This module is part of the

Memobust Handbook

on Methodology of Modern Business Statistics

26 March 2014

Theme: Design of Data Collection Part 2: Contact Strategies

Contents

General section

1.   Summary

This modules deals with setting up a contact strategy to reach respondents. It deals with several sub-themes; the first part deals with choosing a contact strategy regarding when to send material to the respondents and what material to send, strategies for reminders, using penalties and fines et cetera. The second sub-part deals with finding the right respondent at an enterprise. These first two parts cover the whole of the sample, while following sub-parts deal with specific treatment of sub-populations. The third sub-part covers whether or not to treat new enterprises in any specific way, and the fourth sub-part covers specific treatment of large enterprises. Since giving feedback to respondents is a way of encouraging participation, a specific sub-part is devoted to that area. And finally, the last sub-part deals with making changes to the design during the data collection based on the outcomes so far, i.e., responsive design. Together with the previous module ("Data Collection – Design of Data Collection Part 1: Choosing the Appropriate Data Collection Method"), the module gives an overview of the decisions to make before starting the actual collection process.

2.   General description

2.1   *Choosing the appropriate contact strategy*

When the appropriate data collection mode(s) has been chosen, the next step is to choose an appropriate contact strategy. A contact strategy consists of when and how respondents are contacted, and what material (questionnaire, cover letter, instructions et cetera) is used in each contact. As was stated above (see "Data Collection – Design of Data Collection Part 1: Choosing the Appropriate Data Collection Method" regarding mixing modes), it is in many cases wise to exploit a data collection mode to the maximum before switching to another mode, i.e., using a sequential mixed mode approach. Another possibility when mixing modes is to utilise a parallel mixed mode design, giving the respondent the choice to choose his or her preferred mode. Some studies have shown that the choice of mode in itself can have a negative effect on response rates. This has been shown for voluntary surveys, where the most important question for the respondent at the initial stage is whether to participate or not in the survey. Making this choice more difficult by offering several possible ways to provide data may affect the willingness to participate in a negative way (not participating may seem as the easiest choice). For mandatory surveys, such negative effects are not as common. Since the respondent has to provide the data, being able to choose the best way to do it may have a positive effect on providing timely high quality data.

The choices to make regarding contact strategy are the following:

1. Timing of the first sending

2. Time given to respond

3. Material to include in the first sending

4. Number of reminders

5. Timing of reminders

6. Material to include in reminders

7. Penalties/Fines

Timing of first sending and time given to respond

Of course, the timing of the first sending is dependent on when the data collection needs to be finished for other processes to have appropriate time before the survey needs to be completed, and also on some of the other choices (number of reminders, using force), but it is also in many cases dependent on data availability. Many surveys (for example most monthly economic surveys) ask for very recent data, as soon as they are available. For example, economic data on revenues and costs, salary data, employment data et cetera are normally registered in the administrative systems at enterprises. But in many cases, there is a need to extract data in the form of reports to be able to provide them to statistical offices. For accounting data, there may be additional work done in order to finish an annual report or such, and there may be internal and external rules on when data may be released and to whom. External rules may even be legislated. The general advice is to consider what data the survey covers, examine when data may be available for statistical purposes and adapt the data collection period and the timing of the first sending after that. Normally, the best time for the first sending is just when data has become available. For an annual survey, it is normally possible to adjust the collection period to when data are available. However, for some short term statistics with a tight production schedule, this might interfere with the second point, time given to respond. It might then be better to send out the request a short time before data are available, so the respondents can prepare for providing data and are ready when data become available. It is also important not to send out the questionnaire too late. Data are in many cases archived at enterprises at specific times, so it is important not to send out questionnaires asking for old data and most importantly not to ask for data that have recently been archived, since that increases the risk of respondents misreporting current data instead. For example, accounting data are often archived one year after the accounting year, so that what are readily available at a specific time are data for the current and the previous accounting year, while older data would have to be retrieved from archives.

When asking for old data, it might be necessary to adapt the given time to respond to the survey, but this issue is important to consider in all surveys. The time given to respond should not be too short, since the respondent must be given a fair chance to fulfil the requests of the statistical office, especially if the survey is mandatory. But the time given should neither be too long, since that increases the risk of respondents forgetting about sending data, and also because it sends a signal to the respondents that the survey is not very important. Normally 10-14 days response time is appropriate for short term statistics, while a response time of 3-4 weeks, but not much longer than that, is appropriate for annual or one-time surveys.

