Redesigning the LFS: solving the puzzle

Martijn Souren (Statistics Netherlands)

1. Introduction

Statistics Netherlands (SN) has been using web interviewing as a central part of the data collection strategy of the Labour Force Survey (LFS) since 2012. Since then, no major changes have been made to either the questionnaire or the data collection strategy. However, times are changing. Online data collection is getting more and more important and respondents expect to be able to fill out a questionnaire on their smartphone or tablet. This requires a questionnaire that is brief, easy and simple to fill out. In the Reykjavik papers and presentations of SN (Janssen, 2018 and Souren, 2018) we explained how the introduction of person sampling should facilitate to achieve this goal. In the other paper of SN for this workshop SN (Cremers, 2019) the principles of the questionnaire design are explained to make the questions as brief and simple as possible. This paper and presentation will explain how the puzzle of a brief and simple (wave one) questionnaire in relation to the demands for subsampling will be solved.

2. Person sampling

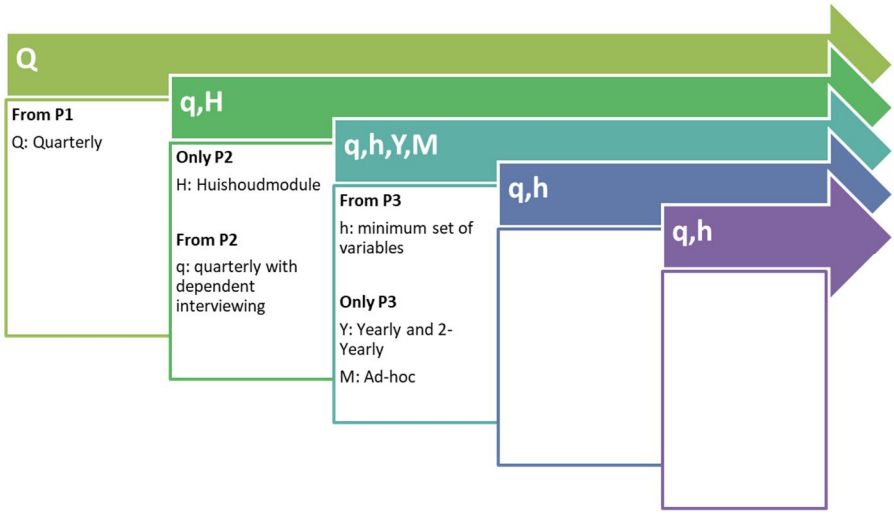
The data collection of Statistics Netherlands is conducted with a CAWI-first philosophy. This means that all the sampled units are in the first instance approached to respond via the web. To make this strategy as efficient as possible SN is always looking to maximize the CAWI response rate. This is already done by for example sending multiple letters and using incentives. To further increase the CAWI-response rate, it was decided that questionnaires should have a mobile lay-out if possible in order to facilitate responses via mobile devices. A sample (and interview) of addresses does not match with the idea of increased use of mobile devices. Interview length and therefore higher drop-out rates together with high rate of proxy answers would become a serious concern. Also, the drop-out rate for CAWI responses from the first to the second wave is considered too high. A shorter questionnaire in the first wave that facilitates mobile device responses will improve the respondent experience and will reduce drop-out rates for the second wave.

For these reasons the redesigned D-LFS is foreseen to become a person sample where in the first wave only the questions for the selected person have to be answered. Moreover, the idea was to minimize the length of the questionnaire by only posing the questions which have a quarterly periodicity. The panel should in turn be used for the output requirements with a lower frequency (f.i. (two-)yearly estimates, ad-hoc estimates and household variables). However, the panel waves should also be as short as possible because there also, CAWI and mobile devices will be used. Therefore, the idea is to apply the wave approach to the fullest extent possible in order to keep the questionnaire of each wave as short as possible.

3. Wave approach

In the Reykjavik paper and presentation (Souren, 2018) the proposed subsampling architecture for the redesigned LFS was presented (figure 1). Wave 1 was planned to be used for the variables with a quarterly periodicity for selected persons only. Herewith making the first wave as short as possible and fit for answering by mobile devices. It was also argued there that the first wave will be used as a benchmark for the key variables where time series modelling will be applied because a first wave for the selected person only does not suffer from high proxy rates and panel attrition. In the second wave the household module will be introduced, determining the composition of the household and asking the minimum set of variables required for all household members. Of course, also the quarterly variables continue in wave 2 but the questionnaire will on average be relatively short due to dependent interviewing for the selected person. From wave 3 onwards the other household members continue with the minimum set of variables using dependent interviewing where possible and there the (two-)yearly and ad-hoc modules will be asked. Wave 4 and 5 continue with the quarterly variables and minimum set of variables for all household members. These waves are available for possible additional questions requested on demand.

