Prepublication Copy
Uncorrected Proofs

Improving Data Collection and Measurement of Complex Farms

Panel on Improving Data Collection and Reporting about Agriculture with Increasingly Complex Farm Structures

Catherine L. Kling and Christopher Mackie, Editors

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

A Consensus Study Report of

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

ADVANCE COPY
NOT FOR PUBLIC RELEASE BEFORE
Monday, October 1, 2018
11:00 a.m. EDT

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by contract number 58-5000-7-0106 between the National Academy of Sciences and the Department of Agriculture. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-XXXXX-X
International Standard Book Number-10: 0-309-XXXXX-X
Digital Object Identifier: https://doi.org/10.17226/25260

Additional copies of this publication are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Copyright 2018 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2018. *Improving Data Collection and Measurement of Complex Farms.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/25260.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.nationalacademies.org**.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on studies' statement of task by authoring committees of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

For information about other products and activities of the National Academies, please visit www.nationalacademies.org/about/whatwedo.

PANEL ON IMPROVING DATA COLLECTION AND REPORTING ABOUT AGRICULTURE WITH INCREASINGLY COMPLEX FARM STRUCTURES

CATHERINE L. KLING *(Chair)*, Dyson School of Applied Economics and Management, Cornell University
J. G. ARBUCKLE, JR., Department of Sociology, Iowa State University
NORMAN M. BRADBURN, Department of Psychology, University of Chicago
RICHARD A. DUNN, Department of Agricultural and Resource Economics, University of Connecticut
ALLEN M. FEATHERSTONE, Department of Agricultural Economics, Kansas State University
JOSEPH W. GLAUBER, International Food Policy Research Institute, Washington DC
BRENT HUETH, Department of Agricultural and Applied Economics, University of Wisconsin–Madison
ANI L. KATCHOVA, Department of Agricultural, Environmental, and Development Economics, The Ohio State University
DORIS MOLD, Sunrise Agricultural Associates
JEAN OPSOMER, Westat, Rockville, Maryland
GREG PETERSON, Agriculture, Energy, Environment and Transportation Statistics Branch, Statistics Canada
KRIJN POPPE, Wageningen University and Research
DANIEL A. SUMNER, Department of Agricultural and Resource Economics, University of California, Davis
JAMES WAGNER, Survey Research Center, University of Michigan
JEREMY G. WEBER, Graduate School of Public and International Affairs, and Department of Economics, University of Pittsburgh

CHRIS MACKIE, Study Director
MICHAEL SIRI, Program Coordinator

COMMITTEE ON NATIONAL STATISTICS

ROBERT M. GROVES (*Chair*), Department of Mathematics and Statistics and Department of Sociology, Georgetown University
FRANCINE BLAU, School of Industrial and Labor Relations, Cornell University
MARY ELLEN BOCK, Department of Statistics, Purdue University (emerita)
ANNE C. CASE, Woodrow Wilson School of Public and International Affairs, Princeton University
MICHAEL CHERNEW, Department of Health Care Policy, Harvard Medical School
JANET CURRIE, Woodrow Wilson School of Public and International Affairs, Princeton University
DONALD DILLMAM, Social and Economic Sciences Research Center, Washington State University
CONSTANTINE GATSONIS, Center for Statistical Sciences, Brown University
JAMES HOUSE, Survey Research Center, Institute for Social Research, University of Michigan
THOMAS MESENBOURG, Retired, formerly U.S. Census Bureau
SUSAN A. MURPHY, Department of Statistics and Institute for Social Research, University of Michigan
SARAH NUSSER, Office of the Vice President for Research and Department of Statistics, Iowa State University
COLM O'MUIRCHEARTAIGH, Harris School of Public Policy Studies, University of Chicago
JEROME P. REITER, Department of Statistical Science, Duke University
ROBERTO RIGOBON, Sloan School of Management, Massachusetts Institute of Technology
JUDITH A. SELTZER, Department of Sociology, University of California, Los Angeles
EDWARD SHORTLIFFE, Department of Biomedical Informatics, Columbia University/Arizona State University

BRIAN HARRIS-KOJETIN, *Director*

Acknowledgments

This report is the product of contributions from many colleagues whom we thank for their time, generosity, and expert guidance. The project was sponsored by the United States Department of Agriculture's National Agricultural Statistics Service (NASS) and the Economic Research Service (ERS) to help them advance methods for collecting data and reporting information about American agriculture given the changes and increased complexity in farm business structure. Our points of contact at NASS and ERS, Kathleen Ott and Jeffrey Hopkins respectively, provided not only administrative guidance throughout the study, but also considerable subject-matter expertise as well, especially for the panel's public information gathering sessions.

The panel thanks the following agency representatives who attended open meetings and generously gave of their time to present material to inform the panel's deliberations: Catherine Woteki, USDA Undersecretary for Research, Education, and Economics; Renee Picanso, Deputy Administrator; NASS, Mary Bohman, Administrator, ERS; Barbara Rater, NASS; and Marca Weinberg, ERS. These individuals clearly articulated to the panel USDA's interests and goals for the study. Joe Parsons, NASS, provided detailed informed about the numbers and types of surveys conducted, the portfolio of products/publications created, and the stakeholders who use them. Marca Weinberg and Jeff Hopkins, ERS, outlined how the Agricultural Resource Management Survey (ARMS) fits into ERS goals and allows the agency to deliver information to stakeholders. Mark Apodaca, NASS, provided insights about how data are kept on NASS's list frame and about the sampling units used by NASS more broadly. Jim MacDonald, ERS, and Kathleen Ott, NASS, presented a statistical profile of the role and characteristics of complex farms in the United States. Kathy answered panel questions about NASS and ERS cognitive testing procedures for ARMS and the Census of Agriculture. Donald Buysse, NASS, presented an overview of data collection and methods for creating estimates from the Census of Agriculture. Jeff Hopkins, ERS, and Andrew Dau, NASS, provided a similar overview of ARMS, also describing key uses of the data. Jeff Bailey and Jody McDaniel, both of NASS, discussed the full range of production surveys conducted by the agency and explained how the roles of these surveys differ from the Census of Agriculture and ARMS. Kevin Barnes, NASS, described the agency's data collection practices across state offices. Linda Young, NASS, and Dan Prager, ERS, presented information to the panel about farm typology, including the uses and purpose of the "principal operator" construct. Chris Messer, who directs the NASS's pacific

region field office, discussed the effectiveness of current survey designs as they are implemented in the field. Lance Honig, NASS, discussed the agency's expanding use of GIS and remote sensing data for improving data collection and research, and reducing respondent burden. Cynthia Nickerson and Steve Wallender, both of ERS, presented information to the panel on current uses of administrative and commercial data in the production of agricultural statistics and visions for next steps.

Beyond USDA, Gaetan St-Louis, Director of Statistical Registers and Geography Division, Statistics Canada, described methods for dealing with complex statistical structures in Statistics Canada programs. Emily Berg, Iowa State University, presented methodologies for using administrative data across the agricultural statistics system. Amy O'Hara, then of the U.S. Census Bureau, described data linkage programs and use of administrative data across the U.S. statistical system more broadly. Ron Jarmin, U.S. Census Bureau; Dave Talon, BLS; Carrie Litkowski, BEA; and Patrick Canning, ERS provided comments on this topic from the perspectives of their respective agencies.

The panel also heard from farm operators and decision makers who, crucially, provide data to USDA through survey participation. Don Brown, Commissioner of the Colorado Department of Agriculture and himself a farmer, provided insights into workings of complex operations, about responding to the Census (and other surveys), and about complications created by the interaction of operations and family members. Providing insights that were intended to help NASS and ERS understand the respondent experience, improve the comprehensiveness and relevance of their surveys, and increase the value of resultant data and statistics to users were Kevin Phillips of Michael David Winery; Jim Rickert of Prather Ranch, Shasta County; and Tony Turkovich of Button and Turkovich Ranch, Yolo County. Each attended a panel meeting to share their observations about NASS data collection activities. Mr. Turkovich gave generously of his time by leading a tour of his farm so that panel members could observe first-hand the business relationships involved in the production processes of a large multi-product operation.

The panel was able to learn more about the demand for USDA data beyond "official" uses (e.g., mandates, principal economic indicators, administration of strategic goals, and enforcing laws and regulations) from Jody Campiche, V.P., Economics & Policy Analysis, National Cotton Council; Bob Young, Chief Economist and Deputy Executive Director, Public Policy, American Farm Bureau Federation; and Mitch Morehart, Authoritative Analytics. Jody McDaniel, NASS and Jim MacDonald, ERS, provided additional information about data demands for informing public policy issues and generating information used by stakeholder groups.

