Enhancing Data for Complex Agricultural Establishments
June 26-28
Niagara-on-the-Lake, Ontario, Canada

Session:
Current and Imminent Data Collection Challenges Faced by Statistical Agencies

Jeffrey Smith, Director, Agriculture Division
Statistics Canada

Introduction

This paper formed the basis for the remarks presented by Statistics Canada in the session "Current and Imminent Data Collection Challenges Faced by Statistical Agencies", during which representatives from several agencies[1] listed and described challenges they faced (and/or expected to face in the near future) in collecting data from large and/or complex agricultural operations. In this paper, the challenges facing Statistics Canada in this regard are presented in two parts: those challenges faced generally when collecting data in business surveys; and those faced more specifically in collecting data from agricultural operations (surveys of agricultural operations are classified as business surveys at Statistics Canada).

General issues

Response burden

Generally, governments are sensitive to the amount of reporting burden placed on businesses, especially smaller ones. They are also generally responsive to calls for reduction in this burden. The outcomes of this sensitivity and responsiveness are usually initiatives that bear names such as "paper burden reduction" or "red tape reduction" and may include goals such as an x% reduction in the forms businesses have to file. Government statistical agencies are similarly sensitive and responsive to issues surrounding response burden since, even in cases where response is mandatory (e.g., specified in legislation), these agencies prefer to gather data co-operatively, rather than by resorting to enforcement of legal sanctions.

The Canadian Federal government has launched several of these exercises in recent years. For example, in its Budget 2007, the Federal government committed to a *"... goal of reducing by 20 percent the burden of paperwork required of Canadian small businesses."*, and announced in 2009 that this had been met by the thirteen key departments and agencies: *" … by streamlining regulations, eliminating duplicate requirements, getting rid of overlapping obligations, and reducing how often documents need to be filed."* An inventory of *"… administrative requirements and information obligations with which businesses must comply"* was compiled by September 2007 including information on the burden they imposed such as counts of forms and other measures of compliance burden. Recently, another round of reductions has been announced. On January 13,

[1] The five agencies presenting in this session were: European Commission – Directorate-General Agriculture and Rural Development (DG.AGRI), European Commission – Eurostat, Instituto Brasileiro de Geografia e Estatistica (IBGE), National Agricultural Statistics Service (NASS) and Statistics Canada.

2011 the Prime Minister announced the creation of the Red Tape Reduction Commission, which would hold public consultations, including online forums, and report in the fall of 2011.

Although temporarily "on hold" during the last federal election campaign in the spring of 2011, the Commission has been very active and has provided federal departments, including Statistics Canada, with information it has collected, during in-person and on-line forums, concerning red tape issues reported by businesses. This information is also available on the Commission's website (www.reduceredtape.gc.ca). As expected, a number of Federal departments were specifically cited as problematic by people offering input to the Commission, Statistics Canada included. The volume of "complaints" aimed at Statistics Canada was not high given the very large number of survey contacts per year; nonetheless, this exercise is being taken very seriously. The general themes of the comments specific to Statistics Canada were:

1. Business surveys are seen as burdensome (some more than others);
2. Some questions asked are felt to duplicate information asked by other federal departments;
3. The purpose, relevance or use of the data collected is not always apparent to respondents;
4. Collection practices are not always appreciated;
5. Businesses feel they should be compensated for the time it takes to respond to surveys.

While reducing response burden on business is a desirable goal for both respondents and statistical agencies (the latter since happier respondents should be more cooperative), high profile efforts to reduce it can pose challenges for collectors of data. In the case of Statistics Canada, efforts have been in place for a long time to ensure that burden is minimized and therefore a highly publicized campaign to reduce by, say 20%, at the government-wide level may be difficult to achieve in a particular agency, such as Statistics Canada, once all (or most) non-essential burden has already been removed. In addition, for some respondents, the presence of highly publicized initiatives to eliminate burden may be used as a reason or justification for not participating in surveys, even those that are mandatory.

