DE GRUYTER OPEN

Journal of Official Statistics, Vol. 32, No. 2, 2016, pp. 259–286, http://dx.doi.org/10.1515/JOS-2016-0013

On a Modular Approach to the Design of Integrated Social Surveys

Evangelos Ioannidis[1], *Takis Merkouris*[2], *Li-Chun Zhang*[3], *Martin Karlberg*[4], *Michalis Petrakos*[5], *Fernando Reis*[6], *and Photis Stavropoulos*[7]

This article considers a modular approach to the design of integrated social surveys. The approach consists of grouping variables into 'modules', each of which is then allocated to one or more 'instruments'. Each instrument is then administered to a random sample of population units, and each sample unit responds to all modules of the instrument. This approach offers a way of designing a system of integrated social surveys that balances the need to limit the cost and the need to obtain sufficient information. The allocation of the modules to instruments draws on the methodology of split questionnaire designs. The composition of the instruments, that is, how the modules are allocated to instruments, and the corresponding sample sizes are obtained as a solution to an optimisation problem. This optimisation involves minimisation of respondent burden and data collection cost, while respecting certain design constraints usually encountered in practice. These constraints may include, for example, the level of precision required and dependencies between the variables. We propose using a random search algorithm to find approximate optimal solutions to this problem. The algorithm is proved to fulfil conditions that ensure convergence to the global optimum and can also produce an efficient design for a split questionnaire.

Key words: Efficient design; respondent burden; sample allocation; simulated annealing; split questionnaire.

1. Introduction

In recent years, several national statistical agencies have explored the possibility of integrating social surveys, as a means of meeting growing demand for new and improved

[1] Athens University of Economics and Business, Patission str. 76, 10434 Athens, Greece. Email: eioannid@aueb.gr
[2] Athens University of Economics and Business, Patission str. 76, 10434 Athens, Greece. Email: merkouris@aueb.gr
[3] University of Southampton, Southampton, SO17 1BJ, United Kingdom and Statistics Norway. Email: l.zhang@soton.ac.uk
[4] European Commission (Eurostat), L-2920 Luxembourg. Email: Martin.KARLBERG@ec.europa.eu
[5] Agilis SA, Acadimias 96-100, 10677 Athens, Greece, Email: Michalis.Petrakos@agilis-sa.gr
[6] European Commission (Eurostat), L-2920 Luxembourg. Email: Fernando.REIS@ec.europa.eu
[7] Agilis SA, Acadimias 96-100, 10677 Athens, Greece. Email: Photis.Stavropoulos@agilis-sa.gr
Acknowledgments: The work presented in this article was part of a methodological consultancy project (contract number 61001.2011.005-2012.426) undertaken by Agilis SA (Athens, Greece) on behalf of the European Commission (Eurostat). The participation of the members of the Department of Statistics of the Athens University of Economics and Business in this project was coordinated by the University's Research Centre. We would like to thank the Associate Editor and the three referees for their valuable comments and suggestions, which have led to a considerable improvement of the manuscript.
Disclaimer: The views expressed here are those of the authors and do not necessarily reflect the official views of the European Commission (Eurostat).

statistics, while at the same time streamlining the survey operations in order to curb costs and limit respondent burden. See, for example, the experience of the UK Office for National Statistics (Smith 2009), the Australian Bureau of Statistics (2012), and the Dutch Central Bureau of Statistics (Cuppen et al. 2013).

In this article, we consider a general approach to survey integration that includes all the social surveys managed by a national statistical agency. This approach is designed to provide a flexible framework that will be able to accommodate the needs of social statistics as they change over time. The starting point for this research was the Eurostat project *Streamlining and Integration of the European Social Surveys* (see Reis 2013). This project was set up to support the implementation of the ESS Vision (European Commission 2009), the strategy for modernising the European Statistical System (ESS), in particular the European system of social statistics, in accordance with the commitments made in the Wiesbaden memorandum (European Statistical System Committee 2011).

The approach presented in this article is based on a modular design. The basic features are as follows. The existing ‘items’ (variables) in social surveys are restructured into a number of mutually exclusive groups, called modules, each of which consists of a small number of items. The modules are themselves distributed among an appropriate number of ‘instruments’, in such a way that each instrument consists of a fixed set of modules and each module is present in one or more instruments. These instruments will together replace the current set of survey questionnaires. Each instrument will be administered to a probability sample of population units within a specified time period. All the units in a sample will be asked to respond to all the items in the instrument assigned to them, except where routing determines otherwise. The modular design is therefore mainly characterised by: (i) the composition of the instruments; and (ii) their associated sample sizes.

The modularisation approach offers a number of potential advantages over the traditional ‘stovepipe’ survey programmes. Both the cost of carrying out the survey and the burden placed on respondents can be reduced, subject to precision requirements, for example by putting modules with lower precision requirements in instruments that are administered to fewer sample units, or by combining several instruments that all contain the same module. The modularisation can enhance the analytical potential of survey data by means of a suitable composition of the instruments, instead of by simply adding more modules to existing surveys. Likewise, new modules can be introduced with a short lead time and in a more cost-effective way, thus allowing emerging needs to be better met.

Designing and implementing a modular approach is by no means straightforward, however. There are specific methodological and technical problems that need to be addressed, and there are also a number of broader issues, some of which we will discuss here. In this article we focus on two questions: (1) how to develop a modular design that is sufficiently versatile to support an integrated social survey system; and (2) how to find a solution to the dual design problem of instrument composition and sample size allocation. We also refer to Karlberg et al. (2015) for a nontechnical description of this approach to modular design, and to Reis (2013) for a discussion of related survey-management and subject-matter issues.

The existing survey methodology of split questionnaire design (SQD) is most akin to the modular design. For a single survey the SQD allows different sets of items to be collected from different sample units; the full questionnaire consists of the union of all these *split* *questionnaires*. The SQD has attracted increasing attention in recent years. See, for example, Raghunathan and Grizzle (1995) for a Bayesian approach, or Chipperfield and Steel (2009) for a design-based approach. While the SQD is traditionally used to lessen the high respondent burden caused by a long questionnaire, it also provides a possible design for survey integration. For example, two existing separate and independent sample surveys can be considered jointly as a single survey with a two subsample SQD. Issues relating to efficient design and estimation for a single SQD have been studied recently in Chipperfield and Steel (2009; 2011). Efficient estimation in a broadened context of SQD has been studied more extensively in Merkouris (2010; 2015), building on earlier methods for composite estimation that combined data from different sample surveys, in order to increase precision and align estimates relating to the same item (see Renssen and Nieuwenbroek 1997 and Merkouris 2004; 2013).

There are, nevertheless, a number of difficulties that make it impossible to fully endorse an SQD model for the integration of social surveys. For example, it is not clear how to accommodate the difference in the timing and frequency of different surveys, or how to implement rotating panels overlapping over time. As we explain in Section 2, these situations can be accommodated under the approach proposed in this article.

To formulate the modular design problem, denote the set of all, say m, modules by $\{M_1, \ldots, M_m\}$, and let the variables in module M_i be $\{X_{i,j}, j = 1, \ldots, \nu_i\}$. Denote the set of all, say k, instruments by $\{I_1, \ldots, I_k\}$ and the associated samples of sizes $n_1, \ldots, n_k$ by $\{S_1, \ldots, S_k\}$. Further, denote the number of modules in instrument I_j by m_j and the set of instruments containing module M_a by $\Theta_a \subseteq \{I_1, \ldots, I_k\}$. Then, for a given set of modules and a predetermined number of instruments, the design problem is to determine the composition of instruments $I_1, \ldots, I_k$ and the corresponding sample sizes $n_1, \ldots, n_k$, in such a way as to minimise a given cost function while respecting multiple constraints. The cost function is a generic loss function determined by the sample sizes and a number of other features of the integrated survey. The constraints are primarily lower bounds on the precision of the estimation for individual modules, and for groups of modules that have to be placed in the same instrument. Such groups will be called 'mandatory crossings'. Other constraints that need to be taken into account relate to features of particular modules and to the choice of which instrument to put them in. For example, modules that have to be administered with a certain periodicity (e.g., quarterly or annually) should be put in instruments of the same periodicity. These constraints mean that the solution to the optimisation problem (i.e., constrained minimisation of the cost function) can only be sought among certain instrument compositions, which we will call 'admissible'. It should be noted that we do not consider the alternative setup of the optimisation problem — in which the information collected is maximised, subject to constraints on cost, precision and admissibility — because it is unclear how the objective function would need to be defined in this case.

Table 1 provides a generic illustration of the structure of instruments and modules. Suppose, initially, that there are two usual sample surveys. Survey A is carried out quarterly, with sample size of 1,000 per quarter, and contains modules M_1, M_2, and M_3. Survey B is conducted annually, with sample size of 10,000, and contains only module M_4. The precision requirements are given in terms of the required annual sample sizes for the respective modules. Assuming for simplicity that all modules contribute equally to the cost function, these two surveys together cost $4*3*1{,}000 + 10{,}000 = 22{,}000$ person

Table 1. Illustration of the composition of instruments in terms of modules

Possible optimal reorganisation of a quarterly survey of size 1,000 per quarter with modules M_1, M_2, M_3 and a further annual survey of size 10,000 with a single module M_4

					Instruments (after optimisation)					
					Name	$I_{1.Q1}$	$I_{1.Q2}$	$I_{1.Q3}$	$I_{1.Q4}$	I_2
					Sample	S_1	S_2	S_3	S_4	S_5
					Sample size	1,000	1,000	1,000	1,000	8,000
						Instrument composition (modules) (after optimisation)				
Modules	Module variables	Required periodicity	Required Annual sample size	Annual sample size before optimisation	Annual sample size after optimisation					
M_1	$\{X_{1,1},\ldots,X_{1,v_1}\}$	Quarterly	4,000	4,000	4,000	**1**	**1**	**1**	**1**	0
M_2	$\{X_{2,1},\ldots,X_{2,v_2}\}$	Annual	2,000	4,000	2,000	**1**	0	**1**	0	0
M_3	$\{X_{3,1},\ldots,X_{3,v_3}\}$	Annual	1,000	4,000	1,000	0	**1**	0	0	0
M_4	$\{X_{4,1},\ldots,X_{4,v_4}\}$	Annual	10,000	10,000	10,000	**1**	**1**	0	0	**1**
						Instrument composition (modules) (after optimisation)				
Crossings	Crossing members	Required periodicity		Annual sample size before optimisation	Annual sample size after optimisation					
G_1	$\{M_2, M_4\}$	Annual	1,000	0	1,000	**1**	0	0	0	0
G_2	$\{M_3, M_4\}$	Annual	1,000	0	1,000	0	**1**	0	0	0

modules. Reorganising the modules into the four quarterly instruments $I_{1.Q1}$, $I_{1.Q2}$, $I_{1.Q3}$, $I_{1.Q4}$ and the annual instrument I_2, as illustrated in Table 1, reduces the cost to $(4+2+1+2) * 1{,}000 + 8{,}000 = 17{,}000$ person modules. Moreover, the reorganisation can be used to accommodate additional information needs, illustrated here by the crossings $G_1 = \{M_2, M_4\}$ and $G_2 = \{M_3, M_4\}$, with the required respective sample sizes. Normally, the required crossings could be obtained by including, for example M_2 and M_3 in Survey B, or M_4 in Survey A, thus further increasing the total cost of the survey.

