# Complex Agricultural Establishments

*Prepared by the Planning Committee for an International Workshop on Enhancing Data for Complex Agricultural Establishments, Niagara-on-the-Lake, Ontario, Canada, June 26-28, 2011*[1]

## Background

The number and organizational structure of farming establishments varies across and within countries.[2] Some organizational attributes of a farming establishment present challenges and new opportunities in the development of agricultural and rural statistics and data bases. These attributes can affect data collection, accuracy of estimates, and the use of data, e.g., in multivariate and policy analysis, disclosure, and dissemination of estimates. We refer to establishments[3] as "complex" when they pose a high level of these types of attributes in the process of developing statistics and research/policy data bases representative of the target populations.

It is not possible to precisely define a complex establishment or the degree of complexity of an establishment, especially since establishments evolve over time as they optimize their objectives, plus effective complexity may vary by the environmental context. Nevertheless, the purpose of this paper is to provide a common understanding of the population of complex farming operations in light of the missions of federal statistical agencies[4]. The paper was developed to provide guidance and insights to an international workshop on the topic of data collection for complex agricultural establishments held June 26-28, 2011 in Niagara-on-the-Lake, Ontario, Canada.

The paper will be used to begin the discussion on alternatives for improving the ability of statistical agencies to provide users with statistical estimates and data bases. Increasing organizational complexity may offer data collection opportunities. For example, as one stage in production or marketing is concentrated in fewer entities, one aspect of data collection can potentially see increased efficiency. However, this will require additional resources to cultivate cooperative relationships with the population and coordinate the linkages across data collection activities in nontraditional ways. More often than not, increasing complexity of farms brings greater difficulty in data collection. We begin by discussing (1) the uses of the data bases and statistics and (2) the causes of complexity in farming establishments and identification of the organizational attributes of farming establishments that could be considered as "complex" for our

[1] Mary Ahearn (co-leader, mahearn@ers.usda.gov), Kevin Barnes (co-leader), David Culver (co-leader), Sheldon Jones (co-leader), Jeffrey Smith (co-leader), Koen Boone, Flavio Bolliger, Bill Iwig, Ashley Leduc, Jaki McCarthy, Jim MacDonald, Joe Parsons, Krijn Poppe, and Daniela Ravindra

[2] There are approximately 2.1 million farming establishments in the U.S., 230 thousand in Canada, 4.8 million in Brazil, and 14.5 million in the EU-27.

[3] We use the terms establishment, farm, and operation interchangeably.

[4] We have a broad definition of statistical agencies, to include economic agencies.

purposes, based on quantifiable characteristics of the establishments. (3) Finally, we consider the issues of complexity in the context of systematically managing other agency goals and responsibilities.

**Uses**

Farm survey and census data contribute to a multitude of end uses, too numerous to describe here. We consider two general classes of products developed from survey and census data, estimates and data bases. Surveys and censuses are the key information sources for critical production and economic *estimates* developed and periodically released, sometimes by economic class of farm, and often times for standard disaggregated geo-political units, e.g., states or provinces. An important part of the value of statistical estimates is that they are part of a long time series which place the current situation in context.

The second type of product developed from survey and census data are *data bases* for policy analyses. Policy analyses are often focused on distributional issues, e.g., they address questions about how current policies have affected economic performance or how proposed policies affect economic performance for certain subpopulations, as well as the aggregate population. For example, what factors influence farmers' decisions about adoption of new technologies, many of which are related to larger household issues? Policy makers must understand farmer decision making in order to institute policies that promote the farmer behaviors they are interested in encouraging. For policy purposes, therefore, it is important to have complete farm-level data because responses to policies will vary by farm and farm household characteristics and the subpopulation of focus will vary depending on the issue.

Experience in the U.S. shows that the average or mean of many indicators mask differences that matter. For example, of the 70 thousand farms with milk cows in 2007, the average dairy farm has 133 cows, but 2.3 percent of farms with 1000 cows or more produced 42 percent of all dairy product sales. Similarly, in Brazil, 40 percent of the largest farms (with 26 hectares or more) account for over three-quarters of total grain, oilseed, and meat production. Ignoring the distribution of economic activity can lead to unintended consequences for a policy to provide assistance to small farms and fails to provide information about the extent of the farm population in vulnerable financial positions.

Important policy issues in today's world are broad and encompass not just food production, but the rural economy, household incomes and environmental issues such as water quality, water availability, and climate change. Given the complexity of these issues, they often require that data from agricultural surveys and censuses be linked to other data sources, i.e., on communities and natural resources. Hence, farm data should have the capacity to be georeferenced or otherwise linkable to relevant data sets.