Statistics Canada's efforts to control response burden are numerous, including using administrative data where possible to replace direct surveying of businesses, ensuring forms ask only essential content, making extensive use of sampling, ensuring sample sizes are minimized given the statistical quality required, and so on. Statistics Canada's success in managing response burden is reported in its Departmental Performance Report (DPR), which is prepared and submitted to Parliament annually. Response burden hours per Canadian business establishment are expressed as an index relative to 1991 hours, and the latest published DPR (for the fiscal year ending 2009-10) showed a value of 68[2]. That is, on average, for each hour of time spent responding to Statistics Canada surveys in 1991, a Canadian business was spending just 40.8 minutes in 2010.

Interaction or confusion with other collection activities

Given that respondents do not always clearly distinguish among data collection activities carried out by various parties (different levels of government, private entities, etc), a declaration by one, such as making their survey voluntary may spill over onto other surveys. The many discussions surrounding the 2010 decision to replace the mandatory "long form" Canadian Census of Population (previously sent to 20% of households) with the voluntary National Household Survey

[2] Agriculture *surveys* are included in the overall response burden numbers described above, but the Census of Agriculture is not.

(sent to about one-third of households), and the numerous statements in the media and elsewhere that even mandatory surveys should have sanctions for non-response removed (as for example, are provided for in the *Statistics Act*[3]), may lead respondents to conclude that all government surveys are optional. This could have some impact on response rates and the amount of follow-up effort (and cost) needed to obtain data of acceptable quality. At the time of writing, this type of confusion was not evident to any great extent.

Another challenge related to penalties is that although these penalties are clearly stated in the *Statistics Act*, the approach has always been to seek cooperation from respondents and to only exercise them via prosecution in the most extreme cases of wilful non-response. However, even this limited, highly selected use may not be effective given that in some cases where convictions have been obtained, penalties have been waived by the courts. The challenge then is to preserve the goodwill and cooperation in an environment which seemingly conveys the message to respondents that surveys may be ignored without much consequence.

Respondents may also confuse data collection activities from different organizations, especially if the topics or content have some similarity. The challenge for the statistical agency is to be aware of other collections that may be taking place, to collaborate where this is possible (e.g., joint collection, data sharing agreements), or at the very least, be able to distinguish and justify its collection to highlight the uniqueness and importance to respondents.

Mode of collection, collection practices

Modes of data collection for business surveys have evolved over the years. Different modes, such as mail, telephone and various forms of electronic reporting can prompt different reactions and perceptions on the part of respondents. Mail can offer the respondent great flexibility as to when they complete and return the questionnaire (and so may be perceived as less intrusive and less burdensome than telephone surveys), but this can be negative from the point of view of the data collector faced with a short collection window. Statistics Canada currently focuses on telephone collection for most business surveys, although there is considerable work underway to develop "e-questionnaires" as the primary mode, as well as to create a true multi-mode environment where respondents have a choice of how to receive and how to return their questionnaire.

When some form of interviewing is used to collect data (whether in person or by telephone), and especially where large numbers of interviewers are employed, the need for standardized training and procedures is clear to ensure the soundness of the collected data and minimize interviewer effects. Couple this with the need to often collect data in short time frames, and the result is that interviewers may come across as rigid (as when giving standard replies to respondent questions, for example) and even seem rude, aggressive and threatening (as when responding to questions about penalties for refusal on a mandatory survey). Clearly some interviewers are extremely effective,

---

[3] Section 31.

Every person who, without lawful excuse,

(a) refuses or neglects to answer, or wilfully answers falsely, any question requisite for obtaining any information sought in respect of the objects of this Act or pertinent thereto that has been asked of him by any person employed or deemed to be employed under this Act, or

(b) refuses or neglects to furnish any information or to fill in to the best of his knowledge and belief any schedule or form that the person has been required to fill in, and to return the same when and as required of him pursuant to this Act, or knowingly gives false or misleading information or practises any other deception thereunder

is, for every refusal or neglect, or false answer or deception, guilty of an offence and liable on summary conviction to a fine not exceeding five hundred dollars or to imprisonment for a term not exceeding three months or to both. 1970-71-72, c. 15, s. 29.

but in large calling operations not all will be, and it is a challenge to manage the positive dynamic between respondent and interviewer.