In this article, we propose to perform the optimisation described for all admissible instrument compositions using a random search algorithm known as simulated annealing (see e.g., Kirkpatrick et al. 1983). We apply the simulated annealing algorithm in a manner adapted to this problem, and show that certain conditions ensuring convergence with probability 1 to the global optimum are satisfied. This means, in practice, that a sufficiently large number of iterations of this algorithm should yield an acceptable approximate solution to the optimisation problem. At each iteration of the search algorithm, we use the simplex algorithm to determine the optimal sample sizes for each instrument, subject to the given precision requirements. This allows us to evaluate the cost of each admissible instrument composition.

The rest of the article is organised as follows. In Section 2, we consider the various constraints that may be encountered in practice and describe how these are built into the optimisation procedure. To simplify the exposition, we start by assuming simple random sampling, and discuss the case of complex sampling designs later. In Section 3, we describe the proposed random search algorithm, which can be used to find approximate solutions to the optimisation problem in practice. The framework of Sections 2 and 3 is extended in Section 4 to allow for complex sampling designs. In Section 5, we give an illustration of the modular design approach using the EU Labour Force Survey as an example. Section 6 contains some concluding remarks.

2. Constraints on Modules and Instruments

In this section, we discuss a number of constraints on modules and instruments. We explain how they can be incorporated into the optimisation framework, so that the algorithm only visits admissible solutions. The main set of constraints arises from the requirement to achieve a given level of precision in estimating parameters (e.g., totals, means and proportions) for the variables of each module. Other constraints relate to the joint observation of modules and the periodicity of modules and instruments.

2.1. Precision Requirements for Modules

Precision requirements for a variable are often formulated in terms of the variance, denoted by $V(\hat{\theta})$, of an estimator $\hat{\theta}$ of some finite-population parameter θ (total, mean or proportion). The first step is to determine the sample size that is required to achieve a specified level of precision. We start by assuming that the samples for all instruments are selected using simple random sampling and independently from one another (in Section 4 we discuss departures from this assumption). Let us first consider the Horvitz-Thompson (HT) estimator $\hat{\theta}$ of θ based on the sample of a single instrument. Then the variance of $\hat{\theta}$ is $V(\hat{\theta}) = \sigma_\theta^2/n$, assuming a negligible sampling fraction n/N. Here N is the population size and $\sigma_\theta^2 = S^2$, the finite-population variance, if θ is a mean, and $\sigma_\theta^2 = N\theta\,(1-\theta)/(N-1)$ if θ is a proportion. Assuming that $\hat{\theta}$ is normally distributed, the precision requirement defined as the attainment of a certain margin of error $e = \left|\hat{\theta} - \theta\right|$ with probability $1-\alpha$ is satisfied if $V(\hat{\theta}) \leq \left(e/z_{1-\alpha}\right)^2$, where $z_{1-\alpha}$ is the $1-\alpha/2$ standard normal quantile. Thus, in the case of a single instrument, the precision requirement is satisfied by

$$n \geq n^*, \text{ where } n^* = z_{1-\alpha}^2 \sigma_\theta^2 / e^2. \tag{1}$$

We could alternatively start with the coefficient of variation $CV(\hat{\theta}) = \sqrt{V(\hat{\theta})}/\theta$ and formulate the precision requirement as: $CV(\hat{\theta}) \leq c$ for some constant c. This then translates into $n \geq n^*$, where $n^* = \sigma_\theta^2/c^2\theta^2$.

If a variable belongs to a module that is present in more than one instrument, then θ may be estimated, using the combined data from the associated samples, as the weighted average $\hat{\theta} = \sum w_i \hat{\theta}_i$, where the summation extends over all instruments that contain the variable of interest, $w_i = n_i/\sum n_j$, and n_i is the size of the ith sample. It is easy to show that if these simple random samples are independent, this composite estimator will have a minimum variance given by $V(\hat{\theta}) = \sigma_\theta^2/\sum n_i$, for negligible sampling fractions n_i/N.

Where a variable is present in a number of instruments, the precision requirement can therefore be satisfied by

$$\sum n_j \geq n^*. \tag{2}$$

When considering all the variables in a module M_a, which is present in instruments $I_j \in \Theta_a$, this process could either be repeated for all variables in the module, with the maximum of the required sample sizes being chosen, or one of the variables could be considered as the 'main variable', and the sample size requirement of the module set in accordance with the precision requirement for this variable. In both cases, the module's precision requirement would be satisfied by

$$\sum_{\Theta_a} n_j \geq n_a^*, \tag{3}$$

for some appropriate n_a^*, which is obtained as above. Here, the summation extends over Θ_a, the set of all instruments containing module M_a. Thus, if, for example, $\Theta_a = \{I_3\}$, then we require $n_3 \geq n_a^*$, while if $\Theta_a = \{I_1, I_3, I_7\}$, then we require $n_1 + n_3 + n_7 \geq n_a^*$.

These m linear constraints, one for each module, on the sample sizes $n_1, \ldots, n_k$ of the various samples, can be represented in the $m \times k$ 'composition matrix' $\mathbf{A}$, with elements $a_{i,j}$, where $a_{i,j} = 1$ if the module in row i is present in the instrument in column j and $a_{i,j} = 0$ otherwise. The precision requirements for all modules may thus be expressed as the following set of constraints on the vector $\mathbf{n} = (n_1, \ldots, n_k)'$ of the sample sizes of all instruments:

$$\mathbf{An} \geq \mathbf{n}^*, \ \mathbf{A} \text{ is an } m \times k \text{ matrix}, \tag{4}$$

where the inequality is to be understood componentwise, and $\mathbf{n}^* = \left(n_1^*, \ldots, n_m^*\right)'$ is the vector of the minimal sample size required for each module.

2.2. *Mandatory Crossings*

In the previous section we discussed sample size requirements determined by the need for a certain level of precision in estimating the parameters of the marginal distribution of the variables in a module. In this section, we consider the sample size requirements relating to mandatory crossings, that is, the sample sizes needed to ensure that the requirement for a group of modules to be simultaneously present in at least one instrument is met. Mandatory crossings make it possible to carry out multivariate analyses as they ensure that all the relevant information is collected from the same sample units. Multivariate analysis can include, for example, estimating regression coefficients or estimating the average of a variable conditionally on the value of another variable (e.g., unemployment given the level of educational attainment). Clearly, for a mandatory crossing a suitable sample size requirement must be specified, so as to ensure estimation of desired precision for the parameters of the multivariate analysis.

Crossings are added as separate rows, $i = m + 1, \ldots, m'$ in the $\mathbf{A}$ matrix. If the crossing is, say, in the i_0th row, then the corresponding sample size requirement is given by $n_{i_0}^*$. The idea is that row i_0 will act as a 'constraint', forcing all modules that are members of the crossing in row i_0 to be jointly present in some instruments, which have a total sample size of at least $n^*_{i_0}$. If the crossing is included in the instrument corresponding to column j_0, we set $a_{i_0,j_0} = 1$, and $a_{i,j_0} = 1$ for all rows i corresponding to modules that are members of the crossing. This means that for a given column of the matrix **A** some rows will have to be set jointly due to mandatory crossings. Note, however, that each member M_i may be included in further instruments beyond the ones in which the crossing is included, for example, if M_i has a sample size requirement n^*_i greater than $n^*_{i_0}$.

Thus, the extended sample size constraints in modules and crossings can be expressed as

$$\mathbf{An} \geq \mathbf{n}^*, \ \mathbf{A} \text{ is an } m' \times k \text{ matrix.} \tag{5}$$

2.3. *Periodicity and Concordance of Modules and Instruments*

In addition to guaranteeing a certain sample size for individual modules and crossings, the choice of instruments must also take into account the periodicity of the modules: for example, certain modules may be administered on a quarterly basis, others on a yearly basis. In order for each module to be administered to a population sample at the required periodicity, the instruments will also be assigned a periodicity, at which they will be administered to sample units. The resulting instruments, one for each wave of data collection, belong together to a common 'parent' set of instruments. Thus, a quarterly parent instrument is sprouted into four quarterly 'sibling' instruments and a monthly parent instrument is sprouted into twelve monthly sibling instruments. We will label all sibling instruments using two indices, the first of which will indicate the common parent, and the second the siblings. Thus, for example, a parent set of instruments I_i with monthly period will consist of sibling instruments $\{I_{i.1}, \ldots, I_{i.12}\}$, one of which is to be administered to a population sample each month. It should be noted that although our approach does not require this, sibling instruments belonging to the same parent will usually have a similar core of modules, as each parent instrument would typically have been created to accommodate the repeated administration of a set of modules. Furthermore, although I_i is actually a set of instruments, we will use the same notation as for a single instrument when the context is clear. We will need twelve columns in the composition matrix, one for each sibling instrument, to represent the parent set of instruments $I_i := \{I_{i.1}, \ldots, I_{i.12}\}$ accurately.

In order to ensure that a module is administered with the appropriate periodicity and with the correct sample sizes, we will assume that all instruments belonging to the parent I_i are administered equally spaced in time and to samples of identical sizes. Moreover, a module may only belong to members of I_i if the period of the module (e.g., twelve for an annual module, three for a quarterly module) divided by the period of the parent instrument I_i is an integer, say r. The module must thus participate in a 'complete periodic subset' of the parent instrument, which is defined as a subset of the form $\{I_{i.s}, I_{i.(s+r)}, I_{i.(s+2r)}, \ldots\}$ of I_i or as a union of such subsets. For example, if a quarterly module is included in a monthly parent instrument I_i, then it must be included either in $\{I_{i.1}, I_{i.4}, I_{i.7}, I_{i.10}\}$, or in $\{I_{i.2}, I_{i.5}, I_{i.8}, I_{i.11}\}$, or in $\{I_{i.3}, I_{i.6}, I_{i.9}, I_{i.12}\}$, or in a union of such subsets.

This means that for a given row (i.e., a module) of the matrix **A**, its elements for the columns corresponding to a complete subset of instruments for the given module will have to be set jointly. See the 'proposal mechanism' discussed in Subsection 3.2 for an explanation of how the proposed random search algorithm deals with such constraints.