First sending

The decision on what material to include in the first sending is based on the choice of modes to use:

Paper questionnaire: Given that the choice is a mail-out paper questionnaire, the material that is sent is the questionnaire itself and a cover letter describing the purpose of the survey, the due date et cetera. In most cases, a pre-paid return envelope with the sending address pre-printed is included so that no additional cost except the response burden is put on the respondent. If a parallel mixed mode approach

is chosen using paper as the main mode, the cover letter or the questionnaire contains a link to reply to the survey by web.

Web survey: For a pure web survey, what is sent is an e-mail with the description of the survey and everything needed, including a link to the survey and, if appropriate log-in information such as user-ID and password. However, a survey that is run by a National Statistical Institute and uses web as the primary mode can also choose another way – to use a mail-out sending for a survey with web as the response alternative. In that case, the sending consists of only a cover letter, including a web address to enter into the respondent's browser and, if appropriate, user-ID and password to log in. This method makes it easier to reach all respondents, since registers of postal addresses are normally more complete and more up-to-date than registers of e-mail addresses. This method has proven effective in many countries, and is recommended for all surveys where you don't have regular and close contact with all respondents (monthly surveys where enterprises selected remain in the sample over a longer period of time being the exception).

CATI and CAPI: For an interviewer-administered survey, the initial sending is a pre-notice letter that tells the respondent that the survey has started and that he or she will be contacted by an interviewer. In an interview survey, no additional material to this letter is needed.

Regarding pre-contacts with respondents in other surveys than interviewed-administered ones, it is sometimes done for new respondents to ongoing surveys. This is described further in Section 2.3 below.

Reminders

Enterprises that haven't sent their data by the due date need to be reminded. There are several methods to use for reminders:

- Ordinary mail

- E-mail

- Telephone

In CATI/CAPI surveys there are no reminders as such, rather a number of contact attempts to reach the respondents.

The number of reminders to use is a delicate matter, especially in voluntary surveys. On one hand, a voluntary survey that is not given a high priority by an enterprise may need many reminders to have a good enough response rate. On the other hand, sending too many reminders angers respondents, who may feel that not responding is the way to show that they will not participate in the survey. Very few respondents actually contact the statistical office to notify the survey that they are not willing to participate. Two or three reminders is the common practice in most countries. The last reminder may also be targeted at specific groups or strata where the response rate is low.

The timing of reminders is also to be decided. If the survey has a due date, which is common practice in enterprise surveys, the first reminder should be sent right after the due date. Waiting a long time after the due date to send a reminder may be interpreted by respondents that the survey was not that important, that the need for data is not that big. This may in turn have a negative effect on response rates. In short-term surveys, especially monthly surveys, there might actually be such a tight production schedule that the first reminder must actually be sent before the due date. This is normally

done by sending a "thank you and reminder" card or letter, thanking those respondents that have already sent data while at the same time reminding the rest that the due date is approaching.

A final decision regarding reminders is what material to include in reminders, this is of course decided by the chosen mode. The decision between ordinary mail and e-mail is the same as when considering the first sending. Of course, if the first sending was done by e-mail, reminders should normally also be sent by e-mail, unless a reminder includes a paper questionnaire, in which case ordinary mail is necessary. Including a paper questionnaire as an attached file in an e-mail is also an option, but it may be difficult to manage the practicalities if the questionnaire includes individualised information like pre-printed data or such. If reminders are carried out by phone, no material is included. The contact strategy may also involve a mixing of modes at the reminder stage as well, using for example mail or e-mail for some enterprises and phone for others.

Using penalties or fines

When the previously described ways of managing reminders do not work, country legislation in many countries gives Statistical Offices the possibility to use the force of the law to ensure that responses to mandatory surveys are received. In this case, the forms and methods available differ between countries, but may involve fines or other penalties for non-respondents. Therefore, it is difficult to make a general description of how this process should be run. However, it can be said that if it is up to the Statistical Office to use the force of the law or not, it is important to actually use it, to show non-respondents the importance of complying with mandatory requirements. If, for administrative reasons, cost reasons or other reasons it is not possible to use this force against all non-respondents, it is most important to use it against new respondents to ensure their future cooperation, and also to the enterprises that are most important to the quality of the survey (i.e., the largest enterprises).