Technology can offer some further reductions in actual or perceived response burden. A well-designed electronic questionnaire (internet response channel) or telephone interview (computer-assisted telephone interview) has the ability to quickly take the respondent through a long questionnaire asking only the relevant portions. On the other hand, although a printed questionnaire can also contain skip patterns, there may be a perception of burden in receiving a long, multi-page printed form. The perception of burden may be fuelled by seeing or hearing the number of pages or number of questions quoted publicly (such as in the media) without all of the context, especially when (as is usually the case), the fact that a given respondent will likely need to respond to only a portion of the questions is not reported.

Technology may also work against us as in the case of the rise of methods to screen and avoid telephone contact (caller id, answering machines, do-not-call lists[4], cell-phones, etc.). The cell-phone issue is complex as not only do more people use them, more rely on them exclusively (no land line). Typically, businesses will want customers to be aware of them and able to reach them, and therefore probably will have a land line even if they do use cell-phones extensively. The lack of cell-phone directories means it will be especially difficult to reach such cell-only respondents. Thus, telephoning respondents, a once very productive collection method, may be in the twilight of its usefulness and the time may be ripe for a next generation of collection mode.

Collection costs / budget pressures

Costs for collection operations generally rise over time and these costs are managed in a variety of ways such as continually refining sample designs to allow sample size reductions, ensuring all content asked continues to be relevant (and eliminating content that is no longer relevant), working with collection partners to ensure we are using the infrastructure optimally (how many and which offices to use), using technology and other means to ensure efforts are producing maximum benefit (using best time to call, appointments, call schedulers, etc). Since there is typically a point of diminishing returns (or a point at which quality would suffer too much), it continues to be a challenge to squeeze more and more savings out of the collection budget using only these types of approaches.

Another aspect of cost is the opportunity cost that respondents seem to feel. In an age of increasing information needs and demands, businesses feel more challenged to meet the many demands on their time while running their business, let alone respond to government enquiries, especially if they see no immediate benefit. Businesses generally exist to earn returns and so being asked to donate their time to what they may see as an unproductive activity is understandably somewhat unnatural to them. The challenge then is to minimize the time intrusion, convince them of the value to their business of the time spent on responding to a survey, and to find some effective, legal and cost-efficient way to compensate respondents. For example, in the absence of the ability to make direct monetary payments, the offer of other benefits to participants such as access to special websites containing information of value to their business, could be explored.

[4] Statistics Canada is exempt from the national do-not-call list; nevertheless, respondents may not make this distinction and may think joining such a list will prevent all incoming telephone calls soliciting information, including from StatCan. They thus may take exception when we do call.

In some cases business respondents do see the value of the information produced from surveys but fail to grasp the need for statistical rigour in the conduct of surveys. Related to the "time is money" argument above, respondents may also feel that their "information is money" and that they should not have to provide it, rather – "go ask one of the many other businesses" (a sort of the variant of the not-in-my-backyard feeling) or "surveys are fine as long as you bother my competitors, but not me". In other words, businesses may feel that being burdened with surveys hampers their ability to compete by diverting time from their productive work (they say) to time spent in wasteful (they say) "red tape." There may also be a feeling that if not every business is asked to respond, then the statistical agency is therefore providing their competitors with a competitive advantage (by not making demands on the competitors' time, and by publishing their information and not the competitors' to boot – the latter also indicative of not fully understanding the statistical survey design). The challenge then is to effectively explain to respondents the need for statistically sound and representative samples and then actually design surveys that spread the burden as evenly as possible, so no one business or group of businesses feels like they are being targeted. This is especially challenging for sectors with populations of businesses that are such that sample designs with high sampling rate strata (or even a take-all stratum) are the most efficient, since in that case some businesses will indeed be surveyed more than others.

Dealing with large or complex businesses

Collecting data from these businesses can be a challenge since, by virtue of their size and/or diversity of activity, they may have units or subunits engaged in many activities and industries, and therefore may be included in many surveys and other information gathering requests. The volume of requests, the timing of them and the possibility that some of the same information may be requested on different occasions (such as basic identification and operating information) puts additional burden on these important units and may lead to frustration and possibly less cooperation in reporting. The challenges and how Statistics Canada deals with these larger and more complex businesses is covered in detail in the paper *Canada's Enterprise Portfolio Management and Large Agricultural Operations Statistics (LAOS) Unit Approach* presented in the session "Current Best Practices from Statistical Agencies", and therefore are not presented here.