A module (or mandatory crossing) cannot be included in an instrument whose period is greater than that of the module. Thus, a quarterly module may be present in quarterly or monthly instruments, but not in annual instruments. If it is nevertheless desired to include a module in an instrument of a lower frequency, for example in order to cross it with a module appearing only in that instrument, then this module should be duplicated and should appear in the list of modules once with the greater period and once with the lower period. The same restriction holds for crossings, their period being defined as the lowest of those of the member modules. Indeed, the period of each member module must be an integer multiple of the period of the whole crossing.

The constraint for sibling instruments to be allocated equal sample sizes is represented by a balancing matrix $\mathbf{B}$, with k columns. Each set of sibling instruments with p siblings corresponds to a set of $p - 1$ rows of the matrix $\mathbf{B}$. Each such row, say i_0, will ensure that $n_j = n_{j'}$ for one of $j' = j + 1, \ldots, j + p$. We therefore set $b_{i_0,j} = 1$ and $b_{i_0,j'} = -1$, and $b_{i_0,j''} = 0$ for $j'' \notin \{j, j'\}$. We then require $\mathbf{Bn} = \mathbf{0}$. It should be noted that the balancing matrix $\mathbf{B}$ is a constant in the optimisation process, unlike the composition matrix $\mathbf{A}$.

2.4. *Other Constraints on Modules and Instruments*

In this section we describe some additional constraints that may arise in practice and indicate how they could be expressed in terms of admissible instrument compositions.

2.4.1. Constraints on the Presence of Modules in Specific Instruments

There are certain modules that should be included either a) in all instruments or b) in specific instruments only:

a. Modules containing variables used as auxiliaries in weight calibration. These are usually categorical variables for which the population totals by category are known, such as demographic variables. As calibration increases the precision of estimates of correlated variables, and as we are interested in bringing together information from different instruments, it is advisable to require a core set of such modules to be included in all instruments. Similarly, it would be possible to include certain modules related to rare population groups or low prevalence characteristics of interest in all instruments, so as to gather information on these items from all units of the integrated sample.
b. Modules for which longitudinal information is required. These modules should only be included in appropriate instruments (e.g., where there is a specified sample overlap between successive waves).

It should therefore be possible, within the process of optimisation, to specify whether a certain module should be present in all instruments, or only in a specified group of instruments. This is achieved by specifying which instrument compositions are 'admissible' (those satisfying all necessary specifications) and by allowing the algorithm to visit only these compositions.

2.4.2. Constraints on the Joint Presence of Modules in Some Instruments

a. Dependencies between modules. It may only make sense for respondents to answer questions from one module if they have responded in a particular way to another module. For example, it only makes sense for an individual to answer a question on hours worked if he/she has responded positively to the question as to whether he/she works at all.
b. Many surveys currently being used may include groups of modules that relate to the same thematic block and that are thus logically related to one another. Breaking up such thematic blocks may be confusing for respondents, may make it difficult for the questionnaire to focus on certain issues and may make training of interviewers more difficult. It is therefore advisable to allow the grouping of modules into thematic blocks, and to either include or exclude whole thematic blocks from any instrument.

In view of the above, the optimisation algorithm should exclude from the admissible compositions any composition where an instrument includes a module, say M, but not all the modules on which M depends. The requirement that certain modules only appear as part of a specific thematic block could be imposed by building a mandatory crossing, the sample size requirement of which would be equal to the maximum of the sample size requirements of the modules contained in the block. Alternatively, if a group of modules needs to appear together in all instruments, then such a group could be treated as a single 'supermodule' rather than as a mandatory crossing. This would reduce the computational complexity of the optimisation. Finally, if a certain group of modules should not all appear in the same instrument under any circumstances, then any composition where an instrument contains all the modules from this group will be made inadmissible.

2.4.3. Putting Limits on the Questionnaire Size of An Instrument

There are two ways of ensuring that the questionnaire is of a reasonable size:

The first is the 'hard' or rule-based approach: instrument compositions are only considered admissible if the sum of the burden of the modules belonging to the instrument is below a certain upper bound.

The second is the 'soft', more flexible approach: the questionnaire size is incorporated into the cost function. The cost function — which we discuss in Subsection 3.1 below —is mainly determined by the sample size and the unit cost of a questionnaire, with the unit cost depending on the size of the questionnaire. In Subsection 3.1, we model unit cost as increasing linearly with the size of the questionnaire. Instead, it could be set to increase linearly up to a certain threshold, and then to increase at an accelerated rate (e.g., quadratically) above the threshold. This creates a pressure for the questionnaire size to stay below this threshold, while leaving some flexibility such that exceptions could be allowed, the 'price' of going above this threshold being a higher unit cost.

3. Optimisation over Instrument Compositions and Sample Sizes

In this section, we describe an approach to optimisation over instrument composition and sample size allocation that respects the constraints described above and yields an approximately optimal modular design.

3.1. The Cost Function

The cost function should ideally capture all the relevant contributing factors, including both the cost of producing the survey and the respondents' burden. In reality, this has proven to be impossible in any strict sense, not least because a national statistical agency will necessarily manage multiple surveys in parallel, and the methods and technical systems for conducting surveys are constantly evolving. For the sake of simplicity, we will follow the standard approach and use a linear cost function.

The overall cost, C, is assumed to be the sum of the cost of all instruments (samples) considered:

$$C = \sum_{j=1}^{k} C_j.$$

The cost C_j of sample j is assumed to be a linear function of the number of individuals in the sample. This follows Groves (1989, 51) and Cochran (1977, 280). It is composed of a survey-specific fixed cost $C_j^{(f)}$ — the cost incurred regardless of the sample size chosen — and a variable cost, which is assumed (as an approximation) to increase linearly with the sample size,

$$C_j = C_j^{(f)} + \alpha_j n_j.$$

The variable cost coefficient α_j is the unit cost. Its dependence on j allows different unit costs to be specified for different instruments, reflecting, for example, different modes of data collection.

The linearity of C in the sample sizes n_j, as expressed by the two preceding equations, is an important assumption, and one which is necessary for the optimisation algorithm (Subsection 3.2.2, Step ii). A realistic calculation of the cost of each instrument is, of course, important to the design; see, for example UNSTATS (2005, 249–300) for an extensive discussion on the assessment of survey costs. In what follows, we further elaborate on a possible decomposition of the unit cost α_j. The proposed optimisation algorithm is, however, applicable regardless of the details of this decomposition.

Let α_j have a household and individual component,

$$\alpha_j = \overline{q}^{-1} \alpha_j^{(hh)} + \alpha_j^{(ind)},$$

where $\alpha_j^{(hh)}$ is the cost of including a household in the sample j, $\alpha_j^{(ind)}$ the cost of including an individual in the sample and $\overline{q}$ the average number of individuals in the household. The coefficients $\alpha_j^{(hh)}$ and $\alpha_j^{(ind)}$ may consist of a fixed cost β_0 (per person or household) and a variable cost, which increases with the respondent burden of the respective questionnaire.

The burden of the questionnaire can be measured in such a way as to appropriately reflect its cognitive and operational burden, taking into account, for example, whether the respondent needs to look at specific personal records or use complex recall processes in order to be able to answer the questions. The burden could, for example, be measured as the average amount of the interviewer's time required, as in Chipperfield et al. (2013). Alternatively, it could be approximated as the number of questions in the modules included in the instrument, if the burden created by each question is judged sufficiently similar. The unit cost α_j of a questionnaire may, in general, depend on the burden of the modules contained in the instrument. The form of this dependence may be arbitrary, suggested by the specific application. A simple linear specification can be expressed as

$$\alpha_j^{(hh)} = \beta_0^{(hh)} + \beta_1^{(hh)} \sum_{\Lambda_j^{(hh)}} l_\gamma \ \text{ and } \ \alpha_j^{(ind)} = \beta_0^{(ind)} + \beta_1^{(ind)} \sum_{\Lambda_j^{(ind)}} l_\gamma,$$

where l_γ is the respondent burden of module M_γ. The first summation extends over $\Lambda_j^{(hh)}$, the set of all modules concerning households contained in I_j, and the second extends over $\Lambda_j^{(ind)}$, the set of modules concerning individuals contained in I_j. Finally, $\beta_0^{(hh)}$ is the fixed cost per household visited (independently of the number of individuals in the household), $\beta_1^{(hh)}$ is the cost per unit of response burden for the household, $\beta_0^{(ind)}$ is the additional fixed cost per person in the sample (if any), and $\beta_1^{(ind)}$ is the cost per unit of response burden per person in the sample. Thus, we obtain that the unit cost α_j of instrument j has the form

$$\alpha_j = \bar{q}^{-1} \beta_0^{(hh)} + \bar{q}^{-1} \beta_1^{(hh)} \sum_{\Lambda_j^{(hh)}} l_\gamma + \beta_0^{(ind)} + \beta_1^{(ind)} \sum_{\Lambda_j^{(ind)}} l_\gamma.$$

A similar idea for defining the unit cost is used in Chipperfield and Steel (2009; 2011), where α_j is defined as the fixed cost per unit plus the sum of the marginal data collection cost across variables in pattern j.

3.2. *Dual Optimisation: Simulated Annealing and Simplex*

3.2.1. The Optimisation Problem

Assuming a predetermined and fixed number k of instruments, our objective function C will depend on the composition of these instruments and on the sample sizes $\mathbf{n}$. The composition of the instruments is determined by the $m' \times k$ matrix $\mathbf{A}$, where $[\mathbf{A}]_{i,j} = a_{i,j}$: the module i belongs to instrument j if and only if $a_{i,j} = 1$. Thus, we have $C = C(\mathbf{A}, \mathbf{n})$. Our aim is to find

$$\begin{aligned} &\min{}_{\mathbf{A},\mathbf{n}} C(\mathbf{A}, \mathbf{n}), \text{ under the conditions : } \mathbf{A} \text{ admissible}, \ (\mathbf{An})_i \geq n_i^*, \\ &i = 1, \ldots, m' \text{ and } \mathbf{Bn} = \mathbf{0} \end{aligned} \tag{6}$$

that is, over the dual-space of $\mathbf{A}$ and $\mathbf{n}$. Now, for any given composition of instruments, that is, for any given $\mathbf{A}$, it is easy to optimise an objective function that is linear in $\mathbf{n}$ under constraints that are also linear in $\mathbf{n}$ via the simplex algorithm. This yields an optimal sample size allocation $\mathbf{n}^{opt} = \mathbf{n}^{opt}(\mathbf{A})$. In this way, our problem can be translated into a minimisation problem in $\mathbf{A}$ alone, that is,

$$\begin{aligned} &\min{}_{\mathbf{A}} C(\mathbf{A}, \mathbf{n}^{opt}(\mathbf{A})) \text{ under the conditions : } \mathbf{A} \text{ admissible}, \ (\mathbf{An}^{opt}(\mathbf{A}))_i \geq n_i^*, \\ &i = 1, \ldots, m' \text{ and } \mathbf{Bn} = \mathbf{0} \end{aligned} \tag{7}$$

The space to which $\mathbf{A}$ may belong to is a subspace of $\{0,1\}^{km'}$, the space of all sequences of zeros ('0') and ones ('1') of length km', the number of instruments multiplied by the number of modules and crossings. The size of this space depends exponentially on km'. For this reason, the space cannot be searched exhaustively to find the minimum. Heuristic and metaheuristic methods have been proposed as a way of solving combinatorial optimisation problems, such as the one we are facing here. They yield approximate solutions while being computationally feasible (see e.g., Blum and Roli 2003). Algorithms of this type 'search' the state space in such a way that there is a good chance of finding an approximation to the minimum (even if the state space is not searched exhaustively). One of the most prominent such algorithms is simulated annealing. It was inspired by a physical process for growing crystals: a molten fluid is slowly cooled until crystals are formed, the slow cooling rate being crucial for crystal formation. The theoretical properties of simulated annealing have been studied extensively. In particular, the conditions under which the algorithm will converge to the optimum with probability 1 are given, for example in Mitra et al. (1986) and Hajek (1988). The main reason for favouring this algorithm is that we could come up with an idea of implementing it in such a way as to accommodate the complicated constraints of the problem (e.g., different periodicities of modules and instruments and the presence of crossings), while respecting the theoretical conditions known to guarantee convergence.