**Causes and Characteristics of Complexity**

Causes and characteristics of complexity include:

- production contracts
- marketing contracts
- vertical integration

- dispersed asset ownership, management, and returns
- use of farmland
- output sales discovery for open market sales

Many complexities are associated with the size of establishments. Large farms, in particular, are more likely to have more complex organizational structures than traditional, midsized family farms. In the U.S., a long-running trend of increasing concentration of production is expected to continue and this is expected to lead to increasing challenges in future data collection activities. There were 5,541 farms in the U.S. 2007 Census of Agriculture that sold more than $5,000,000 in the census year. All but two states (Alaska and Rhode Island) had farms of this size. The majority of farms in this group produced livestock or specialty crops, and produced more than 25% of the total value of agricultural production. Across all specialties, 449 operations produce 10% of total agricultural products, and 4,009 operations 25% of the total. (See Appendix table for information on the number of farms accounting for certain shares of commodity categories.) Furthermore, the data collection challenges are heightened for a concentrated industry because estimates from surveys that rely on stratified sampling often require complete enumeration of the largest operations. The levels that qualify an operation as a "largest" operation vary by survey. In some survey estimates, only a handful of operations may produce a large percentage of the total estimated amount. Maintaining the cooperation of the very large operations in data collection activities is essential to providing accurate estimates, and these are the very operations that must be contacted often for a number of surveys.

Many of the most important policy issues relate to the people engaged in agriculture, and the majority of people in agriculture operate small farms. Extremely small farm sizes can also pose challenges in data collection. Of course, the extent of this issue varies across countries, in part, because of differing definitions of a farm or holding. In the U.S., approximately 25 percent of all farms are point[5] farms. Although some small farms may be start-ups seeking to build their production over time, it is likely that others intend to stay extremely small farms for other financial reasons, such as to lower local property taxes, income tax management, and realization of capital gains on their farmland. The major data collection challenge for the extremely small farms comes, primarily, in their identification as a farm and, secondarily, in the separation of business and household expenses.

- Production Contracts

A production contract is a contract in which a producer produces, cares for, or raises commodities not owned by the producer, using land, equipment or facilities owned or leased by the producer, in exchange for payment. A production contract specifies, in detail, the production

[5] In the U.S., if a place does not have $1,000 in sales, a "point system" assigns dollar values for acres of various crops and head of various livestock species to estimate a normal level of sales. Point farms are farms with fewer than $1,000 in sales but have points worth at least $1,000. Point farms tend to be very small. Some, however, may normally have large sales, but experience low sales in a particular year due to bad weather, disease, changes in marketing strategies, or other factors. For farms with production contracts, the value of the commodities produced is used, not the amount of the fees they receive. Changes are made to the point system over time. For example, beginning with the 1997 Census of Agriculture, operations receiving $1,000 or more in Federal government payments were counted as farms, even if they had no sales and otherwise lacked the potential to have $1,000 or more in sales. And, for 2002, a farm that had $500 point value and $500 in government payments is considered a farm.

inputs supplied by the contractor (processor, feed mill, other farm operation or business), the quality and quantity of a particular commodity, and the type of compensation to the grower (contractee) for services rendered. Almost all broilers in the US are produced under production contracts, as well as the majority of hogs, and other livestock sectors. Production contracts are less common in crop production.

For establishments with production contracts, data collection for some items is a challenge because other parties may contribute inputs to the production and the operator may not be able to accurately report either the amount, the cost of inputs, or quality variations provided by others. Similarly, they may not be able to report the value of production. This lack of information seriously hampers the ability of a data user to understand differences in productivity and returns across operations. In addition, because of the competitive nature of the industries involved, there are sensitivities on the part of both contractees and contractors in providing detailed contract information. Even if it were the practice to contact the contractor for follow-up, some values may not be known to them because many contractors are vertically integrated establishments.

Farm establishments can also be the contractor in production contracts with other farms. For example, a livestock operation may contract with another operation to feed/raise livestock it owns and markets. For accurate accounting of net returns, the livestock sales will be included with the returns of the operation and any expenses paid by the operation for this service must be included in expenses.

On the other hand, some single data series, particularly inventory data, may be easier to collect when production contracting is adopted. If one entity owns the livestock raised on a number of contract operations, it is necessary only to contact the owner of the livestock to estimate inventories, not the person raising each barn of chickens or hogs.
Production contracting is one type of complexity that is related to farm size—large farms are more likely than small farms to engage in production contracting.