Agriculture specific issues

Response burden

The universe of farms is shrinking, so, even in sample surveys, a farm may be randomly selected more often today than in past years. Agricultural producers will likely notice that "their turn" seems to be coming up more often and this may impair response rates and/or increase effort needed to elicit a response. It is a challenge to manage this effect and one approach is to grant, on a case-by-case basis, temporary relief from responding after a period of frequent selection where the respondent has been co-operative in providing information and where such a temporary exemption does not impair the statistical representativeness of the sample.

Some agriculture surveys (in particular where the population is small) are censuses or near censuses, such as when a large take-all stratum is part of the design – for these respondents, the burden may seem to never end. For those operations growing in importance rather quickly and therefore transitioning from a lightly sampled stratum to a take-all, the increased frequency of contact may also be quite noticeable. In some cases, strata have relatively high sampling rates so

that the chance of *not* being selected in a given year will be fairly low. Operations in these strata may not take much comfort in knowing that there is a (small) chance of non-selection, especially after being included in the sample for several years in a row.

Sometimes there is a threshold (size or dollar value) below which no operations are selected. Obviously, this removes all burden for those operations below the threshold, at least for that survey (although they may be in others). However, it is possible that the threshold merely eliminates units that would have been very lightly sampled (e.g., if the threshold was designed to cut off a small portion of the production of a commodity by removing a large number of small units). At some point, in particular where the population is small, to select a high proportion of those above a threshold may lead to resentment among those "in sample", especially if they become aware that they are always selected while other certain operations seem never to be selected. The challenges are to ensure that sample designs remain optimized to produce good estimates and suitable for the population; that exclusion thresholds remain relevant (adjusted) so that they do not exclude too many or too few; and that surveyed units do not get the impression that they are carrying a disproportionate amount of the response burden. In respect to the last challenge, there may be some simplicity and equity, especially for small populations of interest, in simply conducting the collection as a periodic census – every unit is "in sample", so no one operation feels overburdened or left out.

Mixed farms can be in scope for a number of agriculture surveys. Even if each contact is relatively short, the accumulated response burden may be seen as high and there may also be the sense that efforts are duplicated, for example if each different survey asks some of the same or similar questions (such as confirming frame information). Here the burden may not be the cumulative time to answer two or three short surveys, but may stem from the attempts needed to get the respondent on the line. This may be especially noticeable if the surveys come fairly close together in time. The respondent may rightly wonder why the information could not have been gathered in one session, or at least in fewer sessions.

Any agricultural operation can be in-scope for many data-gathering activities by governments, private entities, associations and so on. An operation will naturally be eligible for 1) Statistics Canada agriculture surveys, but also 2) other types of Statistics Canada business surveys, triggered by the presence of activities such as transportation, storage, processing, tourism, employing people, etc., as well as 3) environmental surveys, 4) household surveys, 5) social surveys, 6) surveys by vendors (e.g., of feed, equipment, etc), 7) surveys by industry associations (sometimes by regulation in order to operate in that industry), 8) surveys of type 1 to 5 conducted by their provincial statistical agency, 9) surveys or information requests from their provincial ministry of agriculture, 10) filing tax returns, possibly both personal and business, 11) completing forms to apply for federal, provincial or other agriculture programs, 12) applying for bank or other financing, 13) applying for crop or other insurance, 14) supplying information needed to get permits for use or storage of chemicals or other substances, etc., etc. The challenge is to co-ordinate survey efforts, certainly within a given agency, but also to the extent possible across agencies.

Agriculture is characterized by long stretches of intense work, often squeezed into short periods (such as planting or harvest times), and the weather frequently makes scheduling tasks difficult. So, farmers are busy, and may work off the farm as well, thus there is almost no good time to survey many farmers by telephone. For some information, it is important to gather it right at the time (e.g.,

in Canada, for many crops, there would not be much point to ask about planting intentions in August), and for surveys from which in-season estimates are produced, it is critical to collect and process the data in as short a time period as possible.