Simulated annealing is a probabilistic search method: it moves in the state space from one element to the next by applying small random changes to the current state in the search for the minimum of a certain cost function. If the new state has a lower value of the cost function, it is accepted. If it has a higher cost, it can still be accepted with some small probability. Allowing a higher cost to be accepted in this way allows the algorithm to climb out of local minima of the cost function, so that it can converge towards the global minimum. The probability of accepting a state with a higher cost than the previous one depends on the difference in cost and on a 'cooling schedule', given by the 'temperature' T_t. This temperature decreases towards zero at a specified rate. Accepting a state with a higher cost than the previous one becomes increasingly less likely as the temperature decreases. This enforces convergence to the optimum as the system 'cools down'. Mitra et al. (1986) note that the condition $T_t \rightarrow 0$ is not sufficient for the resulting time-inhomogeneous Markov process to converge to the optimum. They prove that for discrete state spaces an assumption of T_t decreasing at a logarithmic rate, that is $T_t = \gamma / \log(t + t_0 + 1)$, with a suitable lower bound set for γ, is sufficient for this convergence. These results are in line with Hajek (1988), who gives a necessary and sufficient condition on the cooling schedules for convergence, and also suggests a logarithmic cooling schedule. For continuous state spaces, faster cooling rates may be sufficient for an appropriately chosen distribution of random changes from the current state, see, for example Tsallis and Stariollo (1996).

3.2.2. Implementing Simulated Annealing

Simulated annealing can be described as follows: starting with some arbitrary (but admissible) matrix $\mathbf{A}_0$, the following steps are repeated until no further cost reductions are achieved. Setting $C(\mathbf{A}) = C(\mathbf{A},\mathbf{n}^{opt}(\mathbf{A}))$:

i. *Proposal of a new state*. In each step t from the current $\mathbf{A}_t$, a new composition $\mathbf{A}_{t+1}$ is proposed by randomly perturbing $\mathbf{A}_t$. If $\mathbf{A}_{t+1}$ is not admissible, it is rejected by setting $\mathbf{A}_{t+1} = \mathbf{A}_t$

ii. *Sample size optimisation*. For the new candidate state $\mathbf{A}_{t+1}$, the simplex algorithm determines $\mathbf{n}_{t+1}^{opt} = \mathbf{n}^{opt}(\mathbf{A}_{t+1})$, which allows calculation of $C(\mathbf{A}_{t+1}) = C(\mathbf{A}_{t+1}, \mathbf{n}^{opt}(\mathbf{A}_{t+1}))$.
iii. *Acceptance of the proposal.* If the proposal does not increase the cost, that is $C(\mathbf{A}_{t+1}) \leq C(\mathbf{A}_t)$, it is accepted. If the cost increases, that is $C(\mathbf{A}_{t+1}) \geq C(\mathbf{A}_t)$, then the step may still be accepted, with the probability of acceptance being given by $\exp(-[C(\mathbf{A}_{t+1}) - C(\mathbf{A}_t)]/T_t)$, where $T_t = \gamma/\log(t + t_0 + 1)$.
iv. If accepted, $\mathbf{A}_{t+1}$ becomes the new current state, and the new sample size allocation is given by $\mathbf{n}_{t+1}^{opt} = \mathbf{n}^{opt}(\mathbf{A}_{t+1})$.

A further assumption needed for convergence (with probability 1) to the optimum — in addition to the assumption of the logarithmic rate of the cooling schedule, concerns the proposal of $\mathbf{A}_{t+1}$ from the current state $\mathbf{A}_t$ (see point i above). This assumption ensures that, were the temperature frozen at some fixed value, the resulting time-homogeneous Markov process would have a limit distribution. Mitra et al. (1986) note that if, additionally, $T_t \rightarrow 0$ and $T_{t+1} < T_t$ are assumed, the limit distribution would be concentrated on the optimum. This would, however, require that at each fixed value of the temperature T_{t} a number of steps be undertaken large enough that this limit distribution is reached (see e.g., Angelis et al. 2001).

'Proposal-symmetry' condition: the conditional probability of composition $\mathbf{A}'$ *being proposed if the current state is* $\mathbf{A}$, $P(\mathbf{A}'|\mathbf{A})$), *equals the probability of composition* $\mathbf{A}$ *being proposed if the current state is* $\mathbf{A}'$, $P(\mathbf{A}|\mathbf{A}')$). Thus $P(\mathbf{A}'|\mathbf{A}) = P(\mathbf{A}|\mathbf{A}')$. *(See e.g., Mitra et al. 1986)*

We describe below a random perturbation mechanism that respects this symmetry condition and the constraints from Subsections 2.1–2.3.

We use the notation introduced in Section 2: single modules correspond to rows $\{1, \ldots, m\}$ of the matrix $\mathbf{A}$, and crossings correspond to rows $m + 1, \ldots, m'$. We also make use of the concept of a 'complete periodic subset of instruments'. We denote as $J_{j_0}(i_0)$ a subset of a parent instrument that is complete for the module in i_0, and to which the instrument in j_0 belongs (see Subsection 2.3). (For example, if the module in row i_0 is of quarterly frequency and the instrument in column j_0 belongs to a monthly parent instrument, and if j_0 corresponds, say, to the February sibling, then $J_{j_0}(i_0)$ consists of the column indices of the siblings of the same parent instrument corresponding to February, May, August and November.) Finally, we say that the module in row i_0 is 'switched on' in the instrument in column j if $a_{i_0,j} = 1$, while if $a_{i_0,j} = 0$ we say that it is 'switched off'. We also say that we 'reverse a module's switch position' if we set $a_{i_0 j}$ to $1 - a_{i_0,j}$.

For admissibility we must respect the following properties:

a. If a mandatory crossing is switched on for some instrument, then it must also be switched on for all modules that are members of the crossing for the same instrument.
b. For each row of the matrix $\mathbf{A}$ corresponding to a module (or a crossing) of a certain periodicity, there is at least one complete periodic set of instruments for this module (or crossing) that is switched on. This is necessary in order to be able to allocate some sample size to this module.

c. Possible further constraints, such as those discussed in Subsection 2.4, are satisfied.

From a practical point of view, the value of the parameter γ of the cooling schedule should be chosen in such a way as to strike a good balance between an in-depth local search, that is exploring the area in the immediate vicinity of the current state space, and a global search, that is exploring areas of the state space that have not yet been sufficiently searched. For example, a value of $\gamma = 0$ would enforce quick convergence to the next local minimum, but would leave other areas of the state space unexplored. A quick apparent convergence is therefore not necessarily desirable. According to the exposition in Cicirello (2007), Lam and Delosme (1988) proposed a so called *"D-equilibrium"* as a trade-off between speed of convergence to an optimum and quality of this optimum in simulated annealing. Their analysis suggested that this is achieved when the acceptance rate, averaged across the previous 500 steps of the algorithm, is kept around 44 %. One might set the parameter γ in our algorithm so as to achieve this target. Reference is made to other approaches to nonmonotonic cooling schedules (alternating cooling and reheating) in Blum and Roli (2003).

A possible stopping rule for the algorithm would be one of the type: 'exit when the reduction in cost achieved during the last t_{crit} steps did not exceed eps_{crit}', for values of t_{crit} and eps_{crit} chosen by the user. The choice of the initial composition $\mathbf{A}_0$ may affect the number of steps required to achieve the optimum. A possible approach is to start with a composition in which each module participates in all siblings of all instruments that have a period required by the module, provided the constraints of Subsubsection 2.4.1 are also respected. If an admissible initial composition is found, this guarantees the existence of a solution to the optimisation problem: since the set of admissible compositions is nonempty and finite, the minimum will be attained. In more sizeable problems, that is those with a large number of instruments, modules and crossings, an alternative possible strategy would be to first find the optimum for a reduced problem, and then to use this to obtain a starting value for the full problem.

3.2.3. The Proposal of a New State: Simple Cases

The first step (i) in the implementation of simulated annealing, as set out in the previous section — the proposal of $\mathbf{A}_{t+1}$ from $\mathbf{A}_t$ at step t — can be defined as consisting of three phases:

1) Choose a row (module or crossing) i_0 at random from the uniform distribution on $\{1, \ldots, m'\}$ and a column (instrument) j_0 from the uniform distribution on all instruments that are of an appropriate periodicity for this module (i.e., have the same period or a period that is an integer fraction thereof).
2) Modify a_{i_0, j_0} and other related elements of $\mathbf{A}_t$ (see below for details).
3) If $\mathbf{A}_{t+1}$ is not admissible, discard it and set $\mathbf{A}_{t+1} = \mathbf{A}_t$.

Phase (2), the modification of a_{i_0, j_0} and other related elements, must be done in such a way that the 'proposal-symmetry' condition is satisfied.

To illustrate the idea, let us first consider the simple case where there is a unique frequency and there are no crossings. In this case, there are no further 'related elements' to be taken into account in Phase (2). It is only a_{i_0, j_0} that is modified by having its switch position reversed, that is by being set to $1 - a_{i_0,j_0}$. Then, starting at some state $\mathbf{A}$, the probability of proposing a neighbouring state (defined as those to which the probability of moving from $\mathbf{A}$ is positive), say $\mathbf{A}'$, is equal to $1/(mk)$. The same holds for the probability of proposing $\mathbf{A}$ when starting at $\mathbf{A}'$. Thus, the condition is fulfilled.

Let us now consider a slightly more complicated case, where there is one crossing. The constraint created by the presence of the crossing is that when the crossing is switched on, the modules belonging to that crossing must also all be switched on. This is not true the other way around: modules have different constraints on their sample sizes and may, moreover, be members of many crossings. They should therefore be given the freedom to participate in an instrument, even if a crossing to which they belong is not in this instrument. Crossings require the joint presence of modules, but do not set any constraints on their joint absence.