The panel could not have conducted its work efficiently without the capable staff of the National Academies of Sciences, Engineering, and Medicine. Brian Harris-Kojetin, director of the Committee on National Statistics, provided institutional leadership and substantive contributions during meetings; Kirsten Sampson-Snyder, Division of Behavioral and Social Sciences and Education, expertly coordinated the review process; and Marc DeFrancis provided thorough final editing that improved the readability of the report for a wide audience. We also thank program coordinator Michael Siri for his well-organized and efficient logistical support of the panel's meetings, as well as his contribution to assembling and formatting of this report. On behalf of the panel, I thank the study director, Christopher Mackie, for his unfailing good humor, excellent organization, patience in working with an often distracted chair, and attention to detail in all dimensions of producing the work. The quality and timeliness of this report would not have been possible without his substantial contributions.

Finally, and most importantly, a note of appreciation is in order for my fellow panel members. This report reflects the collective expertise and commitment of all panel members: J. Gordon Arbuckle, Iowa State University; Norman Bradburn, NORC; Richard Dunn, University of Connecticut; Allen Featherstone, Kansas State University; Joseph W. Glauber, International Food Policy Research Institute; Brent Hueth, University of Wisconsin–Madison; Ani Katchova, The Ohio State University; Doris Mold, Sunrise Agricultural Associates, LLC; Jean Opsomer, Westat; Greg Peterson, Statistics Canada; Krijn Poppe, LEI Wageningen UR; Daniel Sumner, University of California, Davis; James Wagner, University of Michigan; and Jeremy Weber, University of Pittsburgh. This group—chosen for their diverse perspectives, backgrounds, and subject matter knowledge—gave generously of their time to attend meetings and to apply their expertise in the production of this report.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Academies' Report Review Committee. The purpose of this independent review is to provide candid and critical comments that assist the institution in making its reports as sound as possible, and to ensure that the reports meet institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

The panel thanks the following individuals for their review of this report: David E. Bell, Agriculture and Business, Harvard Business School; Mary Ellen Bock, Department of Statistics, Purdue University; Katherine Smith Evans, Government Relations, American Economic Association; David Freshwater, Agricultural Economics, University of Kentucky; Eldon Gould, Producer, Gould Farms, Maple Park, IL; Carol House, Independent Consultant; Philip L. Martin, Department of Agricultural and Resource Economics, University of California, Davis; and Mitchell Morehart, Owner, Authoritative Analytics, LLC., Palmyra, VA. Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of the report was overseen by Cynthia Clark, National Agricultural Statistics Service (retired), and Charles Manski, Professor of Economics at Northwestern University. Appointed by the National Academies' Report Review Committee, they were responsible for making certain that the independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the report rests entirely with the authoring panel and the National Academies of Sciences, Engineering, and Medicine.

Catherine Kling, *Chair*
Panel on Improving Data Collection and Reporting about Agriculture with Increasingly Complex Farm Structures

Table of Contents

Summary

America's farms and farmers are integral to the U.S. economy and, more broadly, to the nation's social and cultural fabric. A healthy agricultural sector helps ensure a safe and reliable food supply, improves energy security, and contributes to employment and economic development, traditionally in small towns and rural areas where farming serves as a nexus for related sectors from farm machinery manufacturing to food processing. The agricultural sector also plays a role in the nation's overall economic growth by providing crucial raw inputs for the production of a wide range of goods and services, including many that generate substantial export value.

If the agricultural sector is to be accurately understood and the policies that affect its functioning are to remain well informed, the statistical system's data collection programs must be periodically revisited to ensure they are keeping up with current realities. Perhaps the most obvious change in recent decades is that large, complex farms have grown in number and are now responsible for the majority of agricultural production in the United States. Furthermore, the traditional portrayal of farms as self-contained, family-operated businesses does not accurately characterize these entities. The goal of this study is to review, assess, and make recommendations to the U.S. Department of Agriculture's (USDA's) National Agricultural Statistics Service (NASS) and Economic Research Service (ERS) to help identify effective methods for collecting data and reporting information about American agriculture, given this increased complexity and other changes in farm business structure in recent decades.[1]

A wide range of research and policy questions create the imperative for government-collected data on farms and farming. Beyond the value of agricultural statistics in creating a complete economic profile of the country (for example, for the National Income and Product Accounts), their role is crucial in informing policies on the environment, climate change, biodiversity, food security and safety, population health, land use planning, and natural resource management. The safety and quality of the nation's food supply and the health and environmental impacts of production processes are among the most important policy areas that agricultural data and statistics help inform.

[1]The full charge to the panel is presented and discussed in Chapter 1, section 1.2.

NASS and ERS publish statistics and reports that regularly and extensively detail the number of farms in the United States, the quantities and types of commodities they produce, the incomes of both the farm businesses and the farm households that run these businesses, and the status and conditions of the agricultural economy. To justify this public investment in the nation's statistical system, surveys and other data collection instruments must satisfy a range of demands. The justification for data collection by USDA is particularly compelling in cases where (i) data are needed to effectively administer government programs; (ii) data are used in the analysis of policy design and impacts; (iii) data are essential for research on agriculture, health, food, and environmental concerns; or (iv) data improve the workings of markets. NASS and ERS reports and data products often serve one or more of these purposes. In order to maximize the value of data programs, there is a responsibility to make the downstream products as accessible and useful as possible to policy makers, researchers, and other data users while maintaining both privacy (individuals are only asked to reveal information that is necessary to fulfill approved tasks) and confidentiality (the information provided is only shared with appropriate individuals for approved purposes).

The centerpiece of NASS's mandated responsibilities is to administer the Census of Agriculture, as required by law under the Census of Agriculture Act of 1997. The main objective of the Census of Agriculture is to provide an accurate portrayal of farming in the United States in terms of the number of holdings, their activities, their size distribution, and other characteristics. For the Census of Agriculture, NASS attempts to collect information from all of the nation's farms.[2] Because the Census of Agriculture is used to produce county-level estimates, the production and land associated with each farm need to be attributed to a county or counties.

USDA is required by Congress, through authorizing or appropriations legislation, to produce statistics on a range of topics, many of which are estimated using data collected through the Agricultural Resource Management Survey (ARMS), which is jointly conducted by NASS and ERS. ERS is also mandated to publish cost-of-production information for a number of commodities. ARMS is an annual cross-sectional survey that is unique in that it collects, in a representative sample, information on (i) field-level farm practices, (ii) the farm business, and (iii) characteristics of the household operating the farm. Roughly 30,000 farms are sampled each year for ARMS, and the sampling unit for this effort is the operation-operator pair, that is, the farm operation and the associated person who runs it.

The mandates handed to NASS and ERS typically specify the type of information required, but not how that information should be generated. As such, the agencies have considerable latitude in how they collect data from different kinds of farms and how statistics on their activities and finances are produced. For this reason, efforts to improve or streamline the Census of Agriculture, ARMS, or other surveys—in terms of content, questionnaire structure, and design—are unlikely to hinder the agencies' ability to fulfill their mandates.

Furthermore, USDA has considerable flexibility to explore nonsurvey sources of data, such as tax and Farm Service Agency records, which have already been used to improve a number of their data products. Mandates to NASS and ERS generally do not constrain the use of administrative, commercial, web-based, or other sources that could complement or, in some

[2]USDA defines a farm as "a place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the census year." As explained throughout the report, the way terms are defined directly affects many of the measures produced by NASS and ERS. The Glossary contained in Appendix 2.1 to Chapter 2 includes definitions of key terms as used by USDA and as used in this report (and will be explained as in some cases they differ).

cases, possibly substitute for elements of the current survey-centric apparatus. Indeed, given the kinds of information on the agricultural sector that have high value to stakeholders, expanding the breadth and diversity of the data sources drawn from represents a natural evolution for the statistical agencies measuring the sector.

THE ESSENTIAL PERSPECTIVE OF DATA PROVIDERS: BURDEN, RESPONSE RATES, AND DATA ACCURACY

When assessing a data collection infrastructure, key considerations are the burden imposed on respondents and the accuracy of the statistics produced. Minimizing the burden placed on survey respondents is a matter of deep concern at statistical agencies for several reasons. The most obvious reason is that time is a valuable productive resource, and thus mitigating respondent burden reduces the total cost of data collection. Another important motivation for reducing burden is provided by survey research,[3] which suggests that increased burden may reduce the willingness of farmers to respond to the entire survey (unit nonresponse), to respond to particular questions (item nonresponse), or to give careful and accurate responses (measurement error). A high respondent burden therefore can have a deleterious effect on the robustness of findings and conclusions based on analyses of the resulting data. Minimizing respondent burden is an especially pressing challenge in today's climate of declining survey response rates and increasing survey costs.