The tables below show some high level indictors of size of the agriculture population (number of census farms measured in the past six censuses) and the total sample size for a set of major surveys (field crops, livestock, fruit and vegetable, greenhouse).

Table 1
Number of census farms, Canada 1981 to 2011

Year	Number	Ratio (1981=1.00)	
1981	318,361	1.00	
1986	293,089	0.92	
1991	280,043	0.88	
1996	276,548	0.87	
2001	246,923	0.78	
2006	229,373	0.72	
2011	????	???	To be determined by the 2011 Census of Agriculture and published in May 2012

Source: Census of Agriculture

Table 1 shows that total burden is coming down on the Census of Agriculture mostly due to the declining number of farms (the content has been fairly constant in terms of length and time to complete). An internet response channel was introduced in 2006, which was used by about 5% of respondents. In that year, the time to complete the form on-line was about 50 minutes (median). So far in 2011, it looks like more will use the channel (the pre-collection estimate for take-up in the internet channel on the Census of Agriculture was 10%, which looks very achievable at the time of writing and could be surpassed), and the online (median) completion time in 2011 of approximately 30 minutes is below the 2006 figure, likely due to improvements in the software application and wider availability of high-speed internet service in Canada's agricultural regions compared with 2006.

Table 2
Sample sizes, major agriculture surveys 2006-07 to 2011-12 (excluding Census of Agriculture)

Year	Aggregate n	Ratio (2006-07 = 1.00)
2006-07	192,170	1.00
2007-08	192,132	1.00
2008-09	177,652	0.92
2009-10	175,209	0.91
2010-11	166,815	0.87
2011-12	157,177	0.82

Source: Main estimates funded surveys, as per loaded budgets.
Note: Aggregate n is the sum of the sample sizes for major agriculture surveys conducted in year shown; an agricultural operation may be selected into more than one survey, so the number of operations contacted will be less than the number shown in a given year.

This table shows that some significant progress has been made in reducing response burden as the sample contacted in 2011-12 will be only about 82% of what it was a few years ago. Note that part of the reduction from 2010-11 to 2011-12 is due to the cancelation of two smaller surveys and the

reduction by 60% of the sample size of a larger survey in recognition that 2011 is a Census of Agriculture collection year.

Timeliness

Even a field collection window of eight to ten days and estimates published 20 days later may be called "hopelessly outdated" by the agriculture press and other users of the information. Achieving extreme timeliness via surveying means short collection windows are needed and strong efforts in gathering the data during those short windows, both of which may be seen as burden by the respondents. The challenge is to design surveys that can be quickly answered correctly and that ask just the essential information. Meeting these objectives will also aid in keeping the processing/analysis period as short as possible, and a short time between the end of collection and the publication date is also crucial in publishing timely statistics.

On the other hand, such short collection windows will mean that follow-up of non-response will have to be relatively aggressive in order to achieve an acceptable response rate. The challenge is then to ensure that interviewers are well trained to do this kind of work while not upsetting or alienating respondents, and that procedures are as free as possible from irritants so that when the respondent is reached, the actual information collection can take place in a pleasant atmosphere.

Mode of collection, collection practices

As many agriculture surveys using CATI are quite short and quick to learn, it is not unusual for new interviewers to be assigned to these surveys as their first work assignment. Farmers often express frustration when asked agricultural questions by someone they perceive as not sufficiently knowledgeable about agriculture, and the fact that they may not yet have a great deal of interviewing experience may compound the problem. The challenge is to maintain a good core of experienced and agriculture-savvy interviewers for the agriculture surveys. This is not always easy as the higher profile "mission critical" surveys (such as the Labour Force Survey) have a continual demand for the top, experienced interviewers and will draw from agriculture and other interviewer pools. Also, interviewers themselves may view their graduation to more prominent surveys as a positive career move. The challenge here is to ensure that the training provided is of high quality, and can be offered quickly and regularly to maintain a staff of effective and productive agriculture survey interviewers.