Assuming there is a crossing, and that the row selected in Phase (1) corresponds to a module, that is if $i_0 \leq m$, Phase (2) may be defined in exactly the same way as in the case considered above; the argument on symmetry still applies, as may be easily verified. We now consider the case where in Phase (1) a crossing, rather than an individual module, is selected (i.e., if $i_0 > m$). To illustrate this situation, let us assume that M_1 in Table 1 is also of annual periodicity and that $i_0 = 5$, which means G_1 was selected. Let us also assume that an instrument is selected that is switched off, that is $a_{5,j_0} = 0$, for example $j_0 = 3$. The idea is to decide in a randomised way whether the crossing will be switched on: we will switch it on (by setting $a_{5,3} = 1$) with some probability p_0, or leave everything unchanged (the reason for doing so will become apparent at the end of this paragraph). If we switch it on, then this obviously implies that the members of G_1 should be switched on as well, that is we would also set $a_{2,3} = a_{4,3} = 1$. Ignoring rows and columns that are not involved, we denote this state by $\mathbf{A}'$. Note, however, that, depending on which of the members ($i = 2$ and $i = 4$) are already switched on at the current step t, there are different states $\mathbf{A}$ that may lead to the same state $\mathbf{A}'$. In fact, if the crossing has L members, then there are 2^L different states $\mathbf{A}$ that could all lead to $\mathbf{A}'$. In order to satisfy the symmetry condition, we need to be able to get back to any state $\mathbf{A}$ that may have yielded $\mathbf{A}'$. Assuming that each one of these states is to be proposed from $\mathbf{A}'$ with equal probability, this probability should be $1/2^L$. This can be achieved by switching off each member of the crossing independently with probability $1/2$. This implies that, in order to satisfy the symmetry condition, we should also choose $p_0 = 1/2^L$ in our initial randomised decision on whether to have the crossing switched on or off.

Greater caution is required if there are two or more crossings: if at the step at which, say, the first crossing will be switched off (in a column j_0), some of its members are also members of some other crossing that is currently switched on in j_0, then these modules cannot be switched off, as this would lead to an inadmissible composition. Thus, they should not be counted in the different states $\mathbf{A}$, from which $\mathbf{A}'$ may be obtained. Such modules should therefore not be counted in L (which will thus be a random variable).

If there are a number of different frequencies but no crossings, the situation is simple again: for each selected module i_0 and instrument j_0, reverse the switch position of the module in all instruments in a complete periodic set for module i_0, the one to which j_0 belongs, that is, change the setting $a_{i_0,j}$ to $1 - a_{i_0,j}$ for all $j \in J_{j_0}(i_0)$.

3.2.4. The Proposal of a New State: The General Case

In the general case, where there are both different frequencies and crossings, the approach follows the same principle but is significantly more complicated, due to the fact that a crossing and its members may have different frequencies. Only Phase (2) from the previous subsection needs to be modified to accommodate this situation.

Phase (2) of the first step (i) of the implementation of simulated annealing can be defined by distinguishing between two separate cases:

I) If a single module was chosen (i.e., if $i_0 \leq m$), then all instruments in the complete subset to which the module belongs should have their switch position reversed, that is, if $a_{i_0,j_0} = 0$ then set $a_{i_0,j} = 1$ for all $j \in J_{j_0}(i_0)$. Similarly, if $a_{i_0,j_0} = 1$ then set $a_{i_0,j} = 0$ for all $j \in J_{j_0}(i_0)$.

II) If a crossing of period p_{i_0} was chosen (i.e., if $i_0 > m$), then for all its members M_i that have a period p_i, set $r_{i|i_0} = p_i/p_{i_0}$. Then for each M_i there is a partition of $J_{j_0}(i_0)$ into $r_{i|i_0}$ subsets, each of which is complete for M_i. Let us call these subsets $J_{j_0}(i, 1), \ldots, J_{j_0}(i, r_{i|i_0})$. The number of these subsets equals $r_{i|i_0}$. Let $\rho_{i|i_0}$ be the number of subsets excluding those subsets $J_{j_0}(i,k)$ for which M_i is a member of some further crossing $G_{i_1}(i_1 \neq i_0)$ that is currently switched on, that is, $a_{i_1,j} = 1$ for $j \in J_{j_0}(i, k)$. Thus

$$\rho_{i|i_0} = \#\{J_{j_0}(i,k), k = 1, \ldots, r_{i|i_0} | a_{i_1,j} = 0, \forall i_1 > m, i_1 \neq i_0 : M_i \in G_{i_1}, j \in J_{j_0}(i,k)\},$$

where # denotes cardinality. Note that if the module in row i is a member of another crossing that is currently switched on for all $j \in J_{j_0}(i_0)$, then $\rho_{i|i_0} = 0$.

i. If $a_{i_0,j_0} = 0$, then let $L = \sum \rho_{i|i_0}$, where the summation extends over all members of the crossing. Then, with probability $(1/2)^L$, for all $j \in J_{j_0}(i_0)$, set $a_{i_0,j} = 1$ and $a_{i,j} = 1$ for all rows i containing modules that are members of the crossing; else (with probability $1 - (1/2)^L$), set $\mathbf{A}_{t+1} = \mathbf{A}_t$.

ii. If $a_{i_0,j_0} = 1$, then set $a_{i_0,j} = 0$, for all $j \in J_{j_0}(i_0)$. Also, independently for all rows i containing modules that are members of the crossing considered (in row i_0), and independently for each of the subsets $J_{j_0}(i,k)$ of the partition $J_{j_0}(i, 1), \ldots, J_{j_0}(i, r_{i|i_0})$ of $J_{j_0}(i_0)$, for which the module in row i is not a member of another crossing (beyond the one in i_0), which (other crossing) is currently "switched on", with probability equal to $1/2$ set $a_{i,j} = 0$ for all $j \in J_{j_0}(i,k)$.

Note that the proposal mechanism is such that the constraints set out in Subsections 2.2 and 2.3 are already incorporated in Phase (2), while the constraints from Subsection 2.4 are only checked *a posteriori* in Phase (3), once a new proposal has been made.

Different ways of updating the composition matrix **A** can of course affect the numerical efficiency of the search algorithm. We do not have an optimal search algorithm at our disposal.

On the basis of the above, it is possible to prove the following special case. We are not currently able to provide a proof for cases where other constraints, such as dependencies between modules, are taken into account.

Proposition. *The proposal mechanism described in Steps 1–3 above respects the proposal-symmetry condition, subject to the constraints described in Subsections 2.1–2.3.*

The proof of this proposition may be found in the Appendix.

3.3. Appropriateness of the Simulated Annealing Algorithm

The algorithm, consisting of simulated annealing and simplex, can also be used in the design of an efficient SQD, thus resolving a problem to which feasible solutions exist in the literature only for the case of a very limited number of variables (or modules).

For the single survey SQD, and under simple random sampling, Chipperfield and Steel (2009; 2011) consider the optimal allocation of sample size to minimise cost, subject to precision constraints. The most efficient design is determined on the basis of best linear unbiased estimation (possible under simple random sampling and given correlations between all variables), and therefore comprises all $2^m - 1$ possible patterns (instruments). For the case being studied here — an integrated survey with a complex and modular sampling design and a baseline estimation approach involving HT estimators for single instruments (samples), and composite HT estimators combining data from different instruments — we are also aiming to minimise cost, under similar precision constraints. However, our minimisation is over all possible $c_k = \binom{2^m - 1}{k}$ combinations of k instruments, and over the associated sample size allocations. Noting that the use of $2^m - 1$ samples is not practical even for moderate m, Chipperfield and Steel (2009) also propose the choice of a limited number of k best patterns according to a ranking of all patterns based on their relative estimation efficiency. They thus circumvent the more difficult question of optimisation over all c_k combinations, but at the expense of a loss of efficiency.

Our specification of a fixed number of instruments, and the optimisation over all c_k combinations, coincides with the approach adopted by Adiguzel and Wedel (2008). They, however, consider the distinct problem of finding the single survey SQD that minimises the Kullback-Leibler distance (see Kullback and Leibler 1951) to the full questionnaire, while assuming equal sample sizes for all k instruments. Without the assumption of equal sample sizes, the minimum is attained by the full questionnaire. Thus, their approach does not include optimal sample size allocation. A random search algorithm, a modification of the Fedorov algorithm (see e.g., Cook and Nachtsheim 1980), is then used to find the optimal instrument composition.

Applying this algorithm to our problem, including the simultaneous optimisation over sample sizes, would mean that a single full step of the algorithm should examine each of the $(2^m - 1)$ possible compositions for each of the k instruments. It would thus require the same computation time as $k(2^m - 1)$ steps of the annealing algorithm. For example, with 30 modules and three instruments, this figure is around three billion. The number of modules in an integrated social survey system could range from 50 to 200, making it impractical to apply this algorithm.

There are alternative ways of implementing simulated annealing. In particular, future research may draw on the extensive literature on the application of simulated annealing in the optimisation of experimental designs (see Meyer and Nachtsheim 1988). Two examples of work in this area are the discussions on sequential and nonsequential exchange methods in Lejeune (2003) and on exchange and interchange steps in Jansen et al. (1992), which may lead to improvements to the current proposal mechanism. Cooling rates declining faster to zero, as discussed in Tsallis and Stariollo (1996), and the analysis of the impact of the values of the parameters involved (see e.g., Angelis et al. 2001), are also interesting topics for further research.

Finally, the use of other heuristic and metaheuristic algorithms, such as genetic algorithms and other population-based methods (see e.g., Blum and Roli 2003 for a brief discussion of and further references to work on these types of algorithms) could be explored.

4. Modular Design with Complex Sampling

In Section 2, we assumed the use of simple random sampling. In this section, we extend the approach to include complex sampling designs. We first set out how the precision requirements for estimating θ are translated into constraints on the sample sizes, and then explain how the optimisation framework can be modified accordingly.

4.1. Adjusting Sample Size Requirements

4.1.1. Complex Sampling Design for Independent Samples

Estimating θ from a single sample under some complex design, which, for example, involves stratification and/or multistage sampling, leads to a different variance $V'(\hat{\theta})$. Assuming normality for $\hat{\theta}$, the precision constraint $V'(\hat{\theta}) \leq \left(e/z_{1-\alpha}\right)^2$ is then satisfied by the constraint $n/d \geq n^*$, where d is the customary design effect (defined as $d = V'(\hat{\theta})/V(\hat{\theta})$, with $V(\hat{\theta})$ being the variance of the same estimator under simple random sampling), and n^* is as in (1).