Compared with small farms, large farm operations—where the roles and relationships among multiple owners and managers and the operations they oversee are more difficult to unambiguously identify—have been found to exhibit lower response rates to ARMS and Census of Agriculture questionnaires. Several of the recommendations in this report are intended to reduce this respondent burden by making information requests clearer, with the overarching motivation of seeking to increase the accuracy and interpretability of the information collected.

WHAT IS A COMPLEX FARM?

As complex farms have become commonplace, the traditional portrayal of farms as self-contained, family-operated businesses no longer accurately characterizes the entities responsible for the majority of agricultural production in the United States. There is no set definition of a "complex farm." Rather, multiple factors place farms along a spectrum of complexity. Among these factors are the operational and management organization of a farm business, the number and diversity of commodities produced, and the amount of vertical integration in the business. The following dimensions of operational complexity, as well as *how these complexities affect the collection of data from farms,* are identified and discussed in this report:

- *Farm size*. Larger farms often have multiple (and sometimes nonoperating) owners, multipart management structures, complicated legal entities and relationships, and multiple commodity enterprises operating in far-flung locations.
- *Geographic dispersion*. Beyond the operational and management complexities created, the geographic dispersion of complex farming operations presents data collection and reporting challenges. For example, widely dispersed operations add to the difficulty of

[3]For examples, see Hansen (2007), Galesic and M. Bosnjak (2009), and Beckett et al. (2016).

assigning the production of a single farm to individual counties, which is required for widely used county-level statistics.

- *Multifarm, multibusiness* (including value-added) *operations*. When several operations are overseen by a single management entity that shares capital and other inputs, it can be difficult to isolate the prices and quantities of inputs and outputs associated with any one operation.
- *Farm-connected "nonfarm" output*. Measurement may be further complicated when a business's activities straddle farm and nonfarm production, especially when the latter is closely linked to the former. Determining where to draw the line when reporting production or income is sometimes difficult, both for the business and for the data collector.
- *Use of hired managers and labor-contracting entities*. Hired and family labor are both treated as employees of a farm, but contract labor is not. Consequently, data collected on two operations that acquire labor differently will not be comparable; the farm that contracts its labor will appear to have fewer employees than an otherwise similar farm that hires directly.
- *Multiple and dispersed asset ownership*. The presence of more than one person involved in ownership or management complicates the attribution of production and income and makes linking a farm business with a farm household less straightforward.
- *Management and decision making structures*. Farms range in structure from those run by a single "principal" operator to those operating as part of a large corporate entity. In the more complex cases, decision-making responsibilities for different aspects of the business are distributed across multiple parties.

DEFINING FARMING, FARMS, AND FARMERS

Better measurements of the complex farms responsible for the majority of contemporary agricultural production in the United States can yield more informative answers about important agricultural policy issues. Measuring complex farms accurately and consistently requires carefully specified definitions. In line with the day-to-day use of the word and current practice at statistical agencies, the definition of a *farm* should focus on the productive entity as a business engaged in clearly specified types of activities:

> **For conceptual purposes, NASS and ERS should define a farm as an establishment (single unit with a legal or informal management structure) that (1) has its principal or secondary activity in farming with the production of agricultural products and biological assets such as seeds or animals; and (2) for which full economic data on key business variables, such as costs and revenues, can be collected. (Recommendation 4.2)**[4]

This definition, which is similar to USDA's approach of associating a farm with a management unit, is intended to help unravel the structure of some common types of complex

[4]Not all report recommendations are included in this summary. The recommendation numbering refers to the chapter in which each recommendation can be found in the report. In this case, Recommendation 4.2. is the second recommendation in Chapter 4.

holdings. Defining a farm in this way means that the number of farms in the economy and their average size, as measured by a statistical agency, partly depends on how farmers organize their businesses.

Distinguishing among the different levels of a business operation becomes crucial for the purpose of establishing sampling units on which to base surveys and then for interpreting the resulting data.

USDA should consider adopting definitions of (1) *farm establishment* as a business establishment engaged in farming and (2) *farm business* as a collection of business establishments with at least one farm establishment linked by common ownership or control. (Recommendation 5.1)

The *farm business,* defined as a farm or farms sharing a common management structure, includes both cases where one business owns and operates one establishment (a simple farm business) and cases where one business owns and operates a group of establishments (a complex farm business). Currently, whether respondents consider their productive activities to be one farm or more is at their discretion and open to interpretation. Giving farmers guidance by using clear definitions, while also taking steps to increase the likelihood that data collection aligns with the way farmers organize their businesses, is important for reducing farmers' confusion about how to report information.

Defining the *farmer* as the owner of the business entity signals the importance of those who are responsible for decisions made on the farm and who bear all the financial risks. Risk management is an important aspect of U.S. farm policy, implying that this is an important identifying characteristic of a farmer. However, to characterize and understand agricultural production, both ownership *and* management are of interest. Folding both into one category under the term *producer* is one solution. In 2017, NASS introduced this term to replace the term *operator,* and has used it to indicate any person involved in decision making, whether in day-to-day farm management or as absentee owners who may only be involved sporadically in investment or hiring decisions.

As with identifying the farm itself, minimal guidance is currently given on how to identify the farm operator, the person making day-to-day decisions or, if multiple operators are involved, the *principal operator*. Identifying the principal operator, as is requested for ARMS, can be particularly challenging when different people have primary responsibilities for distinct aspects of the farm, such as the management of marketing and the management of crop production, as well as in cases of spouses or partnerships. This vagueness can create confusion for respondents, even if they are making an earnest effort to match their responses to the intent of the question. Interpreting the resulting survey data also requires assumptions about how respondents have understood the terms used in the survey instrument. The way units of measurement are understood ultimately affects statistics on the number of farms, the size and scope of the farm sector, and the farm population.

Because not all establishments owned by a business involved in farming are necessarily farms, another issue arises concerning accounting boundaries. The essence of the problem, and an important source of complexity in agricultural production, is that farming activities take place within a larger food and agriculture supply chain. When policies or programs require information on overall agricultural production in the United States, as opposed to what happens "on the farm," some farming and agricultural activities—carried out by establishments that are not

classified in a statistical framework as farms but instead as agricultural support firms, food processing companies, or retailers—should be considered within the scope.

A related boundary issue is that farms may engage in secondary activities that are not farming but, rather, food processing or retail activities closely linked with farming. Cheese making and roadside or farmer's market stalls selling farm produce are both examples.

These sectoral overlaps mean that a census of farm establishments is not a census of *exclusively* farming or agricultural activities, because some of the included farms also engage in a subset of activities outside farming, some of which could get misreported. Meanwhile, some farming and agricultural activities take place in establishments that are not classified in a statistical framework as a farm, but instead as an agricultural support firm, a food processing company, or a retailer. It would be desirable for some purposes—such as for the national income and product accounts, in which a complete, non-double-counted accounting of farming or farm product retailing is required—to extend survey coverage to include these secondary or smaller activities.

Here again, currently used definitions give considerable flexibility to the respondent concerning how to report the data, which has implications for data accuracy, interpretability, and respondent burden. Therefore:

> **In line with statistics for other parts of the economy, NASS and ERS should apply clear rules based on the nature of its principal productive activities for classifying a business as a farm or as an entity operating in a nonfarming sector with secondary activities in farming. (Recommendation 4.1)**

This recommendation does not imply that only entities classified as farms with farming as a primary activity are of interest to NASS and ERS. On the contrary, the agencies should be interested in all businesses engaging in farm activities, even if those are minority activities. To maintain comparability over time, counting farming activities in businesses that are classified in sectors other than farming becomes even more important when production shifts from simple to complex farm operations.