Farmers do not seem to like getting phone calls, which presents a challenge when a prime collection mode for business surveys (including agriculture) is currently computer-assisted telephone interviewing (CATI). Given the nature of farming, the "telephoning window" is typically longer for agriculture surveys (7 am to 10 pm, including the time when call-back appointments may be scheduled, with the agreement of respondents) than for other business surveys, which would typically focus phone attempts into a shorter, business hours window. In many cases, the contact phone number for the farm business is the home telephone number; whereas for many other businesses, the owner/operator of the business would have separate telephone numbers for home and business. Thus, calls to agricultural producers may seem more intrusive than those made to other businesses. The challenge in managing telephone collection from agricultural producers is to design systems that will intelligently schedule and control calls and will accommodate flexible rules that will minimize the annoyance factor for the respondents. Ideally, collection systems will soon be available to allow wide freedom of choice to respondents in how they receive and return the questionnaires they do receive (a true multi-modal environment).

While the earlier discussion on the topic of cell-phones for general business respondents applies to agricultural operations, there may some unique issues as well. Although most businesses will want customers to be able reach them, and therefore will likely have a well-publicized land-line even if they do use cell-phones extensively, some agricultural operations may be deciding that land lines are less necessary, particularly in cases where no sales to the general public are expected. Also, in the case of the family farm, there may be a land line in the home, but the operator may carry a cell-phone to keep in touch with suppliers or for emergency contact while they are out in the fields or barns, which may be a very large portion of their days, especially at certain times of the year. Even if this cell-phone number was available to an interviewer, it is highly unlikely that being called to respond to a survey would be seen as an emergency, and would probably not be well-received. It will continue to be a challenge to maintain high rates of contact by telephone for agricultural operations.

A lot of hope is being held out for internet response as the mode of the near future, as this would bring numerous advantages such as allowing farmers some flexibility of when they respond (within the collection window), and would take the issue of interacting with an interviewer (possibly one inexperienced in agriculture surveys out of the equation. However, to date, the take-up by agricultural producers has not been high. In the 2006 Census of Agriculture, 5% of farmers chose this response channel (an internet response channel application was offered to all, but mail-back was the predominant method of responding with some also providing information over the telephone). As previously noted, it was hoped that the proportion of 2011 Census of Agriculture responses received through the internet would reach 10% and this should be achieved or even surpassed. The challenge will be to continue to encourage agricultural operators to embrace this response channel so that more of the associated benefits can be realized.

Collection costs / budget pressures

The agriculture data collection program has the challenge of managing in the face of rising survey collection costs. Some drivers are: generally rising costs of doing business (salaries both for interviewers - who became unionization a few years ago - and other staff, postage and telecommunications costs), increased effort needed to reach respondents (which increases staff time per case), more time and effort needed to elicit responses (rising time per unit, sometimes even when content has been streamlined), and the appearance of new costs (such as costs to develop, build, test and maintain electronic collection solutions) which may offset to varying degrees the savings from moving away from traditional methods. The challenge is to keep costs under control within (sometimes shrinking) budgets while maintaining the quality of the outputs.

Some of the ways collection costs are managed include: reducing sample sizes, ensuring content is minimal and equal to what is essential, use of technology and analysis to try to optimize interviewer effort (call schedulers, studies of para-data to identify improvements in methods, etc.), reviewing and reducing frequency of surveying when warranted.

Corporate Business Architecture (CBA) Initiative

There is a move within Statistics Canada towards centralization (frame, collection, processing, and dissemination) – agriculture is included in this initiative and some of the main projects involved are:

- Where the division was performing collection activities (generally, these were the smaller sample, more specialized commodity surveys or administrative data programs), move them into centralized areas;

- Increase the use of the electronic questionnaire mode - Statistics Canada's goal is to have it become the predominant or primary mode, but this may be some years off for agriculture until the coverage of high-speed internet access makes further gains;
- Move the Agriculture Division Farm Register (FR) into the Statistics Canada Business Register (BR), and run the agriculture survey program using the Business Register as the sampling frame;
- Move data processing from divisional systems onto centralized system(s) (Integrated Business Statistics Program);
- Increase the use of taxation and other administrative data, especially where it can replace the need to conduct surveys.

