Statistics Netherlands

Questionnaire communication to collect financial data from large non-financial enterprises

Ger Snijkers (g.snijkers@cbs.nl) and colleagues[1]
Statistics Netherlands

1. Introduction: Redesign of the Quarterly Survey of Finances of Enterprises

In close collaboration with the Dutch Central Bank (DCB), Statistics Netherlands (SN) is redesigning the Quarterly Survey of Finances of Enterprises. Both institutes conduct surveys that collect financial data from the largest 350 non-financial enterprise groups in the Netherlands: SN conducts the Quarterly and Annual Survey of Finances of Enterprises, asking about chart of accounts data; DCB conducts a monthly survey asking about details on financial transactions, which are input for the chart of accounts. These surveys are used together in the Balance of Payments and National Accounts. In 2014, a new set of required data has been developed (the conceptual data model) as the basis for a new quarterly survey. This data model is very complex and requests for a lot of detailed financial information. As a consequence, the response process within businesses could be quite complicated and burdensome, even though the goal of this redesign was to reduce response burden, as well as improve the data quality. The target population consists of the 350 largest non-financial enterprises in the Netherlands.

As input for a new quarterly questionnaire a feasibility study was conducted, studying the internal response process in these businesses, in order to tailor the questionnaire to the response process. This study resulted in a number of design requirements.

This paper discusses the design process, including the background of the combined data model, the feasibility study, the business response processes, questionnaire design requirements, the planning of the questionnaire design, as well as the project team. In the presentation, I will show the new questionnaire, and discuss a number of questionnaire design issues and design solutions, applying a communication approach.

2. Goals of the redesign

In theory, the CBS and DCB surveys should generate approximately the same results, at the Balance of Payments level. In practice, however, large differences in results are perceived between the surveys. Several post-field actions have been taken to eliminate these differences (e.g. with regard to data editing). In order to permanently eliminate the differences, enhance quality and reduce response burden, both institutes concluded that the surveys should be combined into one new quarterly survey, from which the results can be used by both institutes.

3. Conceptual data model

As a consequence of this decision, in 2014 the requirements in the Balance of Payments 6 (BMP6) and the System of National Accounts 2008 (SNA 2008) have been translated into a conceptual data model. A project group of financial experts from SN and DCB developed this model.

This conceptual data model requests for a lot of detailed information to be provided each quarter of the year. It presented as a matrix of line items and columns as is shown in figure 1, defining the information in detail. It was anticipated that this new model would be a risk for the data collection, resulting in item non-response and measurement/unit errors (Haraldsen, 2013; Snijkers 2016). The risks included: are the data available on a quarterly basis, are the data easy to retrieve, and does this actually reduce response burden? In early 2015 it was decided that a risk assessment was needed: a feasibility study and an independent accounting expert review were carried out. As for the planning, the targeted deadline to field the new survey was set to the first quarter of 2017; it turns out that this will be early 2019.


Figure 1. Conceptual data model

4. Research questions

Instead of moving directly from the conceptual data model to the development of a data collection instrument, the feasibility of the data model was investigated (as described by Willimack, 2013; Snijkers & Willimack, 2011; Snijkers and Arentsen, 2015). This study was aimed at getting more insights in the response process within these large enterprises. The main research questions were: (1) Are the data available? And, (2) how much work is involved in collecting the data? In business surveys, these questions address the following issues:

1. What (data): What data do we get? Are the concepts clear and do they match or differ with accounting definitions?
2. Who (units): What entities in the enterprise are involved in the response process? Do we get the data about the correct units?
3. Where (people and sources): what business staffs is involved in the response process and where can the requested information be retrieved?
4. When (time): When is the requested information available?
5. How (questionnaire): What would be the best way to collect the required information?

These questions relate directly to measurement-related quality issues as discussed by Haraldsen (2013a; Bavdaz, 2010; Snijkers, 2016): see figure 2.