In addition to agriculture, an agribusiness complex has been created in the United States and other large economies. This means that agricultural policy, as well as environmental policies that target farming, has effects on sectors other than farming and agriculture. To give policy makers and the public insight into these interdependencies, developing statistics on this agribusiness complex is important. This can be done with a methodology based on the input-output tables of the national accounts (and their satellite accounts) that link farming to the activities in other sectors:

> **NASS, the Census Bureau, and the Bureau of Economic Analysis should report on the size of the agribusiness complex and its components in terms of income, employment, and environmental impacts and develop a program that harnesses existing data collection efforts to create a new satellite account for reporting on food and agriculture industries. (Recommendation 4.4)**

Finally, policy makers and researchers are not only interested in farms and farmers but also in *farm households*, because household dynamics influence the behavior of the farmer and operation of the farm. A well-known example is when investment decisions are made by farmers over long time horizons, where their supply responses to policies are influenced by whether or

not they have a successor. This interest in the total income and well-being of the farmer and the farm household, which factors into ERS mandates, is especially relevant for the family farm. On very large farms and in complex farm businesses, the family dimensions are relatively less important to the functioning and stability of the operation. Such organizations are more like big family firms in other sectors. The NASS and ERS definition of a farm household—as those who share dwelling units with principal farm operators of family farms—is consistent with the way the term "household" is used more broadly across the statistical system.

A DATA COLLECTION STRATEGY FOR IMPROVING THE MEASUREMENT OF COMPLEX FARM OPERATIONS

When considering the appropriate statistical unit for measuring complex farm operations, the motivating question should be, what is the measurement objective? Conceptually, there are three types of statistical units that can come into play, each with a distinct emphasis:

1. The *business*: the farm operation (later redefined as a statistical enterprise/establishment),
2. The *people*: individual farmers and farm households, and
3. The *land*: farmland, subdivided into fields.

Linkages exist between each of these statistical units. For example, ownership, decision making, and employment are associated with the business and the individuals and households involved; and the business is associated with a geospatial coordinate(s). Designing sample frames that maintain reliable linkages between statistical units should be a high priority for a data collection program, because such linkages can be used to indirectly generate representative samples of statistical units across different frames. The USDA already has a well-established sampling frame methodology that deals with some of the complexity brought on by the presence of multiple statistical units that are all of interest.

By using a combined frame of farm businesses and individuals, NASS and ERS track the linkages between the two. This structure works well for simple farms, but measurement issues arise when operation complexity increases. When farms are complex, so that there is no longer perfect overlap between the business unit, the household, and the location, this ambiguity makes it difficult for NASS and ERS to accomplish their missions of providing policy makers, researchers, and producers with reliable estimates of agricultural production activity.

The first step in creating an integrated data collection strategy that can deal with complex cases more systematically is to create a Farm Register. Of course, USDA already has a farm list, which embeds many of the characteristics that are required. However, the existence of multifarm, multibusiness operations, along with the complexity of the management and decision-making structure of these businesses, requires modifications to the current combined establishment-household list-frame approach.

To address the above-described ambiguity that results when farms consist of more than a single-unit farm establishment, the proposed approach would simplify sampling by maintaining separate lists of farms, farm operators, and land holdings, so that the sample unit selected can be the one that is optimal for measuring that characteristic. For instance, information on off-farm income is best obtained from a household-type survey, rather than a survey that targets farms.

NASS should expand on its list frame to create a Farm Register that provides an ongoing enumeration of all farm establishments in the United States. (Recommendation 5.2)

This Farm Register would be similar to the current NASS list frame, but it would focus on the enumeration of farms as businesses and the characteristics of those businesses. It would be an "evergreen" product, regularly updated as new information becomes available.[5] Survey-specific list-frames would be drawn from the Farm Register at a single point in time to support individual statistical programs, including the Census of Agriculture and ARMS.

The Farm Register should follow a farm establishment/farm business structure (as defined above) similar to that of the Census Bureau's Business Register. A farm establishment would be the smallest unit that can report agricultural production, including revenue, expenses, and employment. Each establishment would have an industrial classification, corresponding to its primary activity; however, for the reasons articulated above, secondary activities should also be identified.

Consistent with the above recommendation that NASS and ERS be more prescriptive in their designation of statistical units, a farm business would encompass a collection of farm establishments that are linked by ownership and control:

All farm establishments in the Farm Register should be linked to a farm business. In most cases, farm businesses will include only one farm establishment, but they may include more than one. (Recommendation 5.5)

The following information should be maintained on the Farm Register for each farm establishment:

- Primary NAICS codes for the farm establishment,
- Commodity output flags (North American Product Classification System[6]),
- Name and address of farm,
- Other geolocation indicator,
- Size indicator (sales, number of employees), and
- Linkage variables (e.g., EIN).

The farm business designation corresponds to a statistical enterprise in Census Bureau nomenclature.

The purpose of the agriculture statistics programs in NASS and ERS is to cover all farm activity, regardless of the industry of the statistical unit. The Farm Register may therefore contain enterprises and establishments that do not have agriculture as a primary activity. For instance, an enterprise that is primarily engaged in processing farm products may also operate its own farms. Although most of that enterprise's value-added could be associated with processing, and thus classified as manufacturing, the farming activity still needs to be captured.

[5]Regular, although not continuous, updating of the Farm Register makes sense. NASS is aware of this, because it keeps its area frame's "found" farms separate from the currently listed farms for the purpose of estimation.

[6]Product classification is often more complicated in agriculture than in heavily capital-based industries because its primary asset is land, the utilization of which is more flexible.

Given the requirement to produce statistics on the financial well-being of farm households, the farm register would include linkages between the statistical units (farms or individuals) and households. USDA already identifies *some* people associated with farm operations, such as those involved with decision making, employment, ownership, and contacts for surveys. However, this is not done systematically or completely, as the aim in the past has been to identify a principal operator among all the persons involved with a farm.

Ideally, a household/individual frame would include the households of all the operators. The operators enumerated are not necessarily the same as the people who should be listed as contacts for survey purposes.

NASS should create a separate list frame of farm households within the overall Farm Register that would lead to a more efficient sampling of farm households and/or persons involved in farm activities, since the household list itself can be stratified or augmented with auxiliary data. (Recommendation 5.6)

Building on the existing operator list frame maintained by NASS, the Farm Register should consist of a set of relational databases that include information on places and people and that identify households and businesses with suitable links between the two. This approach would also improve continuity between operator and household records and address problems that arise when the primary operator changes, especially in cases of spouses, two generations of operators (co-principal operators), or partnerships.

Building on this framework, the Census of Agriculture could be recast as a source of basic structural characteristics that in turn creates a sampling frame for more focused surveys. ARMS and its various subcomponents could be reformulated into an annual farm establishment survey, one that collects the information needed for measuring the cost of production and the financial health of farms, including the information needed by the BEA for national economic statistics. Periodic, specialized surveys can be used for any questions not needed for these purposes or for mandates that explicitly require annual collection. Some immediate benefits of such a reorganization would be a reduction in respondent burden and the ability to reallocate USDA resources toward managing the data collection needs of complex operations.

THE POTENTIAL AND LIMITATIONS OF ADMINISTRATIVE AND OTHER NONSURVEY DATA SOURCES

Survey-based products, derived from a well-designed Farm Register, can be combined with other sources of data to improve the overall quality and utility of information on the farm sector while reducing respondent burden.

Broadening the scope of data sources for the measurement of complex farm operations is consistent with efforts across the federal statistical system to increase reporting capacity by exploiting administrative, commercial, and nonstructured (including web-based) sources. Use of nonsurvey data for the production of agricultural statistics is an approach increasingly being undertaken by statistical agencies around the globe and indeed by USDA itself; geospatial data and numerous administrative sources are prominent examples. While both NASS and ERS currently use nonsurvey data sources for statistical purposes, there is even greater potential for their use. For example, they may be used to facilitate the construction of sample frames, to validate data collected from survey instruments, to augment existing collection efforts to handle

nonresponses or missing information, and to contribute to data processing through model-assisted calibration, model-based estimation, and imputation of survey responses.

As has been documented in numerous reports—most recently and prominently that of the Commission on Evidence-Based Policymaking (2017)—the use of administrative data can improve the overall efficiency of data programs by reducing agency expenditures, lowering respondent burden, encouraging the sharing of information across departments and agencies, and potentially increasing the accuracy of information based on survey data. Since administrative data are maintained to support many USDA programs, the scope of these potential applications is vast.

How effectively the federal statistical system can meet future data demands will largely depend on the extent to which data sources—survey and nonsurvey, national and local, public and private—can be combined in synergistic ways. This assertion certainly applies to NASS and ERS programs, since their current surveys alone can no longer provide all the variables and levels of geographical detail necessary to meet the demands of agricultural research and policy making.