Most of these projects are underway or will start soon, with the objective of transitioning the agriculture program to the CBA in 2014-15. Some of these initiatives will result in changes to how data are collected (centralization of collection, moving the FR to the BR, increased emphasis on electronic collection), and a large challenge will be to ensure that discontinuities in statistical series arising from these changes are avoided or at least minimized. In addition, data users are usually understanding of the fact that statistical programs must evolve in order to remain relevant, but they also expect that changes will be as seamless as possible and be well-documented, in order that they may assess the degree to which these changes may affect their use of the data. Thus, agencies must ensure that a portion of their scarce resources are available to analyze the effects of such major transitions and do the necessary thorough documentation.

As a result of the centralization of survey collection, it is possible agricultural producers may resent being contacted by non-specialists based in centralized areas. In the case of Statistics Canada, regional collection centres are located in major urban areas for the most part. This has not been a major issue for the larger volume agriculture surveys that have been collected centrally for some time now, although there are still some comments and complaints. There is the possibility it could be an issue for the smaller, specialized commodity surveys with only a few respondents who may feel that they have lost personal contact with a known, regular data collector from Agriculture Division, with whom they may have been dealing for many years.

"Complex farm" issues

Producers who are complex because they raise a large number of different commodities may experience "survey overload" as they could be in just about every commodity/production survey. If they are a very large quantity producer of a number of commodities, they may be in the take-all stratum in several or many of these surveys. The challenge is to rationalize collection so that the information needed is collected as efficiently as possible from these important agricultural respondents. Statistics Canada deals with these larger and more complex agricultural operations through special handling procedures, and this approach is covered in detail in the paper *Canada's Enterprise Portfolio Management and Large Agricultural Operations Statistics (LAOS) Unit Approach*, presented in the session "Current Best Practices from Statistical Agencies."

If operating arrangements or geographic locations of the operation's parts are diverse, it may require multiple people to fully reply to the content of the questionnaire. This may trigger multiple call-backs (which may add to the feeling of burden for the respondent) and/or additional work at their end to gather all of the requested information. It also may not be completely clear to the respondent which businesses (or portions) to include/exclude since questionnaires will probably have been designed to deal with the most common operating arrangements in their layouts, explanations, instructions and examples. The limited space on most self-enumeration paper

questionnaires makes the situation more problematic whereas with a personal interview or online collection mode, there is the opportunity to provide more information and aid to respondents. The challenge is to "fit" the collection instrument, or a suite of instruments, as closely as possible to the respondent's operating situation and reporting preferences.

The respondent contact person may not be the same from one occasion to the next. This may affect data quality if the transition is not seamless, or may seem burdensome to the new contact while they are "learning the ropes" of reporting the requested information.

Although the current agriculture statistics program aims to cover all business units with agricultural activity (generally, all operations producing agricultural outputs intended for sale, although some surveys use a minimum sales threshold), the responding unit may also belong to a non-agriculture industry (most business surveys are industry-based). In these cases, the main activity will have led to the business being assigned (based on the NAICS – North American Industry Classification System) to an industry other than agriculture, although the unit will be of interest to the agriculture program and will be in-scope for its surveys. If the business engages in yet other activities outside of its industry, there could be other activity-based survey programs interested in that unit and it may be included in those samples as well. This may lead to confusion with respondents getting several (or many) different surveys in addition to agriculture.

Despite best efforts to align questions to the way respondents keep their information, their books may not be configured identically to the way the questions are posed on the survey collection instrument, and the respondent may feel burdened by being asked to process their numbers in order to provide the information the way it is being asked for. It is also somewhat of a paradox that efforts to simplify questionnaires, such as by reducing questions or by removing detail from the questions that are asked, may actually make it more difficult for the agricultural producer to answer. For example, it may seem tedious to ask a cattle operator to separately report bulls, dairy cows, beef cows, calves, steers and various types of heifers separately, but this information is likely readily available and would require simple transcription onto a form or recitation to an interviewer. Asking this detail can be more straightforward than asking for the total number of head on the operation, which may require some arithmetic and may require several rounds of back and forth dialog between interviewer and respondent to clarify what is included and excluded in the total (in the absence of clarifications, the quality of the reported totals may suffer). Simply having a total will also be less useful for data users when the statistics are tabulated. The challenge is to remain current with how operators keep their information (which likely parallels the evolving structure of their agricultural domain) and to balance data users' needs and desires with the reporting burden placed on producers.