Figure 2. Measurement-related quality issues in business surveys: from concept to data
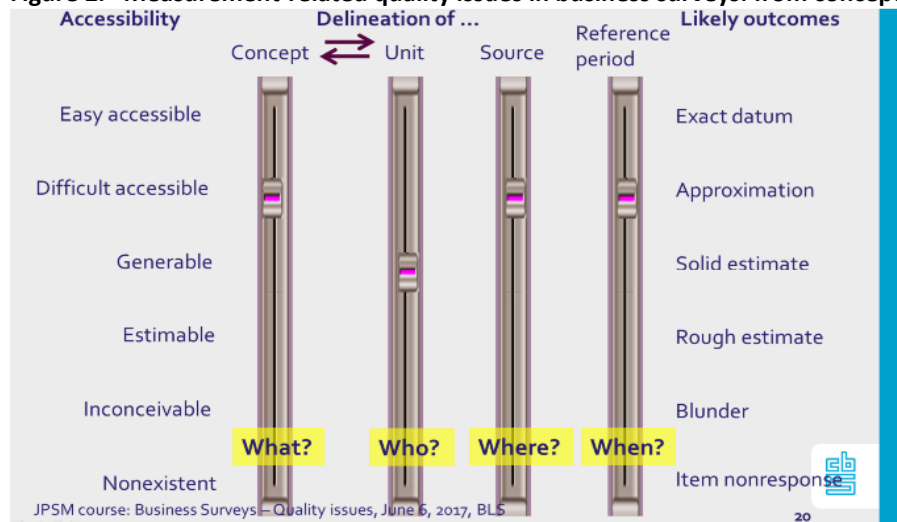
The last question addresses the operationalisation of a data collection instrument. This could be an electronic questionnaire, but also other options were possible. As for the design of a data collection instrument, a number of design issues were raised (see also Haraldsen, 2013b):

- Top-down or bottom-up design? In a data collection instrument two ways of completion can be applied: a top-down approach or a bottom-up approach. In a top-down approach, respondents first have to enter grand totals, followed by details which add up to the totals. The bottom-up approach works the other way round: first respondents have to enter the details; totals are calculated by adding up the details. The question is: what approach works best with regard to response burden and data quality (reduction of measurement errors).
- Do matrix (or grid) questions work for collecting these data, and how should it be designed on a screen?
- How to include instructions in a measurement instrument, in order to highlight the differences between accounting and statistical definitions to respondents?
- What mode of data collection would be the best option? Here a variety of modes can be included, varying from paper and electronic questionnaires, to uploading data using standardized record lay-out into the instrument, and data interchange methods.

Figure 3. Colour-coded Excel sheets

5. Feasibility Study and Accounting Expert Review

The issues listed above have been studied in a feasibility study. Five large non-financial enterprise groups were visited on site and asked about the information in the data model. A topic list for a 2-hour discussion on the data model was prepared on the basis of the research questions. These businesses have been visited in April/May 2015. The contact persons received a small gift as a token of appreciation. During the course of the study

Since the model under investigation is very large, the topic list was carefully time boxed to ensure adequate information was collected to proceed with the development of the data model and the data collection instrument. More specifically, the line items in the model were tackled one by one, whereas the required details corresponding to these line items were dealt with on a higher level. This approach does allow for some subjectivity when coding the results. It was not possible to do a detailed record check on availability and retrieval; this would take too much time.

At the same time, an accounting expert review (Willimack, 2013) was carried out by PricewaterhouseCoopers (PWC), with the same focus: to answer the five main questions mentioned above. The PWC study (PWC, 2015) consisted of a desk research, the accounting export review part, as well as a feasibility study with four different enterprises. This study was carried out independently of the SN/DNB feasibility study: during the execution phase, no information was shared between the researchers to avoid bias in outcomes.

Both studies resulted in the same conclusion: combining the SN and DCB surveys into one new survey seems logical for both organizations, but offers no synergy benefits to the enterprises. Consequently they don't

perceive the new data requirements as a reduction in response burden. As a result, the goals of improving data quality and reducing response burden may not be achieved.

The findings were presented as a short management summary and a color-coded Excel sheet (see figure 3). The coding is based upon the ease of retrieving the required information and the sources that would need to be used. For *each variable* in the data model the color-coded sheet indicates if:

- the information is easily and readily available (at group accounts level): coded green;
- the information is available at a central location, but not in the group accounts (treasury level), which requires more effort: coded yellow;
- the information is available, but decentralized(general ledger level), which requires considerable effort to acquire: coded orange;
- the information is not available: coded red;
- in the meantime, some information was dropped from the data model: this is shown in brown.

The color-coded Excel sheet presents management and researchers with a clear overview of potential risks in the data collection and processing stages: the more steps and the more sources are involved in the response process, and the deeper within the business information has to be retrieved, the higher the risks of survey errors like measurement errors and item non-response.

6. Business response process

Another important result of the feasibility study is that the business response process depends on the structure of the business administration. Figure 4 gives an overview of the processes that were identified. For businesses with a centralised accounting this process was straight forward. For complex structures, this retrieval process can be very complex, involving one or more respondents/data providers and data sources at various locations.