When θ is estimated by a weighted average of estimates from different instruments, as in Subsection 2.1, then the weights should be modified to $w_i = \left(n_i/d_i\right)/\sum\left(n_j/d_j\right)$, where d_i is the design effect for the ith sample. It can be shown that for independent simple random samples and for this choice of weights, the composite estimator $\hat{\theta} = \sum w_i\hat{\theta}_i$ has minimum variance $V'(\hat{\theta}) = \sigma_\theta^2/\sum\left(n_i/d_i\right)$, provided that sampling fractions n_i/N are negligible. The precision requirement $V'(\hat{\theta}) \leq \left(e/z_{1-\alpha}\right)^2$ is then satisfied if

$$\sum\left(n_j/d_j\right) \geq n^*.$$

In the case where the design effects for the different samples are identical, the weights of the composite estimator are $w_i = n_i/\sum n_j$, and the precision requirement is satisfied by $\left(\sum n_i\right)/d \geq n^*$, where d is the common design effect.

4.1.2. Complex Sampling Design for Coordinated Samples

If the samples of the various instruments are negatively coordinated, for example by splitting a sample into subsamples, one for each instrument, then $V'(\hat{\theta}) \leq \sigma_\theta^2/\sum\left(n_i/d_i\right)$, due to the negative covariance term between the estimates $\hat{\theta}_i$. The precision requirement is, therefore, again satisfied if

$$\sum \left(n_j/d_j\right) \geq n^*.$$

Let us finally consider an example where there is positive coordination between the samples and show how this case may also be embedded in the current framework. We assume that there is some overlap between L samples $S_1, \ldots, S_L$, of sizes $n_1, \ldots, n_L$, as is the case in surveys with rotating samples. Let us denote by $n_{i\cap j}$ the number of sample units in $S_i \cap S_j$, let $n_{i\cap j} = \tau_{ij} n_i$ for some appropriate τ_{ij}, and assume that all subsets reflecting common and noncommon parts of these samples are drawn independently from each other by simple random sampling. It is then logical to pool estimates over time, as is done for example when estimating a yearly average from four quarterly estimates. Putting $\hat{\theta} = \sum w_i \hat{\theta}_i$, where $w_i = n_i/n_+$, $n_+ = \sum n_i$, $\tau_{i+} = \sum_j \tau_{ij}$ and $\tau_{avg} = \sum_i \tau_{i+} w_i$, we obtain

$$V(\hat{\theta}) \leq \frac{\sigma_\theta^2}{n_+}\left(n_+^{-1}\sum_{i,j} n_{i\cap j}\right) = \frac{\sigma_\theta^2}{n_+}\left(\sum_i \frac{n_i}{n_+}\tau_{i+}\right) = \frac{\sigma_\theta^2}{n_+}\tau_{avg}.$$

In the specific case of a longitudinal survey with L wave, for which the sample sizes $n_1, \ldots, n_L$ are restricted to be equal to each other, we have $\tau_{avg} = L^{-1}\sum \tau_{i+}$. It is then possible to adopt a 'design' effect $d_j = \tau_{avg}$, such that the precision requirement is satisfied if $\sum \left(n_j/d_j\right) \geq n^*$.

4.2. Adapting the Optimisation Framework

In all the cases described in Subsection 4.1, the precision requirements for a module M_a are satisfied by

$$\sum_{\Theta_a} \left(n_j/d_{a,j}\right) \geq n_a^*, \tag{8}$$

where $d_{a,j}$ is the module's design effect for instrument j, n_a^* is obtained as in (3), and the summation extends over Θ_a, the set of all instruments containing the module M_a.

A way of representing these m' linear constraints (one for each module and each crossing), on the sample sizes $n_1, \ldots, n_k$ of the various samples, is by introducing the $m' \times k$ 'design-composition matrix' $\mathbf{R}$, with elements $[\mathbf{R}]_{i,j} = a_{i,j} d_{i,j}^{-1}$, where $a_{i,j}$ are the elements of the composition matrix $\mathbf{A}$ and $d_{i,j}$ those of the design-effect matrix $[\mathbf{D}]_{i,j}$. For a crossing in row i_0, we set $d_{i_0 j_0} = \max_i\{d_{ij_0}\}$, where the maximum is over all members of the crossing. In the special case where all samples are independent and drawn using simple random sampling, and a weighted average of HT estimators is used, all $d_{i,j} = 1$, and we obtain $\mathbf{R} = \mathbf{A}$.

Generally, the constraints in (8) may now be expressed as

$$\mathbf{Rn} \geq \mathbf{n}^*, \ \mathbf{R} \text{ is an } m' \times k \text{ matrix}, \tag{9}$$

where the inequality is to be understood componentwise, $\mathbf{n} = (n_1, \ldots, n_k)'$ is the vector of the actual sizes of the samples associated with the k instruments, and $\mathbf{n}^* = (n_1^*, \ldots, n_{m'}^*)'$ is the vector of the minimum sample sizes for the m' modules and crossings required under simple random sampling, that is, those given in (1). This relation now replaces (5).

Furthermore, (7) is now replaced by

$$\min{}_{\mathbf{A}} C(\mathbf{A}, \mathbf{n}^{opt}(\mathbf{A}, \mathbf{D})) \text{ under the conditions : } \mathbf{A} \text{ admissible}, \ (\mathbf{R}\mathbf{n}^{opt}(\mathbf{A}, \mathbf{D}))_i \geq n_i^*, i = 1, \ldots, m' \text{ and } \mathbf{B}\mathbf{n} = \mathbf{0} \quad (10)$$

where $\mathbf{n}^{opt}(\mathbf{A}, \mathbf{D})$ is the admissible vector of optimal instrument sample sizes and $\mathbf{n}^{opt}(\mathbf{A}, \mathbf{D}) \neq \mathbf{n}^{opt}(\mathbf{A})$.

In other words, the algorithm is defined exactly as previously on the space of all composition matrices **A**, with the only modification being that the sample size constraints are now formulated using the design-composition matrix **R** instead of the composition matrix **A**.

5. An Illustrative Example

As an illustrative example of how the approach presented above could be used, we consider a hypothetical reorganisation of the Labour Force Survey (LFS) using three quarterly instruments, in place of the current single questionnaire, which is administered quarterly.

The blocking of the LFS variables into 30 modules was prepared by Eurostat (see Table 2). The first six of these modules represent demographic and household characteristics, and will be present in all available instruments. A further group of ten modules contains the structural variables, that is, those on which information must be collected on an annual basis (in accordance with EU regulations). Of the modules containing quarterly variables, five are needed for the definition of "ILO-Unemployment" (unemployment as defined by the International Labour Organization). We therefore group these into a single crossing. For the purposes of this illustration, we split the remaining nine quarterly modules into two crossings: 'Employment conditions 1' ('Empl_cond1'), containing six modules, and 'Employment conditions 2' ('Empl_cond2') with three modules, as indicated in Table 2.

For each module, Eurostat defined the dependencies on other modules and calculated, based on historically observed proportions of respondents, the average respondent burden l_γ as the sum (across individuals in the sample) of the number of questions in M_γ answered by each individual, divided by the total number of individuals in the sample. For the purposes of this illustration, we assume the fix costs $C_j^{(f)}$, $\beta_0^{(hh)}$ and $\beta_0^{(ind)}$ to be 0, and the coefficients $\beta_1^{(hh)}$ and $\beta_1^{(ind)}$ to be equal to 1. The calculations are based on the data for Portugal. The reason for choosing Portugal was that it has a population size close to the EU median.

The required sample sizes are calculated so as to satisfy EU regulations that apply to Portugal, and on the assumption that all instruments are administered to independent samples, drawn using simple random sampling. According to Commission Regulation (EC) No 377/2008, for structural variables, the relative standard error (assuming simple random sampling) of any yearly estimate representing one percent or more of the working-age population must not exceed nine percent for countries with a population of between one million and 20 million inhabitants. Note that, if simple random sampling is assumed, the sample sizes required to satisfy these precision requirements may be derived from the variance of estimating proportions of a specified magnitude, and, therefore, do not need to be estimated by means of pilot samples or other similar single surveys. This implies a

Table 2. *Optimal instrument composition, burden and sample sizes for the illustrative example from Section 5*

Instrument							I1.1	I1.2	I1.3	I1.4	I2.1	I2.2	I2.3	I2.4	I3.1	I3.2	I3.3	I3.4
Final sample size							22,050	22,050	22,050	22,050	7,367	7,367	7,367	7,367	12,585	12,585	12,585	12,585
Burden							35	35	35	35	37	37	37	37	33	33	38	41
Module							Instrument composition (modules)											
Name	Description	Depends on	Burden	Period	Sample size: Required	Sample size: Final												
GEO	Core geographic module		1.25			168,008	1	1	1	1	1	1	1	1	1	1	1	1
DEMO	Core demographic module		4.62			168,008	1	1	1	1	1	1	1	1	1	1	1	1
DEMOHAR	Common harmonised demographic module (migration)		4.00			168,008	1	1	1	1	1	1	1	1	1	1	1	1
HHOLD	Household composition		3.75			168,008	1	1	1	1	1	1	1	1	1	1	1	1
GEN	Inter-generational information		2.00			168,008	1	1	1	1	1	1	1	1	1	1	1	1
GENHAR	Common harmonised inter-generational module		9.81			168,008	1	1	1	1	1	1	1	1	1	1	1	1
ATWORK	Employment status - At work	DOME	1.75	3	168,005	168,008	1	1	1	1	1	1	1	1	1	1	1	1
ABSWORK	Employment status - Absences from work	ATWORK, STAPRO	0.03	3	168,005	168,008	1	1	1	1	1	1	1	1	1	1	1	1
STAPRO	Contract - Professional status	AKROWT	0.74	3	168,005	168,008	1	1	1	1	1	1	1	1	1	1	1	1
WORKSEEK	Looking for work - Search for work	ABSWORK, ATWORK, DEMO	2.20	3	168,005	168,008	1	1	1	1	1	1	1	1	1	1	1	1
WORKSEEKAVAIL	Looking for work - Availability	WORKSEEK	0.25	3	168,005	168,008	1	1	1	1	1	1	1	1	1	1	1	1
OCCUP	Characteristics of the workplace - Occupation	ATWORK, STAPRO	0.90	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
WORKHOURUSUAL	Working hours - Usual hours	ATWORK	0.48	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
WORKHOURACTUAL	Working hours - Actual hours	ATWORK, STAPRO, WORKHOURUSUAL	1.39	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
WORKMORE	Wished hours	ATWORK	0.52	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
ECONACT	Characteristics of the workplace - Economic activity	ATWORK	0.67	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
SIZEFIRM	Characteristics of the workplace - Enterprise characteristics	STAPRO	0.33	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
CONTRACT	Contract - Characteristics of contract	STAPRO	0.74	3	79,805	79,808	0	0	0	0	1	1	1	1	1	1	1	1
WORKPLACE	Characteristics of the workplace - Workplace	ATWORK	1.23	3	79,805	79,808	0	0	0	0	1	1	1	1	1	1	1	1
JOBSTART	Start of job	ATWORK, STAPRO	0.12	3	79,805	79,808	0	0	0	0	1	1	1	1	1	1	1	1
MAINEMPLSTAT	Main employment status	DEMO	1.00	12	12,205	12,585	0	0	0	0	0	0	0	0	1	0	0	0
WORKTIME	Working times	ATWORK, STAPRO	1.95	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	0	1
JOBSECOND	Second job	ATWORK	0.45	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	0	1
JOBLOOKING	Looking for other job	ATWORK, DEMO	0.44	12	12,205	12,585	0	0	0	0	0	0	0	0	0	1	0	0
WORKEXP	Previous work experience	ATWORK	1.24	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	1	0
EMPLOY1Y	Employment Situation 1 year before	DEMO	3.72	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	0	1
EDUCATT_LEVEL	Educational attainment level	DEMO	1.11	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	0	1
EDUCATT_CHAR	Educational attainment main characteristics	DEMO, EDUCATT_LEVEL	1.21	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	0	1
EDUCPART4W	Participation in education and training (4 weeks)	DEMO	1.83	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	1	0
EDUCPART4W_CHAR	Characteristics of participation in education and training (4 weeks)	EDUCPART4W	2.10	12	12,205	12,585	0	0	0	0	0	0	0	0	0	0	1	0