Figure 1: Proposed subsampling design of the new Dutch LFS, version 1



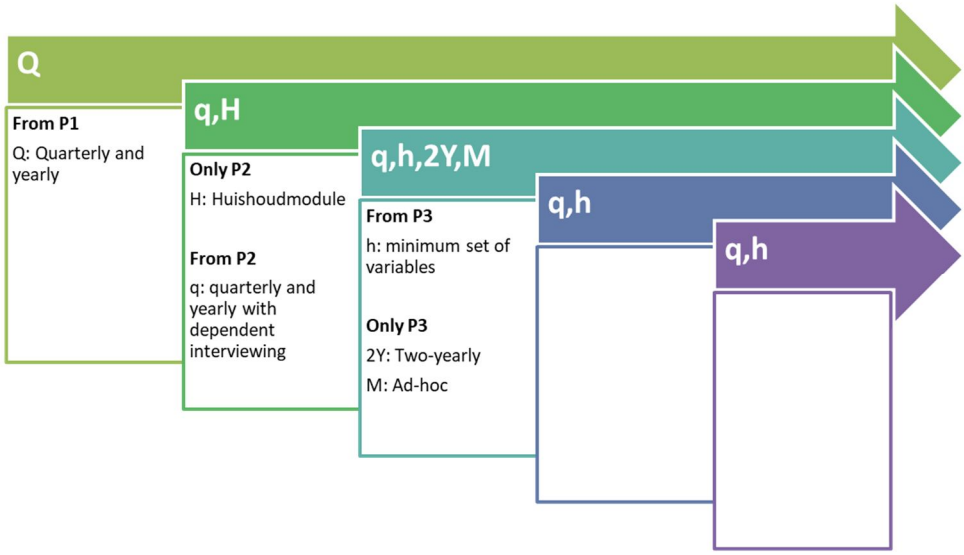
While developing the questionnaire it turned out that a large selection of yearly variables was actually not fit for the third wave. Table 1 presents the list of yearly variables that were to be added to the third wave only. One can see that this is a heterogeneous list and from a questionnaire design perspective it is difficult to make this work properly. If the questions would be placed in the first wave only this is no problem, but if these are posed in a later wave there are some issues. Take for example the education variables. When collecting data in the first wave on the level of education attained it would follow logically to ask for field of education or work experience during education but if this is asked in a different wave this is not possible unless questions are repeated or previous answers are presented to the respondent. The same holds for questions on labour contract and dependent self-employed. For retrospective questions it is rather strange to ask for this after being in the survey for more than half a year. A question on supervisory responsibilities is already needed for the ISCO classification which is quarterly. Furthermore, HOMEWORK and HATYEAR are part of the minimum set of variable and should therefore also be part of wave 2 anyway.

Table 1. List of yearly variables to be collected

Variable	Module
TEMPREAS	Labour contract
TEMPAGCY	Labour contract
MAINCLNT	Dependent self-employed
VARITIME	Dependent self-employed
SUPVISOR	ISCO
SIZEFIRM	NACE
LOOKOJ	Wanting work
HWWISH	Wanting work
NEEDCARE	Wanting work
HATFIELD	Education
HATWORK	Education
HATYEAR	Education
WAYJFOUN	Looking for work
STAPROPR	Retro
NACEPR2D	Retro
ISCOPR3D	Retro
HOMEWK	Workplace
INCGROSS	Register

Since for many yearly variables it is more logical to ask them already in wave 1, SN decided to adjust the subsampling design. Yearly variables will be asked in wave 1. But since the sub-sample for ad-hoc modules and biennial variables shall also be part of the sample for annual variables, the data collection on yearly variables has to continue at least until wave 3. And because it would be a waste not to use the information collected on yearly variables in wave 1 and 2, it is decided to add the variables throughout the panel and thus make them quarterly as well (see figure 2) .

Figure 2: Proposed subsampling design of the new Dutch LFS, version 2



4. Conclusion

With the current heterogeneous list of yearly variables a subsampling approach in a consecutive wave is not a preferred option. The questions for the two-yearly and ad-hoc modules are less diverse but from a questionnaire design perspective one could still argue that also in that case it would be better to use the first wave as a subsample for these modules. For example, the questions on participation in formal and non-formal education and training in the last *twelve* months do not fit well in the third wave either so these may have to be moved as well.

Thus, a wave approach with a heterogeneous list of variables can only work properly when subsampling is applied to the first wave. Moving all of the yearly, two-yearly and ad hoc modules to the first wave would however increase the burden even more and probably too much. We have to keep in mind that the first wave will increasingly be answered by mobile devices where drop-out rates are higher for long questionnaires. The result for the Dutch LFS is therefore a compromise design with the first wave and consecutive waves being longer than strictly necessary to make subsampling still possible and still keep the first wave as short as possible.

References

Cremers, M. (2019), 'Redesigning the LFS - omnimode and b1-language', paper for LFS workshop Budapest 2019, Statistics Netherlands, Heerlen, April 2019.

Janssen, B. (2018), ' Redesigning the Labour Force Survey', paper for LFS workshop Reykjavik 2018, Statistics Netherlands, Heerlen, April 2018.

Souren, M. (2018), ' Redesigning the Labour Force Survey', paper for LFS workshop Reykjavik 2018, Statistics Netherlands, Heerlen, April 2018.