Confidentiality/dominance

Where a large agricultural operation is a dominant player in a commodity grouping or geographic area (or a few are dominant), the result may be a primary suppression, so that there may be some "holes" in published tables. This may lead some respondents (the large, dominant ones who triggered the suppression as well as the others who didn't) to question the point of taking the time to respond if the results are not published. Primary suppressions lead to secondary suppressions (to prevent residual disclosure of the suppressed cells) - so even more respondents may wonder what is the point of responding.

The use of waivers (obtaining the consent of certain respondents to allow statistics to be published that will reveal some information about their individual operation) in order to allow publication of data in these cases may be an answer. However, obtaining the waiver also places further paperwork burden on the respondent (completing the form, perhaps consulting with others in the process of deciding whether to grant the waiver) and creates additional work for the data collector (possibly travelling to negotiate the waiver, renewing the agreements). The challenge is to design surveys to avoid confidentiality problems, but some of the other pressures (reducing response burden, declining number of farms) may make this increasingly difficult.

Conclusion

While the objective of this session, "Current and Imminent Data Collection Challenges Faced by Statistical Agencies", was to enumerate the issues and set the stage for further discussion at the meeting, rather than solve them, some common themes did emerge.

Increasing the use of administrative data already provided by agricultural operations appears to be a direction that must be pursued. Although most statistical agencies make use of these sources now, to make fuller use of them to the point of being able to substantially reduce the amount of surveying will take some time, financial investment and a will among the many holders of such data to work together (for example across provincial boundaries) and to work with the statistical agencies towards building integrated and coherent systems of administrative data.

While reduction of survey response burden was acknowledged as a good and worthy goal, it appeared that among the large and/or more complex operations, there is an equally strong (or even stronger?) desire on their part for the data collector to "get it right", that is, to take the time with them to properly understand the operation so that the data which are collected are correct, meaningful and useful.

It also seemed that the burden of responding may be connected not so much to the fact of being asked to reply to a survey, but to those situations where the operation has already provided the same or very similar information, either to another arm of the government or to the statistical agency on a previous occasion of the survey, or to someone else whom they believe has (or should have) the ability to share the information. Also, when nothing has changed, it can appear redundant to be asked to provide the information again. It seems that greater use of technology (e.g., to allow sharing of information, to allow pre-filling of information and only getting updates on changes, etc.) and innovative methods on the part of the statistical agencies would be welcomed by respondents.

Acknowledgements

The author would like to thank Lynda Kemp, Assistant Director, Agriculture Division, Statistics Canada and Jean-Pierre Simard, Director General, Agriculture, Technology and Transportation Statistics Branch, Statistic Canada for their useful comments and input on the various versions of this paper.

Related Reading

Respondent Relations Challenges for Obtaining Agricultural Data, Kenneth D. Korporal, Proceedings of Statistics Canada Symposium 2005, "Methodological Challenges for Future Information Needs", Statistics Canada

An Analysis of the Relationship Between Survey Burden and Nonresponse: If We Bother Them More, Are they Less Cooperative?, Jaki Stanley McCarthy, Daniel G. Beckler, Suzette M. Qualey, Journal of Official Statistics, Vol. 22, No 1, 2006, pp. 97-112.

Canada's Enterprise Portfolio Management and Large Agricultural Operations Statistics (LAOS) Unit Approach, Francine Lavoie, Chief, Enterprise Portfolio Management Program, Enterprise Statistics Division, and Steven Danford, Chief, Census of Agriculture, Agriculture Division, Statistics Canada. Paper presented in the session "Current Best Practices from Statistical Agencies", Enhancing Data for Complex Agricultural Establishments, Niagara-on-the-Lake, Ontario, Canada, June 26-28, 2011.

Statistics Canada Corporate Business Plan, 2009/2010 to 2011/2012, Available at: http://www.statcan.gc.ca/about-apercu/cbp-pae/index-eng.htm