Crossing					Instrument composition (crossings)											
Name	Members	Period	Sample size: Required	Sample size: Final	I1.1	I1.2	I1.3	I1.4	I2.1	I2.2	I2.3	I2.4	I3.1	I3.2	I3.3	I3.4
Demogr./Housh.	GEO,DEMO,DEMOHAR,HHOLD,GEN,GENHAR			168,008	1	1	1	1	1	1	1	1	1	1	1	1
ILO_Unempl	ATWORK,ABSWORK,STAPRO,WORKSEEK,WORKSEEKAVAIL	3	168,005	168,008	1	1	1	1	1	1	1	1	1	1	1	1
Empl_cond1	OCCUP,WORKHOURUSUAL,WORKHOURACTUAL,WORKMORE,ECONACT,SIZEFIRM	3	117,665	117,668	1	1	1	1	1	1	1	1	0	0	0	0
Empl_cond2	CONTRACT,WORKPLACE,JOBSTART	3	79,805	79,808	0	0	0	0	1	1	1	1	1	1	1	1

required total sample size of 12,205. For the variables in the ILO-Unemployment crossing, according to Council Regulation (EC) No 577/98, the relative standard error when estimating the change in subpopulations representing five percent of the working-age population between two successive quarters must not exceed three percent for countries with a population of between one million and 20 million inhabitants. This implies a required yearly sample size of 168,005 (i.e., 42,001 per quarter). This regulation dominates the others, which require smaller sample sizes. For the other two crossings containing quarterly variables, and in order to introduce some further precision differentiation, we assume sample sizes corresponding to the same level of precision in estimating the change between quarters (3 %) but for subpopulations of seven percent and ten percent of the working-age population respectively. This implies annual sample sizes of 117,665 and 79,805.

The benchmark composition used as the standard against which to compare the cost reduction achieved by the reorganisation is the composition where all modules (including structural modules) are present in all instruments. This is equivalent to the traditional approach of administering a single questionnaire to all sample units. As there is no differentiation between instruments, the total yearly sample size equals the maximum of the required sample sizes, that is, 168,005, and the total yearly data collection cost (taking into account the modules' respondent burden) is around 8.70 million person-questions. This benchmark composition was also used as the initial composition $\mathbf{A}_0$ for the annealing algorithm.

The optimisation algorithm is programmed in R (R Core Team 2013) and used R-package *lpSolve* for the simplex optimisation (Berkelaar et al. 2013). In this example, it was run for 50,000 steps, using the benchmark composition as the initial state. The cost reduction achieved by the optimisation is illustrated in Figure 1 below. It can be seen that most of the reduction has already been achieved after around 4,000 steps. The minimum is

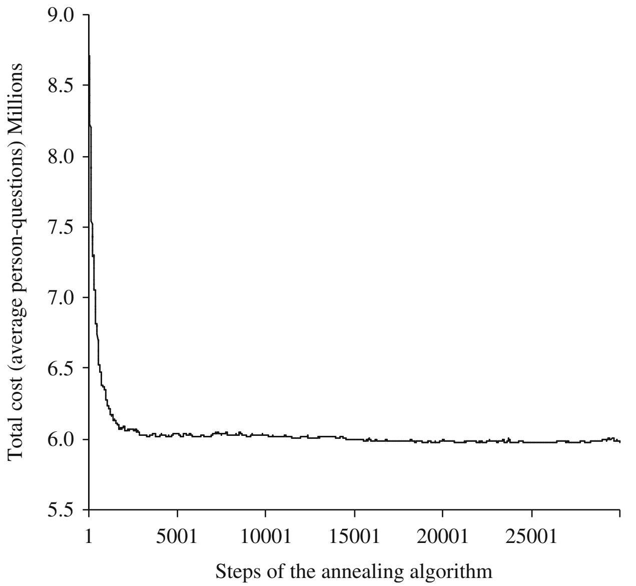

Fig. 1. Cost evolution (first 30,000 out of 50,000 steps) for compositions visited by the simulated annealing algorithm when searching for the optimal composition for the illustrative example described in Section 5.

reached after approximately 28,000 steps, with a cost of 5.97 million person-questions, a reduction of around 31 % from the cost of using the benchmark composition. Had we included a component representing the fixed cost for each instrument, increasing the number of instruments would, of course, have cancelled out some of this cost reduction. The purpose of this article is, however, to illustrate the modular approach rather than to propose a definitive alternative design.

At this 'approximate optimum', the 'ILO-Unemployment' modules are included in all three instruments, and allocated a total sample size (168,008) that is almost identical to the required one. The 'Empl_cond1' crossing participates in the first two instruments, and the 'Empl_cond2' crossing in the last two. For both crossings, the total sample sizes achieved in the optimum (117,668 and 79,808, respectively) are almost identical to the sample sizes required for these modules. The annual modules are distributed across the quarters of the instrument which has the sample size (12,585 per quarter) closest to what is required for these modules (12,205), in such a way as to respect module dependencies. These modules are therefore slightly 'oversampled'. This oversampling would, of course, have been avoided, resulting in a further cost reduction, had one more annual instrument been introduced. Note that with the distribution of modules to instruments described in Table 2, the annual modules and the modules in Empl_cond1 are not jointly administered to any sample units, implying that there is no information on the interaction between them. This could have been avoided by introducing a further crossing comprising the modules in Empl_cond1 and some or all of the annual modules, with a required sample size chosen accordingly. Satisfying this additional constraint would inevitably have meant giving up some of the cost reduction achieved by the design shown in Table 2.

6. Conclusion

In this article, we have formulated a modular design and have proposed an algorithm that could be used to restructure the questionnaires currently used for social surveys into instruments. A better balance could thus be achieved between precision requirements on the one hand, and the need to reduce the cost and burden of the survey on the other. In view of the extremely large number of possible instrument compositions, a random search algorithm has been developed to help find the optimal composition.

There are several methodological issues that should be acknowledged in relation to the modular design. It should be noted that these are all perennial difficulties associated with survey sampling, which apply to any design approach, and a more detailed investigation of these issues is beyond the scope of this study. Nevertheless, it is helpful to be clear about the tacit assumptions that we necessarily have to make as a result, for which we have generally adopted the standard practice.

When determining required sample sizes, the requisite design parameters σ_θ^2 for all θ (means and proportions) and all modules involved and the required design effects are likely to be unknown, and would need to be estimated. This is typically done by means of pilot samples or similar single surveys. For the estimation of design effects under complex survey designs, see, in particular, Gambino (2009) and Park and Lee (2004).

Practical discussions of the modular approach often note that certain features of the modular design may have an effect on various nonsampling errors. The typical errors include those related to the cognitive aspects of the instrument composition, the mode effects and the nonresponse. The cognitive effects relate to the respondents as well as to the data collection personnel. The possible consequences include confusion, increased training costs, measurement errors, and potential nonresponse. Mode effects exist if the obtained measurements differ according to the mode of data collection, with possible modes being, for example, face-to-face interview, computer-assisted telephone interview, and self-administered postal questionnaire. The nonresponse rate, both for unit and partial nonresponse, may vary depending on the items or the combination of the modules.

We do not explicitly account for nonsampling errors while addressing the two design questions (1) and (2) that we posed in Section 1, due to the lack of carefully studied empirical evidence. For example, suppose the nonresponse rate in a forthcoming survey is expected to be around 50% based on past experience. What is then the difference in cost and precision for example between (a) planning a sample that is twice as large as the required net sample, and (b) planning a sample that is three times as large and aiming at a 33% response rate in fieldwork? Meanwhile, we do implicitly assume that the modular design is based on the best current knowledge as to how to structure the items in the different modules. We assume that the constraints on instrument composition reflect choices made as to whether some modules should or should not be present simultaneously in the same instrument, or whether a module should only be assigned to a certain mode of collection. As was demonstrated in Section 2, the modular design approach can incorporate all such constraints.

We would like to emphasise that, in practice, a potentially large error in approximating the optimal design is unlikely to be the most critical concern, not least because the optimum is only such with respect to a certain design parameterisation, which can itself be challenged and certainly will be revised over time. Moreover, attention should be given to addressing and reducing the potential nonsampling errors that may come with the new system in future. In this respect, it is important to use a suitable benchmark in order to ensure a fair assessment under which potential shortcomings are balanced against gains. It is unrealistic to require the new system to meet certain unattainable ideals which today's standalone surveys also fail to achieve.

Moreover, a number of considerable changes would have to take place at national statistical institutes for it to become possible to implement the modular approach. These involve considerable investment in resources, and the complexity and cost of these organisational changes should not be underestimated. The changes fall into three main areas: (i) Information structure and production systems: data collection and processing require a greater degree of conceptual harmonisation and operational standardisation in order to make them comparable, shareable and reusable. Standardised database and warehouse solutions and metadata systems will need to replace end-to-end statistics-specific processes and management, in order to make the data and metadata accessible across the different statistical domains. (ii) Fieldwork management systems: these are needed both as a way of managing the numerous modules and instruments required in fieldwork and so as to be able to update or replace existing modules and instruments over time. (iii) Staff: staff need to adapt to a new situation where the data are collected in multiple instruments and stored, managed and shared using standard system solutions.

Finally, we would like to draw attention to the fact that while the approach presented here could allow efficiency gains to be made compared to the system currently in use (as illustrated in Section 5), and while it is reasonable to explore this (since today's system of social surveys would be the point of departure from which a gradual phasing in of a modular system would take place), a main strength of the modular design lies in its flexibility to meet new information requirements. The system could easily incorporate new modules (or eliminate redundant ones), scale up (or reduce) the precision requirements for individual modules or cross certain modules with each other. Thus, the potential of the new system is not limited to meeting today's information needs. One of its main attractions is that it can accommodate new information requests from policymakers in a much more flexible way, without the constraints of the present survey system.