- Marketing Contracts

Under marketing contracts a producer enters an agreement with a downstream handler to deliver a specified commodity, with specified qualities, at a certain time period, for a specified price or pricing mechanism. Since the producer retains control over production decisions he or she is able to provide information on production decisions, including input prices. Hence, marketing contracts do not pose the same data collection challenges as do production contracts. However, data collection challenges in marketing contracts can arise when the final output prices are not known at the time of data collection due to a complex pricing mechanism or lags in marketing that occurs over multiple periods.

Most of the production marketed through marketing contacts is on relatively large farms, but a small, and increasing number of small farms, use marketing contracts. For example, marketing through community supported agriculture (CSAs) or having a predetermined arrangement with a restaurant to deliver product are types of marketing contracts.

- Vertically Integrated Operations

Vertical integration combines successive stages in the production and marketing process under the ownership or control of a single establishment or firm. Vertical integration poses challenges in data collection because some data items, most notably commodity prices, may not be defined, as they are in open markets. For example, production prices are usually defined at the farm gate. If an operation controls production from the field to the retail chain, a farm gate price may never exist. For example, a livestock slaughterhouse acquiring a cattle operation to better manage their supply target for slaughter is a case of vertical integration.

Individual establishments may engage in multiple marketing channels, further complicating data collection and estimation. For example, an operation that grows grapes for wine may sell some grapes on the wholesale market, and keep a portion of the crop to make into wine (i.e., downstream vertical integration). The value they receive for the wholesale grapes may not be equivalent to the value for grapes kept for value-added processing if the operation chooses to keep higher or lower quality products for in-house activities.

Most of the product produced by vertically integrated firms is from relatively large firms. However, some small farms engage in a type of vertical integration, such as an apple orchard that produces and sells its own cider. The sales of the cider are treated as income of the farming operation, i.e., farm-related income. This also implies that there must be a clear understanding of when a processed or value-added product should be considered as income of the farming operation.

- Dispersed Asset Ownership, Management, and Returns

There are many reasons why an establishment might have multiple asset owners and managers for farming operations. For one, the start-up and expansion costs in farming can be quite high, especially given the price of land. Just like in any business, a farm producer may seek investment partners, some of whom participate in some or all of the management decisions. Since a priority use of data for policy purposes is the development of well-being estimates for farm operator households, if all operators are part of the same household, contacting the farm business can also allow for farm operator household information to be collected. When a farm has multiple operators who do not share a household, developing well-being estimates for all farm operator households requires a follow-up to the farm operators who are not principal operators to determine their households' nonfarm income, net worth, and household characteristics. (The additional contacts have never been made in the U.S.)

Since farmland has historically been a very sound investment, it attracts outside investors, who do not participate in the management of the operation. Sometimes these investors invest in land management companies, contracting out the land management activity, and the companies then rent out farmland. Additionally, farming across the globe is generally a family business and is often left to multiple heirs. Oftentimes, heirs will sell their shares to the principal operator, but not always. For example, among Black farmers in the U.S. it is not uncommon for small farming operations to have many non-operator owners, all of whom are descendants of an earlier land owner. This is sometimes referred to as fragmentation. Data collection for non-operator landowners requires a follow-up visit—last done in the U.S. in the 1999 Agricultural Economics Landownership Survey.

Another cause of dispersed, and complex, farm ownership patterns is the result of government farm programs that set certain limits on the types of farming organizations that can participate in programs, such direct payments programs. In the U.S., corn producers receive the greatest share of direct payments and cotton producers receive the greatest per-acre payments. Although effective payment limits are quite generous to producers, some of the largest producers choose to reorganize their businesses so as to avoid the limitations. (The 2008 Farm Act provisions have sought to reduce that practice.) The effect of payment limit avoidance is to produce more organizational complex establishments with more sharing of ownership, management, and returns, thereby complicating data collection efforts.

Similarly, the organization of farming establishments may become more complex through the increase in owners and managers as a result of owners' motivations to (1) reduce tax burdens resulting from income tax laws and inheritance provisions and (2) reduce legal liabilities. This has implications for how farm operators and owners receive income from the farm establishment. For example, operators of C-corporations do not receive the net income of the farm as a sole proprietor would, rather they might receive dividend income or wage and salary income and might chose to retain some of the earnings with the farm business.