2.2   Defining the respondent

One very important aspect in getting high quality data from enterprises is to be able to reach the best respondent within the enterprise. In small enterprises, this is normally not a big problem, since there are only a few employees, so the chance of reaching a relevant respondent is high. On the other hand, small enterprises often have their accounting done by a service bureau or similar, meaning the people within the enterprise might not always be able to respond to a survey themselves, without consulting the service bureau. But for small enterprises, the best way to get good responses is to address the survey to the enterprise itself. For larger enterprises, it might be more efficient to address a questionnaire to either a specific person or a designated role. The advantage of doing so is minimising the risk that the questionnaire is either held up by a "gatekeeper" within the enterprise – for example a person who opens all mail, but does not know whom to give a specific questionnaire to – or given to the wrong person within the enterprise. If the questionnaire is filled in by the "wrong" person – i.e., someone who is not suitable to respond based on the contents of the questionnaire – the risk of having measurement errors due to misunderstood questions et cetera is increased. Another risk in large enterprises is that if a questionnaire is first given to the wrong person and that person realises he or she is not the suitable person, it takes extra time for that person to find a suitable person in turn. This risks late responses or non-response to the questionnaire floating around within the enterprise. However, addressing a questionnaire to a specific person or a designated role may also have disadvantages. Starting with a specific person, this is normally the person who answered the questionnaire last time. This option is only available in ongoing surveys. If the last time was a year, or sometimes even three

months ago, the person may no longer be the most suitable person. He or she may have changed position (in which case he or she might be suitable in finding the new suitable person, but may not have the time to do so based on the new commitments) or may even have left the company. In the last case, this may lead to the enterprise sending back the envelope without even opening it, meaning extra effort to re-send the questionnaire to someone else. Sending to a specific person may therefore only be recommended in short-term surveys where the statistical office is in continuous contact with most respondents in the survey, or in surveys where the advantage of having the same contact over time is very large, e.g., when the content is extremely complicated. Sending to a designated role (e.g., "head of economic division") may be an alternative to simplify the work of a gatekeeper within an enterprise to find a suitable person. However, this may also be risky. It is difficult to find role names that are known to all enterprises. If the role name chosen does not exist within an enterprise, the enterprise may simply send the envelope back. Otherwise, they will often have to translate the given role to something similar within their enterprise, the advantages of addressing questionnaires to a specific role can therefore be considered small. If a questionnaire is suitable for a specific type of person, it is probably better to state clearly in the cover letter what kind of data the questionnaire asks for and what type of person is a suitable respondent.

If the survey is large and covers several subject matters, it might not be possible for one person to fill in the whole questionnaire, he or she might need help or input from other people within the enterprise. This is more often the case in large enterprises than in small ones. If this is the case, it is good if the questionnaire is designed so that dividing the task is easier. This can be done by specifically putting parts that might need additional respondents on separate pages, and by allowing respondents to electronic questionnaire to save the data without sending and resume the questionnaire at a later time.

2.3 *Specific treatment of new enterprises*

Especially in an ongoing survey where the design normally is such that a sampled unit will be included in the sample over a period of time before rotating out of the sample, it might be worth to consider treating new enterprises in the sample a little differently. This can be considered not only in the data collection process, but also in the processing stage where perhaps they could be edited in another way than other respondents. But at the collection stage, if the questionnaires are sent to a specific person (see Section 2.2), no such person is available for new enterprises. One alternative in such a case is to have a pre-contact with new enterprises in the survey. The pre-contact can be done by phone or by sending new enterprises a specific letter before the collection starts to ask for a suitable person to send the actual questionnaire to. If the questionnaire contains specific data that have to be extracted from accounting systems (or similar) or requires specific actions from the respondents (calculations, estimations, perhaps even changing the accounting system to keep track of the data requested), a pre-notice letter or pre-collection telephone contact can also include such information and instructions. However, if the last example is valid, i.e., the enterprise has to make adjustments to the accounting system or in other ways arrange for data to be stored to make filling in the questionnaire possible, such a pre-contact must be done well in advance, so that the enterprise has a reasonable possibility to take action. Since sampling is often done close to the collection sending pre-notice letters or phoning so long in advance is not possible. So in practice, pre-contacts should mostly be used to find contact persons and inform about an upcoming survey. Since additional letters also increase the response burden on enterprises (it takes time to read a letter, react to it and take action) it

is recommended not to use them for easy surveys, they can only be motivated where data provision can be considered difficult or require special knowledge.

