakel, B. Buelens, J. van der Laan, and T. Klausch. (2013). Disentangling mode-specific selection and measurement bias in social surveys. *Social Science Research* 42: 1555–1570. Available at: https://www.cbs.nl/-/media/imported/documents/2012/32/2012-11-x10-pub.pdf.

[18] Schouten, B., M. Calinescu, and A. Luiten. (2013). Optimizing quality of response through adaptive survey designs. *Survey Methodology* 39 (1): 29–58. Available at: https://www.cbs.nl/-/media/imported/documents/2011/23/2011-x10-18.pdf.

[19] Schouten, B., N. Shlomo, and C. Skinner. (2011). Indicators for monitoring and improving representativeness of response. *Journal of Official Statistics* 27 (2): 231–253. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/indicators-for-monitoring-and-improving-representativeness-of-response.pdf.

[20] Tourangeau, R., M. Brick, S. Lohr, and J. Li. (2017). Adaptive and responsive survey designs: a review and assessment. *Journal of the Royal Statistical Society*, Series A, 180 (1): 203–223. DOI: https://doi.org/10.1111/rssa.12186.

[21] Wagner, J. (2008). Adaptive Survey Design to Reduce Nonresponse Bias. PhD thesis, University of Michigan, Ann Arbor, USA. Available at: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/60831/jameswag_1.pdf?sequence=1&isAllowed=y.

[22] Wagner, J. (2013). Adaptive contact strategies in telephone and face-to-face surveys. *Survey Research Methods* 7 (1): 45–55. DOI: https://doi.org/10.18148/srm/2013.v7i1.5037.