Appendix

Proof of the Proposition on the Fulfilment of the Proposal-Symmetry Condition

Let us denote by $\pi(i_0,j_0)$ the probability of selecting an i_0 and a j_0 in one of the available complete sets of instruments. In the case that $\mathbf{A}'$ was generated from $\mathbf{A}$ in the case described in Subsubsection 3.2.4, Point I) above (i.e., the switch position of a module in row $i_0 \leq m$ was reversed), we have $P(\mathbf{A}'|\mathbf{A}) = P(\mathbf{A}|\mathbf{A}') = \pi(i_0,j_0)$, and the proposal mechanism respects the symmetry condition.

Similarly, let us now consider the case where $\mathbf{A}'$ was generated from $\mathbf{A}$ in the case described in Subsubsection 3.2.4 Point II) (i.e., a crossing in row $i_0 > m$ was selected). First, assume that all the modules contained in the crossing have the same periodicity, implying $r_{i|i_0} = 1$ for all its members. Moreover, for illustration purposes, assume the crossing has only $L = 2$ members, which are assumed not to be members of any other crossings that are currently switched on. Then, in the state, say $\mathbf{A}'$, in which the crossing is switched on, all members are also switched on, while, if the crossing is switched off, there are four possible states in relation to its two members, say $\mathbf{A}_1, \ldots, \mathbf{A}_4$, corresponding to the situations where none, both, or one of the two members are switched on (for all $j \in J_{j_0}(i_0)$). Then, in the case presented in Subsubsection 3.2.4 Point II.i), we get $P(\mathbf{A}'|\mathbf{A}_i) = \left(1/2^2\right)\pi(i_0,j_0)$, while in the case presented in point II.ii), we get $\mathrm{P}(\mathbf{A}_i|A') = \left(1/2\right)\left(1/2\right)\pi(i_0,j_0)$, and the proposal mechanism respects the symmetry condition. The argument is similar for $L > 2$: the number of states from which it is possible to generate $\mathbf{A}'$ equals 2^L, and the transition probabilities to and from $\mathbf{A}'$ are all equal to $1/2^L$. Note that modules that are currently members of other crossings that are switched on are fixed in their current switch position during the transitions considered.

In the case where a crossing may contain modules of different periodicities, again, the number of states from which it is possible to generate $\mathbf{A}'$ equals 2^L, and the transition probabilities to and from S' are all equal to $1/2^L$, which concludes the proof.

7. References

Adiguzel, F. and M. Wedel. 2008. "Split Questionnaire Design for Massive Surveys." *Journal of Marketing Research* 45: 608–617. Doi: http://dx.doi.org/10.1509/jmkr.45.5.608.

Angelis, L., E. Bora-Senta, and C. Moyssiadis. 2001. "Optimal Exact Experimental Designs with Correlated Errors Through a Simulated Annealing Algorithm." *Computational Statistics & Data Analysis* 37: 275–296. Doi: http://dx.doi.org/doi:10.1016/S0167-9473(01)00011-1.

Australian Bureau of Statistics. 2012. *Household Expenditure Survey and Survey of Income and Housing, User Guide*. Australia, 2009-10 (cat. no. 6503.0). Available at: http://www.abs.gov.au/ausstats/abs@.nsf/mf/6503.0 (accessed April 2016).

Berkelaar, M., et al. 2013. *lpSolve: Interface to Lp_solve v. 5.5 to solve linear/integer programs. R package version 5.6.10*. Available at: http://CRAN.R-project.org/package=lpSolve (accessed August 2013).

Blum, C. and A. Roli. 2003. "Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison." *ACM Computing Surveys* 35: 268–308. Doi:http://dx.doi.org/10.1145/937503.937505.

Chipperfield, J.O. and D.G. Steel. 2009. "Design and Estimation for Split Questionnaire Designs." *Journal of Official Statistics* 25: 227–244.

Chipperfield, J.O. and D.G. Steel. 2011. "Efficiency of Split Questionnaire Surveys." *Journal of Statistical Planning and Inference* 141: 1925–1925. Doi: http://dx.doi.org/10.1016/j.jspi.2010.12.003.

Chipperfield, J.O., M. Barr, and D.G. Steel. 2013. "Split Questionnaire Designs: Are They an Efficient Design Choice?" In Proceedings of the 59th ISI World Statistics Congress, 25–30 August 2013, Hong Kong. 311–316. Available at: http://2013.isiproceedings.org/Files/IPS033-P1-S.pdf (accessed June 2015).

Cicirello, V.A. 2007. "On the Design of an Adaptive Simulated Annealing Algorithm." In First Workshop on Autonomous Search, in conjunction with CP'2007, Providence, Rhode Island, USA.

Cochran, W.G. 1977. *Sampling Techniques*. New York: Wiley.

Cook, R.D. and C.J. Nachtsheim. 1980. "A Comparison of Algorithms for Constructing Exact D-optimal Designs." *Technometrics* 22: 315–324. Doi: http://dx.doi.org/10.1080/00401706.1980.10486162.

Cuppen, M.D.J., P. van der Laan, and W. van Nunspeet. 2013. "Reengineering Dutch Social Surveys: From Single-Purpose Surveys to an Integrated Design." *Statistical Journal of the International Association for Official Statistics* 29: 21–29. Doi: http://dx.doi.org/10.3233/SJI-130762.

European Commission. 2009. *Communication on the Production Method of EU Statistics: A Vision for the Next Decade*. Brussels: European Commission (COM(2009)/404). Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF (accessed January 2015).

European Statistical System Committee (ESSC). 2011. *New Conceptual Design for Household and Social Statistics*. Wiesbaden, Germany. Available at: https://www.destatis.de/EN/AboutUs/Events/DGINS/Document_Memorandum.pdf (accessed January 2015).

Gambino, J.G. 2009. "Design Effects Caveats." *The American Statistician* 63: 141–146.

Groves, R.M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.

Hajek, B. 1988. "Cooling Schedules for Optimal Annealing." *Mathematics of Operations Research* 13: 311–329. Doi: http://dx.doi.org/10.1287/moor.13.2.311.

Jansen, J., R.C.M.H. Douven, and E.E.M. van Berkum. 1992. "An Annealing Algorithm for Searching Optimal Block Designs." *Biometrical Journal* 34: 529–538. Doi: http://dx.doi.org/10.1002/bimj.4710340503.

Karlberg, M., F. Reis, C. Calizzani, and F. Gras. 2015. "A Toolbox for a Modular Design and Pooled Analysis of Sample Survey Programmes." *Statistical Journal of the International Association for Official Statistics* 31: 447–462. Doi: http://dx.doi.org/10.3233/SJI-150913.

Kirkpatrick, S., C.D. Gelatt, Jr., and M.P. Vecchi. 1983. "Optimization by Simulated Annealing." *Science* 220: 671–680.

Kullback, S. and L.A. Leibler. 1951. "On Information and Sufficiency." *Annals of Mathematical Statistics* 22: 79–86. Available at: http://www.jstor.org/stable/2236703.

Lam, J. and J. Delosme. 1988. "Performance of a New Annealing Schedule." In Proceedings of the 25th ACM/IEEE Design Automation Conference. 12–15 June 1988, Anaheim, CA, U.S.A. 306–311.

Lejeune, M.A. 2003. "Heuristic Optimization of Experimental Designs." *European Journal of Operational Research* 147: 484–498. Doi: http://dx.doi.org/10.1016/S0377-2217(02)00292-8.

Merkouris, T. 2004. "Combining Independent Regression Estimators from Multiple Surveys." *Journal of the American Statistical Association* 99: 1131–1139. Doi: http://dx.doi.org/10.1198/016214504000000601.

Merkouris, T. 2010. "An Estimation Method for Matrix Survey Sampling." In Proceedings of the Section on Survey Research Methods: American Statistical Association, July 31 to August 5, 2010, Vancouver, Canada. 4880–4886. Available at: https://www.amstat.org/Sections/Srms/Proceedings/y2010/ Files/308769_61580.pdf (accessed March 2016).

Merkouris, T. 2013. "Composite Calibration Estimation Integrating Data from Different Surveys." In Proceedings of the 59th ISI World Statistics Congress, 25–30 August 2013, Hong Kong. 205–210. Available at: http://2013.isiproceedings.org/Files/IPS020-P3-S.pdf (accessed March 2016).

Merkouris, T. 2015. "An Efficient Estimation Method for Matrix Survey Sampling." *Survey Methodology* 41: 237–262.

Meyer, R.K. and C.J. Nachtsheim. 1988. "Constructing Exact D-optimal Experimental Designs by Simulated Annealing." *American Journal of Mathematical and Management Sciences* 8: 329–359. Doi: http://dx.doi.org/10.1080/01966324.1988.10737244.

Mitra, D., F. Romeo, and A. Sangiovanni-Vincentelli. 1986. "Convergence and Finite-Time Behavior of Simulated Annealing." *Advances in Applied Probability* 18: 747–771.

Park, I. and H. Lee. 2004. "Design Effects for the Weighted Mean and Total Estimators under Complex Survey Sampling." *Survey Methodology* 30: 183–193.

R Core Team. 2013. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/ (accessed March 2014).

Raghunathan, T.E. and J.E. Grizzle. 1995. "A Split Questionnaire Survey Design." *Journal of the American Statistical Association* 90: 54–63.

Reis, F. 2013. "Links Between Centralisation of Data Collection and Survey Integration in the Context of the Industrialisation of Statistical Production." Working paper presented at the UNECE Seminar on Statistical Data Collection, 25–27 September, 2013, Geneva, Switzerland. Available at: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2013/mgt1/WP2.pdf (accessed January 2015).

Renssen, R.H. and N.J. Nieuwenbroek. 1997. "Aligning Estimates for Common Variables in Two or More Sample Surveys." *Journal of the American Statistical Association* 92: 368–374. Doi: http://dx.doi.org/10.1080/01621459.1997.10473635.

Smith, P. 2009. "Survey Harmonization in Official Household Surveys in the United Kingdom." In Proceedings of the ISI World Statistical Congresses, 16–22 August 2009, Durban, South Africa.

Tsallis, C. and D.A. Stariollo. 1996. "Generalized Simulated Annealing." *Physica A* 233: 395–406. Doi: http://dx.doi.org/10.1016/S0378-4371(96)00271-3.

UNSTATS 2005. *Household Sample Surveys in Developing and Transition Countries*. New York: United Nations (ST/ESA/STAT/SER.F/96.). Available at: http://unstats.un.org/unsd/hhsurveys/pdf/Household_surveys.pdf (accessed January 2015).

Received February 2015
Revised July 2015
Accepted October 2015