2.4   Specific treatment of large enterprises

Since large enterprises are more important than small enterprises in most business surveys with regards to quality, and since the largest enterprises are sampled in most surveys, many countries have decided to devote special attention to large enterprises. Some of the issues that this treatment includes are the following:

- Providing a single point of contact within the Statistical Office for the largest enterprises. This is both for when the enterprises need contact with the Statistical Office and vice versa.

- Building relations with the respondents within the largest enterprises.

- Profiling the largest enterprises with regards to statistical units – enterprise group, legal units, kind-of-activity units, local kind-of-activity units and so on.

- Helping surveys finding the right contact person within the largest enterprises for a specific survey. This can be aided by identifying functions and roles used in each of the largest enterprises. In some countries, these roles and functions are surveyed specifically and registered for future use.

- Helping the collection and editing staff with contacts with these enterprises.

- Helping the enterprises with support for their data provision.

- Enterprise-specific arrangements in data provision, such as specific questionnaires containing data from several surveys, aid in how to estimate difficult figures and such.

- Coherence analysis of data sent to different surveys.

In most countries that have tried a specific treatment of large enterprises, a special organisational group of people has been created. Within the group, one person deals with a specific number of enterprises, normally the same enterprises over time. This helps to build competence about the enterprise and its data, and it builds a relation with the enterprise. The enterprises themselves appreciate having a single contact person et cetera. Normally, one person deals with between 3 and 10 enterprises, and the group normally consists of 5-10 persons. The number of enterprises that should be covered by such a treatment may differ between countries, but in most countries it would be true to say that covering the 30-50 largest enterprises will cover a large part of the economy.

2.5   Feedback to enterprises

One final thing to consider when designing the data collection is whether to give respondents feedback on their responses or not. Here, we are not talking about a simple receipt that data have been received, which is normal procedure in electronic collection, but rather something more, using the actual data. Feedback can be of several types:

- General feedback, e.g., finalised figures and tables from the survey

- Specific feedback, using the actual data provided by each enterprise to provide an individualised feedback, maybe even comparing the enterprise to similar ones. Examples of such figures are key ratios or market shares.

Giving feedback can have several positive effects – it can show how the information is used by society and what is produced from the respondents' data, it can be, in some cases, of direct use to the respondents themselves and it can lower the perceived burden of providing the data. So it can be recommended to try to give feedback whenever possible. But feedback can be designed in several different ways.

In general, specific feedback is probably more efficient than general feedback, since it is based on the enterprises' own figures and could be of more use to the enterprise itself. The most attractive form of feedback is probably feedback directly after transmitting data. This requires that the transmitted data are immediately transformed into something nice-looking and sent back to the respondent. If this is possible, it can have very positive effect. But it is often not possible to use current data from other enterprises to make comparisons in such a feedback. Then, in this case, a good alternative could be an interesting feedback created by combining current data from the specific respondent and data from previous rounds for the whole sample, such a feedback can be used. But such feedback would have to be judged against better individualised feedback at a later stage. Therefore, if it is possible to create a better feedback by using current data but it is not possible to create that feedback immediately, then it is better to send the feedback at a later stage, after the collection is completed. When such feedback is sent later, it is good to allow respondents to choose (for example by ticking a box in the questionnaire) whether they want to receive the feedback or not, and also to give them the alternative to direct the feedback to a different person than the respondent.

If specific feedback based on data from the individual respondent is not relevant for a specific survey, or if it is very difficult to create, more general feedback based on the whole sample can be considered an alternative, it is still better than nothing. Another version of this is to give feedback of the data from last survey round when the request for data for the current round is sent out. This can result in a higher response rate or a lower perceived burden in itself.

2.6   *Responsive design issues*

Responsive design, which is a special case of adaptive design (Schouten et al., 2013), is a term that is used for modifying the design while the collection process is running. Of course, this is not done in a random manner, but rather following pre-specified rules. When statisticians talk about responsive design, they mostly refer to household surveys using CATI or CAPI modes. In those surveys, a lot of adaptations to the contact strategy can be made based on response rates in different strata, age groups et cetera, and those changes can have a profound effect on the actual collection work, since it deals with when different cases should be called and how many contact attempts should be made. For business statistics, responsive design is not discussed as much, partly because most business surveys use self-administered modes and also since, when CATI/CAPI is used, contacts with enterprises are limited to business hours. However there are still some adaptations to the design that can be made during the collection phase in business surveys:

- Utilising different modes – for example, in strata where the response rate is low, it might be considered to use a different, more expensive mode, than for other strata. This can be using CATI instead of paper/web questionnaires, CAPI instead of CATI et cetera.