USDA should explore opportunities for record linkage at the person level to obtain information on key demographic and off-farm employment variables, and perhaps with IRS on farm income and expense information. These opportunities can be explored through participation in the Federal Statistical Research Data Centers program, a partnership between federal statistical agencies and leading research institutions that provides secure access to restricted-use microdata for statistical purposes. (Recommendation 6.1)

NASS and ERS have already developed a data access mechanism in which ARMS data are accessible for statistical purposes through a cooperative agreement with the University of Chicago's National Opinion Research Center. This arrangement works well for those who want to work with ARMS data alone, but it does not provide opportunities for linking with data from other agencies.

Any redesign of the Census of Agriculture and ARMS should be done with the presumption that these instruments will need to be linkable to other data sources maintained by USDA, to other statistical agencies, and even to nongovernment organizations. The key design element in the data system for promoting data linkages—for example, between household records and farm business records—is created during questionnaire design. If units of measurement are consistent, then in principle a crosswalk among a range of data sources can be maintained.

Given the work of the Commission on Evidence-Based Policymaking to improve the climate for legislative changes that would make data linking more routine across the statistical agencies, now is the time for NASS and ERS to begin mapping out a strategy to coordinate their survey and administrative data programs, both within USDA and across other key agencies such as the Census Bureau and the Bureau of Labor Statistics.

1
Introduction

1.1. WHY MEASURE THE ACTIVITIES OF FARMS AND FARMING?

Agriculture Is a Large and Important Sector of the Economy

America's farms and farmers are integral to the U.S. economy and, more broadly, to the nation's social and cultural fabric. A healthy agricultural sector helps ensure a safe and reliable food supply and improves energy security. It contributes to employment and economic development, traditionally in small towns and rural areas where farming serves as a nexus for related sectors from farm machinery manufacturing to food processing. And it contributes to the nation's economic growth overall, providing crucial raw inputs for the production of a wide range of goods and services, including many that generate substantial export value.[1]

In 2015, farms directly contributed $136.7 billion to the U.S. economy, representing about 1 percent of gross domestic product (GDP).[2] When one includes related sectors, such as forestry, fishing, tobacco products, textiles and apparel, and food service—which add value to raw farm outputs by using them as inputs in downstream production—the overall contribution of the agriculture sector to the economy is considerably larger.[3] The United States Department of Agriculture's (USDA's) Economic Research Service (ERS) estimated that, in 2015, "agriculture, food, and related industries" generated goods and services valued at $992 billion, a 5.5 percent share of the nation's GDP. In terms of its contribution to employment, ERS estimated that, in 2015, 21.0 million full- and part-time jobs were related to the agriculture and food sectors—11.1

[1]Agricultural exports for the United States reached an all-time record high of $152.3 billion in 2014; due to a number of factors, including lower commodity prices, that level has fallen in subsequent years (USDA, Economic Research Service, *Outlook for Agricultural Trade*, May 2016).

[2]See https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/ag-and-food-sectors-and-the-economy.aspx

[3]Defining the "farm" sector and the "agriculture" sector in unambiguous terms, and then factoring in how value-added processes link the two, is far from a trivial task. This task is undertaken in this report, most directly in chapters 2 and 4.

percent of total U.S. employment. Direct on-farm employment accounted for approximately 2.6 million of these jobs—1.4 percent of the total for the country.[4]

Agricultural Activities Have a Large Impact on the Nation's Natural Resources

Farming and ranching activities occupy just over half of the 2.3 billion acres of land coverage in the United States (Nickerson and Borchers, 2012). The size of this footprint varies significantly by state and region, as illustrated by Figure 1-1. In some states, such as Iowa, Kansas, and Nebraska, natural lands have undergone a nearly complete conversion to agriculture, so that farmland makes up roughly 90 percent of all acreage in the state. In others, such as Alaska, Massachusetts, and Nevada, there has been much less conversion (farmland making up less than 10 percent in these three states). These alterations in land use have come with dramatic changes to wildlife habitats, water use, hydrology, water quality, and biodiversity. In addition to agriculture's direct effect on these natural assets through land conversion, agricultural activities use more than 25 percent of all the surface water and 85 percent of all the groundwater withdrawn for human use in the United States.[5] Agricultural activities also contribute about 25 percent of total GHG emissions in the country.[6]

Finally, agricultural activities are significant sources of water pollution. For example, agriculture is responsible for more than 90 percent of the loads in nearly 70 percent of nitrogen-impaired watersheds in the country, contributing to drinking water concerns, lost recreational opportunities, and aesthetic losses. For these reasons, it is essential to capture the contribution of farming when measuring the economic impacts of alternative conservation programs, the efficacy of policies designed to protect the environment, challenges related to climate change and water scarcity, and enhancements to agricultural competitiveness through technology.

[4]See https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/ag-and-food-sectors-and-the-economy.aspx.

[5]See https://water.usgs.gov/edu/wateruse-diagrams.html.

[6]See https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data.

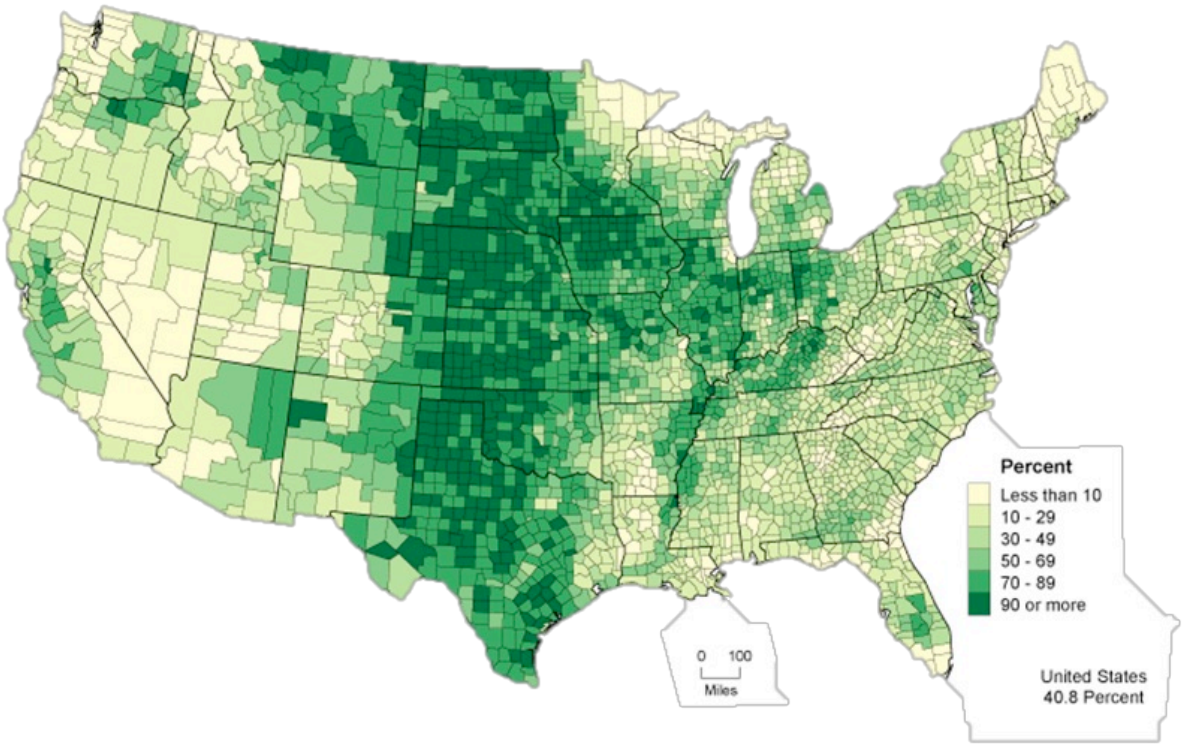


Figure 1-1 Acres of land in farms as percent of land area in acres, 2007
Source: USDA, NASS[7]

The U.S. Government Allocates Substantial Financial Resources to the Sector

The U.S. government has provided income and conservation support to agricultural enterprises since the Great Depression, which motivated the Agricultural Act of 1933. A major objective of federal farm policy then was to increase and stabilize farm household incomes through price and income supports for selected program-covered crops and dairy. In farm bill programs enacted during the 1990s, annual land set-asides and government stockholding to raise farm prices were largely abandoned in favor of payments tied to historical production. A large portion of these payments flows to the largest 2.1 million farms in the United States (USDA, 2014). Government subsidies are highest for grains, oilseeds, cotton, sugar, and dairy products. Most other farms—including those producing beef, pork, poultry, hay, fruits, tree nuts, and vegetables, which together account for about half of the total value of production—receive only minimal government support.