Figure 4. Complexity of response processes
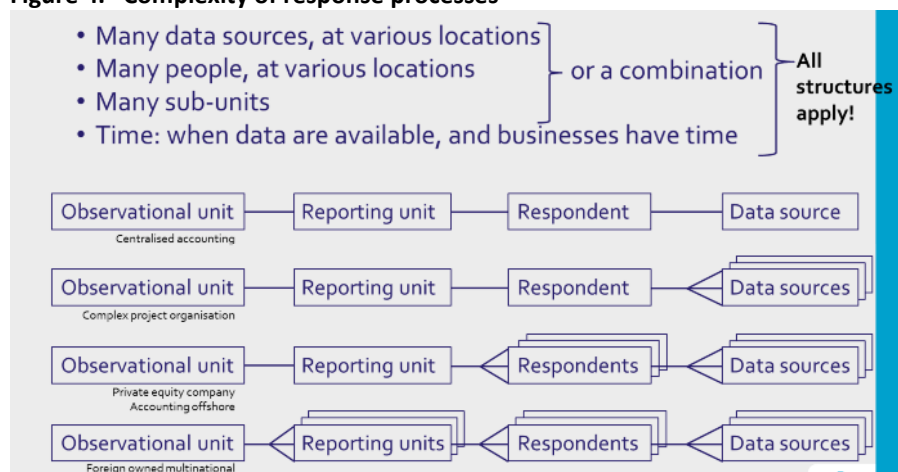
7. Questionnaire design requirements

Following from these studies it was decided to develop an electronic questionnaire; also a number of questionnaire design requirements were identified. These can be clustered into three groups: 1. Content issues, 2. User interface & usability issues, and 3. Recommendations regarding the communication strategy:

1. Content issues:
    - A clear definition of the structure of the questionnaire, identifying each and every data entry box.
    - Clear-cut definitions of terminology: there is a difference between statistical and accounting definitions
2. User interface & usability issues:
    - Both top-down and bottom-up completion should be possible.
    - Data entry not only manually but also by uploading/importing data files; as well as data export options.
    - Indicate where the data come from (to facilitate the internal process)
    - To be accessible and completed from various locations and by various respondents: online application.
    - Provide a clear overview of the questionnaire: use an index to provide overview of the content, for navigation, and progress control.
    - Have a print option of the questionnaire should be available, including an overview of the data asked.
    - Include consistency checks and validation rules

- The questionnaire should be available in Dutch and in English.
- Working with matrixes is not a problem.
3. Communication
  - Communicate the new questionnaire in one year in advance so that businesses can prepare the internal process. In follow-up to this result it was decided to include a pilot year (2018).

8. Questionnaire design

Early 2016 a project team started with the design of the questionnaire. Again this included two aspects: the operationalization of the content and the user interface. As for the content, the conceptual data model was translated into an overview of the entire questionnaire content using Excel (see figure 5a).This was done in such a way that: 1. each individual data item is represented by a data entry box (meaning that if a box was missing here, the variable would not be in the final data file), and 2. each tab represents a screen (thus serving as input for the user interface). This Excel sheet can be seen as a questionnaire schedule. Developing this schedule was a considerable effort and took quite some time. Specifying the questionnaire content in detail turned out to be a cyclical process: it required going back and forth to the conceptual data model, as the model needed additional specifications.

An initial visualisation of the user interface was designed in Powerpoint, as is shown in figure 5b. A draft version of this design was presented to a small number of businesses (those we had participated in the feasibility) as a first check to see if this user interface would work in practice. Especially we were interested in how people would navigate, finding their way, using the index. This seemed to work well, which gave confidence to proceed.

Even though the data model is quite complex, the structure of the questionnaire is quite simple. The question-naire is structured around the chart of accounts: assets and liabilities (which was the original SN Survey of Finances of Enterprises). For some items more detailed information is requested: a matrix of financial mutations (reconciliations); these items come from the monthly DCB survey.

In addition, a list of usability issues and functionalities was prepared. Each issue was rated according the MoSCoW principles: Must, Should, Could, Would. This was necessary as it was decided that the questionnaire would be developed by an external software developer.

Figure 5a. Translation of conceptual data model into an Excel questionnaire schedule

Figure 5b. Visualisation of the user interface using Powerpoint

9. Project team

It was decided that the electronic questionnaires should be developed by Revolux because of positive experiences with this software house, and the IDEP functionalities (see footnote 1). In a weekly cyclical process the questionnaire development progress was monitored by a project team. This team also provided the above discussed questionnaire requirements. Next to a project leader, this multi-disciplinary team existed of six people: an questionnaire designer from the Data Collection Division (who is owner of the questionnaire, and developed the Excel content sheet, Figure 5a), a content matter expert from the SN statistical department (who had been involved in the data model development, and has contacts with DCB experts), an expert from the large business unit (as most businesses in the target population are from this population), a former field

staff office (knowledgeable about the response processes in businesses), and a methodologist (with expertise in business questionnaire design, the business response process, and business survey communication).