- Utilising different reminder strategies – for example deciding which enterprises should get an e-mail reminder, which should get a reminder by paper mail and which should get a reminder by telephone.

- Utilising penalties/fines. In most cases, penalties/fines are not used for all non-responding enterprises, a responsive design approach can be used to decide for which enterprises penalties/fines will be used.

- Utilising different versions of the questionnaire. Some business surveys use different versions of the questionnaire, for example a longer questionnaire for larger enterprises and a shorter version for smaller enterprises. Responsive design can be used to decide if some enterprises should receive the shorter version instead of the longer one at the reminder stage.

In order to use a responsive design approach, the two following things must be done:

1. Set up decision rules for action. This involves several steps;

   a. deciding on which different paths that can be taken (different modes, reminders, questionnaires, penalties et cetera, see above),

   b. deciding which sub-groups to analyse (different strata or other groups)

   c. deciding the thresholds when action is to be taken

   d. deciding when decisions will be taken

   The four points a-d will then resign in a set of specific rules, e.g., "At date X, all enterprises with a size over Y in all strata with a response rate of Z or lower will be given a telephone reminder instead of the planned e-mail reminder". Of course, parameters like Y and Z can also be relative, compared to other enterprises or groups. For example, Z could be something like a response rate that is a percent lower than the average response rate in all strata. There has been some work done to come up with more advanced indicators that take response bias and the like into account, for example at Statistics Canada and Statistics Netherlands. Those indicators have mostly been designed for household surveys but there is some experience in business statistics as well. At the time of writing, it is too early to make a general recommendation on more advanced indicators.

2. Set up measurements. In order to utilise responsive design, there must be paradata available that can be used to make the decisions based on the rules that were set up.

3. Measure and take action. As the collection moves along and the decided measurement points are reached, measurement is made and appropriate action taken. This may also involve adjusting parameters like Y or Z in the example above, especially if it turns out that the action required is too costly compared to the survey budget.

It should be noted that responsive design is a relatively new concept that is still being developed within many statistical offices, and it is not yet possible to make general recommendations as to when and how to use it or not.

3.   Design issues

4.   Available software tools

5.   Decision tree of methods

6.   Glossary

For definitions of terms used in this module, please refer to the separate "Glossary" provided as part of the handbook.

7.   References

De Leeuw, E. D., Hox, J. J., and Dillman, D. A. (eds.) (2008), *International Handbook of Survey Methodology*. Lawrence Erlbaum Associates, New York.

Schouten, B., Calinescu, M., and Luiten, A.(2013), Optimizing quality of response through adaptive survey designs. Component of Statistics Canada Catalogue no. 12-001-X Business Survey Methods Division – June 2013.

Snijkers, G., Haraldsen, G., Jones, J., and Willimack, D. K. (2013), *Designing and Conducting Business Surveys*. Wiley, Hoboken, NJ.

Interconnections with other modules

8.   **Related themes described in other modules**

    1. Data Collection – Design of Data Collection Part 1: Choosing the Appropriate Data Collection Method

9.   **Methods explicitly referred to in this module**

    1.

10.   **Mathematical techniques explicitly referred to in this module**

    1.

11.   **GSBPM phases explicitly referred to in this module**

    1.

12.   **Tools explicitly referred to in this module**

    1.

13.   **Process steps explicitly referred to in this module**

    1.

Administrative section

14.   Module code

Data Collection-T-Design of Data Collection (Part 2)

15.   Version history

Version	Date	Description of changes	Author	Institute
0.1	17-02-2012	first draft	Johan Erikson	Statistics Sweden
0.2	19-06-2012	second draft after first review	Johan Erikson	Statistics Sweden
0.3	27-06-2012	revised after second review	Johan Erikson	Statistics Sweden
0.4	30-10-2013	revised after EB comments	Johan Erikson	Statistics Sweden
0.4.1	21-11-2013	preliminary release		
1.0	26-03-2014	final version within the Memobust project		

16.   Template version and print date

Template version used	1.0 p 4 d.d. 22-11-2012
Print date	21-3-2014 17:49