Overall for the 2008–12 period, the U.S. government spent $114 billion on 60 programs providing financial assistance to farmers. The largest farm payment programs administered by the USDA's Farm Service Agency (FSA) are those providing subsidies for crop insurance and

[7]See https://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Ag_Atlas_Maps/Farms/Land_in_Farms_and_Land_Use/07-M079.php (Accessed August 2018).

direct commodity payments. For the same period (2008–12), the USDA reported spending about $28 billion on crop insurance subsidies to help farmers manage risks associated with losses in yield or revenue. The Congressional Budget Office (CBO) has projected that crop insurance subsidy program costs will rise to about $8.8 billion annually over the 2015–24 period. The CBO further projects that farm commodity program costs will rise to about $4.2 billion annually for the same period.[8]

In addition to farm support payments, the federal government provides substantial financial resources to support conservation and environmental programs on farmland. The largest program is the Conservation Reserve Program, which has retired more than 37 million acres from cropland production at its peak. That program and others, including the Environmental Quality Incentive Program and the Conservation Stewardship Program, have provided more than $60 billion in recent years to enhance conservation.[9]

Data and Statistics about Agriculture Are Widely Used by the Public, by Researchers, and by Policy Makers

Economists have long argued that agricultural statistics are largely a public good. Bonnen (1977) points out how improving the quality of agricultural statistics can improve public policy through a better understanding of policies' effects on society. Key users of the information produced by the National Agricultural Statistics Service (NASS) and ERS include the USDA and other government policy makers, Congress and the White House, program administrators and managers, federal statistical agencies (including for international reporting), state and local government officials, farm and industry groups interested in public policy issues (including nature conservation), and researchers. Publicly available data also contribute to the efficient operation of markets and are used by farmers, ranchers, and other businesses for planning and forecasting.

A wide range of research and policy questions drive the need for government-collected data on farms and farming. Beyond the value of agriculture statistics in creating a complete economic profile of the country, such as for the National Income and Product Accounts, their role is crucial in informing policy across a broad range of knowledge and activity. This includes policy on the environment, climate change, biodiversity, food security and safety, population health, land use planning, and natural resource management. As a nation, we care deeply about the safety and quality of our food supply and the health and environmental impacts of our production processes. These are the most important policy areas that agriculture data and statistics help inform. To maximize the value from investment in the nation's statistical system, government has a responsibility to make publicly funded data collections as accessible and useful as possible to researchers, policy makers, and the public.

Some data programs are driven by legislative and regulatory requirements. U.S. farm policy typically follows a five-year legislative cycle that produces an omnibus "Farm Bill." The Farm Bill—the most recent being the Agricultural Act of 2014—governs

[8]Congressional Research Service, "Overview of Farm Safety Net Programs," CRS Report IF00025, *In Focus*, May 13, 2014; costs from CBO. The commodities covered under the FSA program are wheat, feed grains, rice, peanuts, soybeans, other oilseeds, dry peas, lentils, and chickpeas. Details at https://www.gao.gov/assets/670/664671.pdf

[9]See https://ageconsearch.umn.edu/bitstream/174075/2/cmsarticle_374.pdf.

> programs related to farming, food and nutrition, and rural communities, as well as aspects of bioenergy and forestry [and] authorizes policies in the areas of commodity programs and crop insurance, conservation on agricultural lands, agricultural trade (including foreign food assistance), nutrition (primarily domestic food assistance), farm credit, rural economic development, agricultural research, State and private forestry, bioenergy, and horticulture and organic agriculture.[10]

Publicly available agricultural statistics are an essential element of good governance and are also essential to enhance competitiveness. Sound administration of government agricultural programs requires information on the uptake of support provided to farmers to guide the development and monitoring of farm business and farm household assistance packages.

Agricultural statistics also support market efficiency, providing information used in management, in research and development, and in the investment decisions made by farmers and other business owners. Data on agricultural production, yield, and prices affect commodity markets and support trading and distribution systems for agricultural products. There is a robust literature (e.g., Myers, Sexton, and Tomek, 2010; Adjemian, 2012; Karali, 2012; Mattas and Silvera, 2014; Sanginabadi, 2018) indicating that USDA reports can have a significant impact on commodity futures markets. Research by Isengildina-Massa et al. (2018) indicates that alternative data sources have a limited impact on the dependence of markets on USDA-published information. They find, for example, that "the largest impact [was] on the Crop Production reports released later in the growing season, when production was more directly measurable using satellite and precision agriculture tools such as combine yield monitors" (p. 21). Statistics on market prices and commodity stocks are also used by industry organizations to monitor demand and supply at local and national levels.

There are also reporting requirements that must be submitted to international regulatory bodies charged with monitoring global food stocks, chemical residues in food and feed commodities, and other food quality controls in order to maintain access to international markets. Statistics on export values and volumes may be used by governments to prioritize products and market access activities.

Much is known about productivity and productivity growth in the manufacturing and retail sectors that allows capital resources to be reallocated from low- to high-productivity activities. Similarly, in agriculture, market characteristics and trends must be measured to answer questions such as: Are the farms that are disappearing the least productive ones? Are farmers with superior methods and processes the ones surviving? How do production processes differ across operations? How is the health of rural economies affected by globalization and by more services being provided by far-away firms? How has the changing value chain of agricultural products affected the lives of farm families? The answers to these questions are not all well understood, but they can be improved upon by tapping into a rich vein of data linking together households, businesses, and land use so that successful businesses, and the characteristics of the individuals who run them, can be examined.

Farming Is a Rapidly Changing Enterprise

If the agricultural sector is to be accurately understood, and the policies that affect its functioning are to remain well informed, the statistical system's data collection programs must

[10]https://www.ers.usda.gov/topics/farm-economy/farm-commodity-policy/.

be periodically revisited to ensure that they are keeping up with current realities. Perhaps the most obvious change in U.S. agriculture in recent decades has been large farms' growth in number and in economic influence. As illustrated in Figure 1-2, in 2014 about 76,000 farms had sales of $1 million or more. These farms accounted for 57 percent of production, while another 128,000 midsize farms[11] accounted for 21 percent of production and the remaining 1.87 million small farms[12] accounted for only 22 percent of production. Indeed, many of these small farms are run by part-time producers who have off-farm jobs to supplement their income (or serve as their primary income) or to provide a source for health insurance.

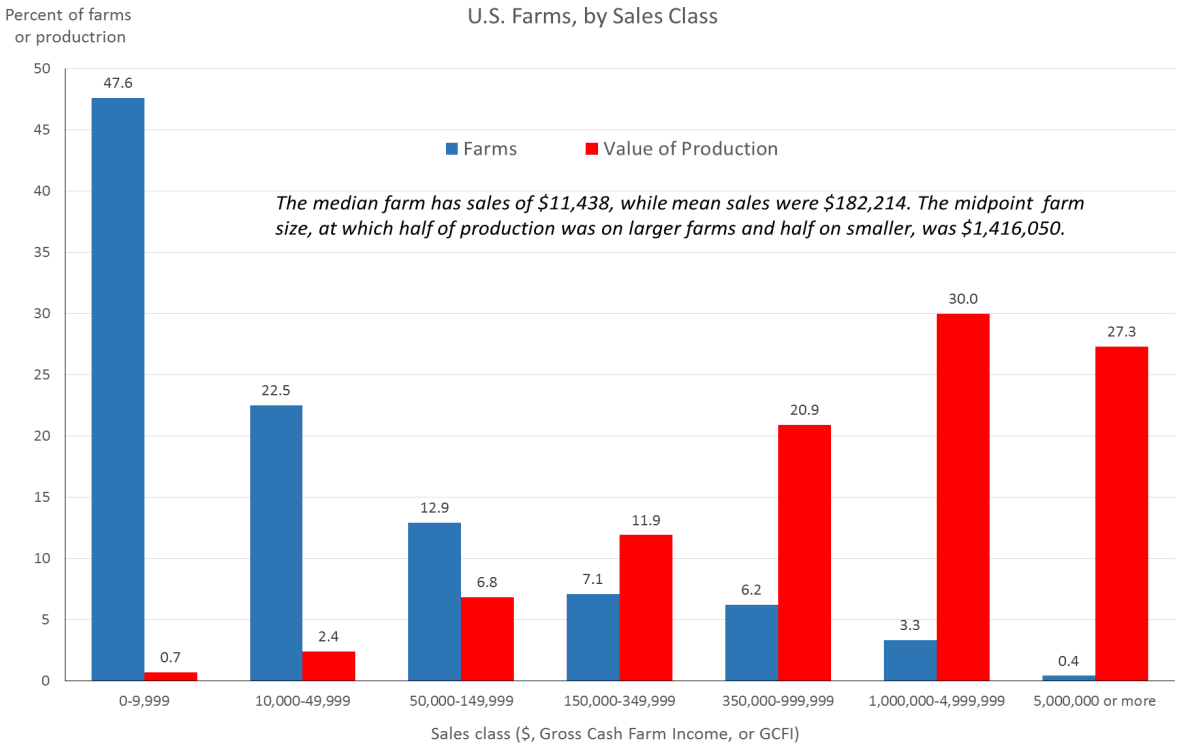


Figure 1-2 Distribution of U.S. farms in number and production value, by sales class, 2014
SOURCE: Presentation by James MacDonald and Kathleen Ott; data from USDA Agricultural Resource Survey 2014. Reprinted with permission.