If an operation has many operators and/or many owners, data collection can be difficult for a variety of reasons. In the U.S., an operator is defined as the person(s) making day-to-day decisions for the farm operation. While the existing definitions of operators and operations facilitate the current approach to list building, the current concepts may no longer be reasonable concepts for some complex farms. For example, a complex operation may consist of multiple enterprises (perhaps in multiple locations) with separate managers for each enterprise. For example, a dairy farm that produces its own crops might have a crop manager and a cow manager. It may also have an accountant or bookkeeper, who manages business records, a human resource manager who controls information about employment and a marketing manager, who makes decisions about pricing and sales, as well as a general manager with overall control of the operation. For a given survey, it may be difficult for a single respondent to provide data for the operation and difficult to identify which respondent can respond to different data items for an establishment. Different persons involved in the operation may even provide alternative responses to survey questions. In addition, the contact person might change more often than a smaller operation with a single owner/operator. Perhaps most problematic, when there are multiple owners and managers, is tracing the net returns of the farming establishment that accrue to each of the parties. This is further complicated by the fact that data collection efforts in very complex establishments sometimes rely on fairly low-level staff to complete survey instruments, while most educational outreach efforts are geared at farm managers, farm owners, or high-level professional staff.

- Use of Farmland

Land management companies rent out land to farm establishments, offer services to farmland owners, and have varying degrees of involvement with agricultural land. If some of the land they manage is managed by them as a place that qualifies as a farm, they are part of the farm population to be sampled. For farmland that is rented out or managed in some manner for multiple farmland owners, a land management company may be the best contact for some information that is collected on surveys and censuses of farms.

Some farm establishments rent-in grazing land on a per-head basis from private and public organizations. In this case, a farm operator may not know and be able to report how many acres are being used exclusively for their purpose; the best source for the acreage information may be the entity renting-out the grazing rights. Land rented on a per-head basis can be rented from public or private agencies, industrial corporations, grazing associations, and from individuals under a short-term grazing arrangement. Knowledge of acres rented on a per-head basis is critical for land use statistics. Accuracy of land use statistics has increased in its importance because of international concerns regarding climate change and potential climate change mitigation policies. For farm financial analysis, being able to accurately measure farmland as a production input is essential.

- Output Sales Discovery for Open Market Sales

Collection of the value of sales for commodities sold on the open market is a basic economic data item for any farm survey or census. However, there is some evidence that the ability to collect this basic item varies by commodity and region. This is because, for some commodities in some regions, the value of sales may be net of marketing expenses, rather than gross of marketing expenses. Given that there may be variation across establishments about how best to collect sales (and marketing expense) information, the preferred approach is not clear. This question has been a long-running question for ERS and NASS experts in the U.S. Currently, the two major U.S. farm data collection efforts, the quinquennial Census and the annual ARMS, take differing approaches.

Some establishments market their output through grower cooperatives. It is not uncommon for producers in cooperatives to not have final sale information for their product at the time of data collection due to a lag in sales. Payments for product sales can come in the form of cooperative dividends. (The lag in information on commodity sales under a cooperative is not unlike the situation for sales under marketing contracts.) In the case of some commodities, e.g., rice, the cooperative receives the government farm payments on behalf of the grower and government payments are transferred by the cooperative to the grower along with dividend payments. In some situations, cooperatives do not attempt to separate out the source of the returns between product sales and government payments. Clarification of the sources of cooperative payments would require direct contact with the grower cooperative or administrative records.

**Complexity in Relationship with the Statistical Agency's Practices and Multiple Objectives: the Case of NASS**

A statistical agency with decentralized data collection must have policies and practices that are consistent across the individual data collection units, such as state offices of the NASS. In practice, standardization in data collection must be balanced with the need to recognize the diversity that exists among farming establishments and to address the agency requirement for efficient and accurate data collection while minimizing respondent burden. It is no small challenge to meet these multiple objectives, especially in a continually evolving agricultural sector and information technology environment. Strategic choices about data collection from complex farming establishments are made in recognition of other agency considerations. These include:

- Production and Management Occurring across Boundaries

There is a demand for NASS to publish a variety of estimates at the national, regional, state, district, and county level. Operations, especially large operations, increasingly operate across geographic borders, whether the borders are of county, district, state, or nation. It may be difficult to collect data from operations with many units for a variety of reasons, or to even identify the boundaries of the operation. Operations may not be able to completely define where production, expenses, and sales occur for a given item. Some survey items may apply to the whole operation, and parsing them to a geographic area may be difficult for a respondent. Depending on the nature of the operation, data may reside at the headquarters of the operation or at individual operating units. If data resides at individual operating units, it is hard to ensure that data is consistent across the units, when it is collected by different field offices.