10. Project planning
The questionnaire design and development started in mid-2016, and was planned in three stages (Göttgens, 2016):
1. Version 1 of the questionnaire: ready by March 2017. This version should include a working user interface, the entire chart of accounts level, as well as two detailed reconciliation modules.
- Pre-testing of version 1 was done in February-May 2017. This test was focussed on usability design (Giesen and Vis-Visschers, 2017). The results indicate the user interface seems to work quite well.
2. Version 2 of the questionnaire: ready by July 2017. In this version all modules should be ready, with a beginning of e.g. instructions and validation rules.
- Since in the feasibility study businesses indicated that they needed one year to get prepared, it was decided that 2018 will be a pilot year. In this year business will get access to the questionnaire, so they can establish the necessary processes. Also the internal processes within SN and DCB need to be prepared, including systems for the field work (e.g. a system test), data editing, and survey communication. In August 2016 summer a detailed communication strategy was designed (Houben and Snijkers, 2016), indicating that in November 2017 businesses will receive the first letter.
3. Version 3 will be the final version of the questionnaire, including all modules, instructions, error and validation checks, and a user manual. This version should be ready by the end of October 2017.
- The fieldwork will start in the first quarter of 2019.

11. Major lessons learned
- The common questionnaire development process taken from social surveys (conceptualisation – questionnaire development – pre-testing – adapting the questionnaire) should be adapted for business surveys and should start earlier (see also Snijkers and Willimack, 2011): conceptualisation – *study the business context* – questionnaire development – pre-testing – adapting the questionnaire.
  In business surveys it is especially important to take the business context into account, i.e. tailor to the business context (Snijkers et al., 2013), both for the questionnaire and the communication strategy.
- Operationalization of the data model into a questionnaire is not a linear process; a stable model is critical.
- It is critical to have a close collaboration with the software developers; it requires weekly update meeting to monitor the development progress.
- This project provides insights in the relationship between unit and measurement errors as shown in figures 2 and 4, indicating that unit errors are part of the Total Survey Error Framework (see Snijkers, 2016, Haraldsen, Snijkers and Zhang, 2015, Haraldsen 2013a;).

References
Bavdaz, M., 2010, The Multidimensional Integral Business Survey Response Model. Survey Methodology 36(1): 81-93.
Giesen, D., and R. Vis-Visschers, 2017, Design and results of usability tests with the new quarterly SN/DCB questionnaire, February-May 2017 (in Dutch). Report, Statistics Netherlands, Heerlen.
Göttgens, R., 2016, Project plan for developing the new SN/DCB questionnaire (in Dutch). Statistics Netherlands, Heerlen.
Haraldsen, G., 2013a, Quality Issues in Business Surveys. Chapter 3 in Snijkers, G., G. Haraldsen, J. Jones, and D.K. Willimack, Desiging and Conducting Business Surveys, Wiley, Hoboken.
Haraldsen, G., 2013b, Questionnaire Communication in Business Surveys. Chapter 8 in Snijkers, G., G. Haraldsen, J. Jones, and D.K. Willimack, Desiging and Conducting Business Surveys, Wiley, Hoboken.
Haraldsen, G., G. Snijkers, and L.-Ch. Zhang, 2015, A Total Survey Error Approach to Business Surveys. Presentation at the International Total Survey Error Conference (TSE15), 19-22 September, Baltimore, USA.
Houben, L., and G. Snijkers, 2016, Business Survey Communication Strategy for the new quarterly SN/DCB survey (in Dutch). Report, Statistics Netherlands, Heerlen.
PWC (PricewaterhouseCoopers), 2015, Results of the study on the Data Model for Non-financial Enterprises (in Dutch). Confidential PWC report, commissioned by Statistics Netherlands.
Snijkers, G., and K. Arentsen, 2015, Collecting Financial Data from Large Non-Financial Enterprises: a Feasibility Study. Paper presented at the 4th International Workshop on Business Data Collection Methodology, 14-16 September, Washington DC, USA.
Snijkers, G., G. Haraldsen, J., Jones, and D.K. Willimack, 2013, Designing and Conducting Business Surveys, Wiley, Hoboken.
Snijkers, G., and D.K. Willimack, 2011, The Missing Link: From Concepts to Questions in Economic Surveys. Paper presented at the 2nd European Establishment Statistics Workshop (EESW11), 12-14 September, Neuchâtel, Switzerland,
Snijkers, G., 2016, Achieving Quality in Organizational Surveys: A Holistic Approach. Chapter 3 in Liebig, S., and W. Matiaske (eds.), Methodische Probleme in der empirischen Organisationsforschung, Springer, Wiesbaden.
Willimack, D.K., 2013, Methods for the Development, Testing and Evaluation of Data Collection Instruments. Chapter 7 in Snijkers, G., G. Haraldsen, J. Jones, and D.K. Willimack, Desiging and Conducting Business Surveys, Wiley, Hoboken.