[11]Midsize family farms are defined by ERS as those with a gross cash income of between $350,000 and $999,999.

[12]USDA defines a farm as "any place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the year." Places without at least $1,000 in sales receive points based on their potential for sales. Places with at least $1,000 in sales and/or points are considered farms. (A farm with less than $1,000 of actual sales but with enough points still to qualify is referred to as a *point farm*.) A "family farm" is a farm in which "the majority of the business is owned by the operator and individuals related to the operator by blood, marriage, or adoption, including relatives that do not live in the operator household." Chapter 3 explores alternative definitions, some of which are adopted by statistical agencies outside of the United States.

These figures portray a markedly different picture than existed in the United States even 20 years ago. Figure 1-3 indicates the extent to which production has shifted from small farms to large farms over this period. The share of total production from small farms fell from 40 percent in 1996 to 22 percent in 2014, while the share from large farms rose from 33 to 57 percent. As production has become increasingly concentrated in a small number of large farms, farm operations have also become more complex.[13] This report's extensive focus on large vertically (and horizontally) integrated farms is motivated by the growing share of economic activity in the sector accounted for by such operations.

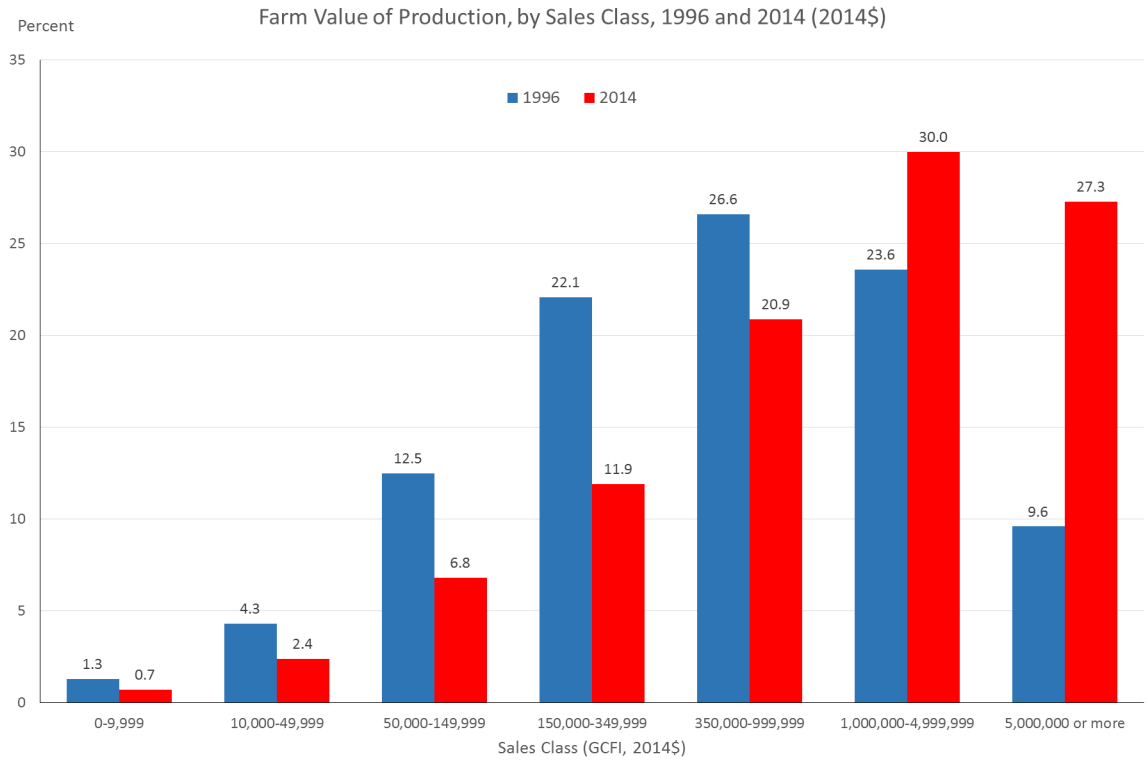


Figure 1-3 Distribution U.S. farms, in production value by sales class, 1996 and 2014
SOURCE: Presentation by James MacDonald and Kathleen Ott; data from USDA Agricultural Resource Survey, 1996 and 2014.
NOTE: Sales for 1996 adjusted to 2014 prices using BLS Producer Price Index for Farm Products.

Because of the skewness in production, simple means and medians focused on average farm size present only a superficial and in some ways misleading portrait of American agriculture. This becomes evident, for example, in statistics on cropland acres: The number of midsize crop farms has declined sharply, while the number of farms at the large and small extremes has grown. As shown in Figure 1-4, the size of the *average* farm (in acres) has changed little over the past four decades or so, but large farms have gotten considerably larger. Simple

[13]At the other end of the size spectrum, increasingly specialized farms have emerged that are smaller, more dispersed, and more transient than traditional farms. Some specialized farms emphasize a single product in traditional rural locations; others are emerging in less traditional locations, such as high-tech, vertical operations (often indoor) in urban areas.

means and medians focus on the average farm, but relatively few cropland acres reside on average-size farms.[14] Likewise, given the wide array of statistical measures of mean and standard deviation, some measures are more appropriate than others for reporting farm household income and wealth and communicating trends effectively.

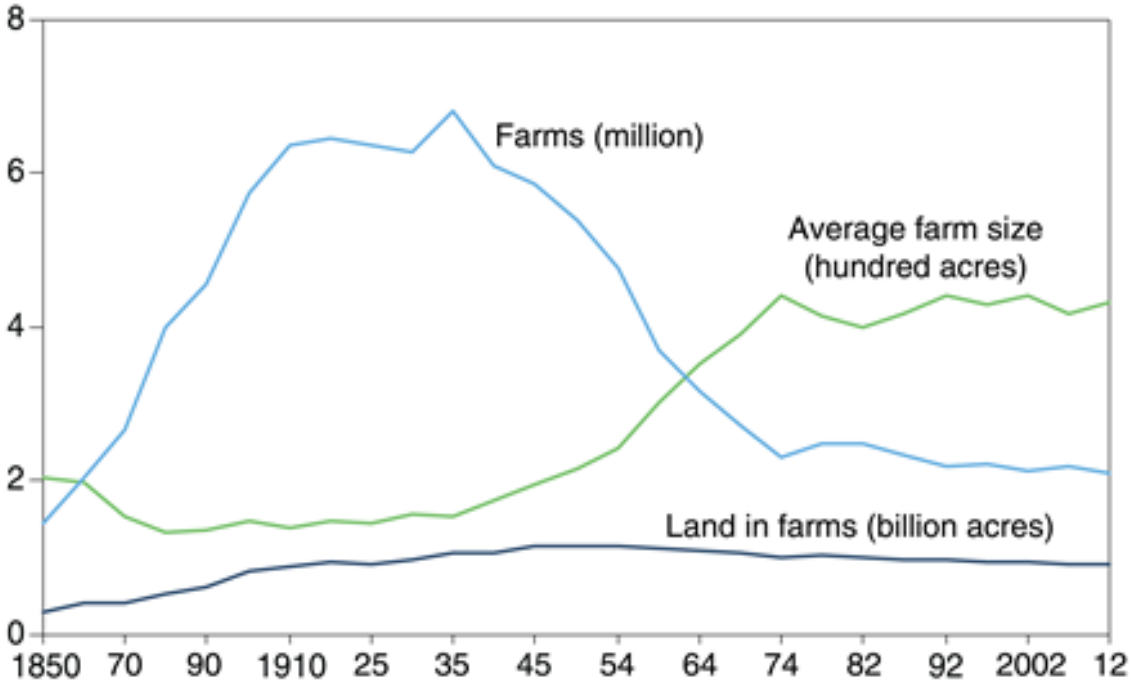


Figure 1-4 Farms, land in farms, and average acres per farm, 1850–2012
SOURCE: USDA, ERS, using data from USDA, NASS, Census of Agriculture (https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/farming-and-farm-income/). Accessed August 2018.