- Disclosure Complexity

Very large operations, though not organizational complex, sometimes pose difficulties in the publication of estimates. If an operation makes up a large share of the total estimate, it may not be possible to publish the estimate without disclosing the data value of the largest operation. In other cases, the largest operations may be reluctant to provide data, because they do not want to provide any information to competitors in the market. Statistical agencies must be very strategic in collecting data from the very large operations because (1) it is a time-consuming activity on their part to respond to surveys and censuses, (2) they are contacted more frequently than smaller farms, and (3) their cooperation is essential for accuracy in many types of estimates.

- List Frame Construction and Maintenance

The structure of the sampling frames and the business rules of the survey organization need to accommodate each other. NASS uses both a list sampling frame and an area sampling frame to conduct their survey program. The list frame includes all known farming operations, compiled from various list sources. The area frame includes the establishments operating land for agricultural production in randomly located segments of land across the U.S. NASS maintains the operation name, if available, and the operator name for each establishment on each frame. Strict rules are needed for handling name changes and matching names between frames to ensure that data for a farm are represented on only one frame and not duplicated. For some surveys, it is also important to track historic data over time for an operation, even if the operator changes. Again, strict coding rules are needed to accommodate the survey requirements. Any enhancements to data collection procedures to better address collection from complex establishments need to recognize the list frame requirements.

- Complex Organizations and Administrative Data

In the U.S., NASS relies extensively on administrative data from USDA farm program agencies to build lists for sampling purposes and, in some cases, to link administrative data to survey data. The use of administrative data is especially important for respondents who fail to report on surveys and censuses. However, farms are often defined differently for administrative purposes than they are for data collection purposes, and this is especially problematic for farming establishments with a complex organizational structure. Linking the administrative unit to the defined operating unit (for data collection purposes) may be difficult, and subjective, based on multiple pieces of information that define a relationship.

## Implications

Farm surveys and censuses collect a variety of data, including inventory, expenditures, prices, and farm and household level financial data. The increasing organizational complexity of farming establishments offers both opportunities and challenges for improving the accuracy of estimates and policy data sets. In part, this will vary by the use and type of data required. For certain items in sectors dominated by production contracting, for example, inventory data may be relatively easy to collect. One contact with the contractor may be enough to collect individual inventory data for many contractees. Similarly, collecting acres rented on a per-head basis for individual farms and ranches may be collected by contacting public and private grazing agencies. While collecting the data from an entity, including administrative sources, that has knowledge of a particular data item for multiple farming establishments minimizes respondent burden and data collection costs, it increases the cost associated with the coordination of the process to link these data to individual farm records. For policy analysis purposes, accurate data must be available at the farm-level because flexible distributional information is essential.

Increasing complexity generally means collection of farm financial data will be more difficult. For example, with production contracting, a price may be impossible to obtain at the farmgate, since the contractor owns the livestock until sold to retail markets. Or, for whole farm financial accounting, when multiple parties have an economic stake in the operation, it may be necessary to make multiple contacts to ascertain the full picture of the economic health of the farm. Increasing concentration in production also implies that contact with the very largest operations be managed strategically and systematically across surveys.

Careful consideration of the complexity that exists in today's—and future—agriculture, in light of the multiple objectives of statistical agencies, will allow for the development of recommendations for improved data collection and better position statistical agencies to inform the important decisions of the day.

Appendix Table. Number of U.S. Largest Farms Producing Shares of Production, by Commodity Group, 2007

| *Percent of total sales* | *vegetables* | *fruits* | *grains* | *cotton* | *nursery* | *dairy* | *cattle* | *hogs* | *poultry* |
|---|---|---|---|---|---|---|---|---|---|
| *10%* | *27* | *37* | *2,913* | *106* | *29* | *125* | *38* | *33* | *87* |
| *25%* | *120* | *214* | *12,119* | *534* | *128* | *551* | *173* | *255* | *576* |
| *50%* | *551* | *1,339* | *41,100* | *2,213* | *614* | *2,418* | *2,862* | *1,713* | *3,643* |
| *75%* | *2,086* | *5,893* | *99,807* | *5,284* | *2,460* | *10,407* | *34,857* | *5,311* | *9,886* |
| *100%* | *69,100* | *112,690* | *479,467* | *18,591* | *50,784* | *69,763* | *798,290* | *74,789* | *148,911* |

*Source: special tabulation of the 2007 Census data*