The historical change in farm size can also be represented by the distribution of farmland. Figure 1-5 tracks the "midpoint acreage," which shows the land-size point at which half of all cropland acres are on farms with more cropland and half are on farms with less cropland (MacDonald, Korb, and Hoppe, 2013). By this measure, the midpoint amount of cropland acreage on U.S. farms has nearly doubled since 1982. Meanwhile, the size of the average (mean) farm has changed little.

[14]See https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=76862.

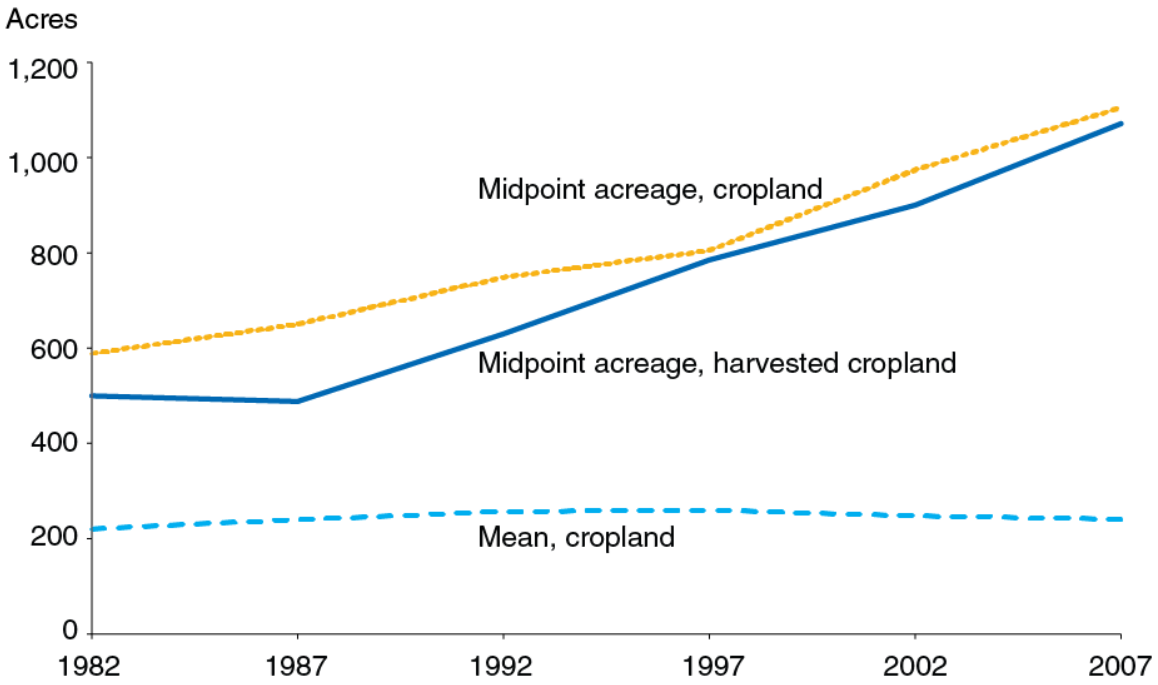


Figure 1-5 Midpoint acreage versus mean cropland trends, 1982–2007
SOURCE: Mean cropland from MacDonald et al. (2013), using USDA NASS data from the Census of Agriculture. Midpoint acreages from ERS calculations, based on unpublished Census of Agriculture data.

It is important to note that a large farm is *not necessarily* a complex farm, while a small farm may have organizational and production complexities. However, as discussed in detail in Chapter 3, large farm businesses are more likely than others to display dimensions of operational complexity, such as the presence of multiple owners, operations, and legal entities; more complicated management and decision-making structures; greater geographic dispersion of operations; and more extensive contractual relationships. Farming is also a rapidly evolving sector of the economy, influenced by changing technology, increased automation, and modernizing business structures, and large-scale farms are leading the automation charge in the United States. Examples abound: Driscoll's Berries in California is reconfiguring its operations to allow for the introduction of robot harvest crews to its fields and to bring industrial automation to its supply chain.[15] Companies like PrecisionHawk offer farmers real-time crop monitoring through use of drones in the field that communicate via a tablet or smartphone. In Kansas and elsewhere, farms use automation to monitor water application and operate irrigation activities from the office. In California's Salinas Valley, Taylor Farms uses water knives operated by artificial intelligence and, at Hahn Family Wines, a sophisticated water monitoring system is in place. Almond farms are using moisture sensors to monitor soil conditions, with the resulting data passed back as inputs into the automated irrigation systems.[16] And of course tractors are now equipped with GPS sensors, which improves their efficiency in covering ground in the fields and,

[15]See https://www.forbes.com/sites/currentaccounts/2017/03/02/with-farm-labor-getting-scarcer-big-u-s-farms-are-preparing-to-turn-to-robots/#15865f1e7bb6.

[16]See http://www.economist.com/technology-quarterly/2016-06-09/factory-fresh.

in turn, reduces fuel bills and improves the uniformity of fertilizer, herbicide, and pesticide spraying.

These complexities pose challenges to the statistical agencies (NASS and ERS) whose program missions involve collecting data used to measure and report on the activities of farms and farmers, and in some cases on agriculture more broadly, while seeking to limit the respondent burden of those supplying the information. For example, the operational complexity of a farm could create uncertainty about who should be contacted to respond to surveys, what entities within the business the respondent represents and has accurate information about, and even (in the presence of value-added activities and multi-operations) what to report on. These questions were less pressing when conventional single-family farm and principal operator concepts—of the kind that much of the data collection and methodology is still based on—dominated production. The composition of farm households' income, a key indicator of U.S. farms' well-being, has also changed in ways that add complexity to measurement. Specifically, off-farm income has been rising sharply as a portion of total household income, which naturally means that on-farm income has declined.[17] Data collection and statistical programs must account for these changes in farm businesses and households if agricultural statistics are to remain accurate and useful for the broad range of purposes to which they are put.

Sorting out the measurement complexities created by modern agriculture is important even for generating the most basic summary statistics for the sector. Table 1-1 provides a different cut on the distribution of farms by size (in this case by the amount of acres harvested). It is difficult to assign a level of confidence, for example, to the estimate that there were 496 farms with between 10,000 and 25,000 harvested acres each in 2012, because it depends on how farmers define their "operations" when interpreting a census or survey question, and likewise on how the statistical agency processes the information. One large farm may divide operations in such a way that it does not reach the acreage threshold, while a similar sized farm may divide things in a way that does. As farm structures have changed, reporting to NASS and ERS surveys has become more difficult for respondents because more of them have to track entities with multiple operations and operators, with value-added activities, or with geographically dispersed operations and ownership structures.

[17]ERS, "2017 Farm Income Forecast," February 7, 2017. All values are adjusted for inflation using the chain-type GDP deflator, 2009 = 100 (OMB, Historical Tables, Table 10.1).

Table 1-1 Distribution of farms, by amount of harvested acreage, 1987, 1997, and 2012

Farm size (in harvested acres)	Number of farms			Harvested acres (000,000)		
	1987	1997	2012	1987	1997	2012
2,000 - 4,999	6,570	16,692	26,404	17.7	46.4	75.3
5,000 - 9,999	513	1,267	3,323	3.3	8.0	20.8
10,000 - 24,999	99	173	**496**	1.3	2.3	6.8
25,000 or more	11	18	26	0.6	0.8	1.1
All	7,193	18,420	30,249	22.9	57.5	104.0

SOURCE: Presentation by James MacDonald and Kathleen Ott; data from Economic Research Service summaries of farm level records in National Agricultural Statistics Service, Census of Agriculture.

1.2. STUDY OBJECTIVES AND STATEMENT OF TASK

The goal of this study is to review and assess current methods for collecting data and reporting information about American agriculture, to help identify those that are most effective, and to make recommendations on the same to USDA's NASS and ERS. This has become increasingly important to undertake given the rising complexity of, and other changes in, farm business structure in recent decades. Guidance on how to coordinate and prioritize data collection is part of an overall strategy to continue building toward a data system that is well suited to measure modern farming and agriculture.

The agencies (NASS and ERS) are seeking a structure for reporting on complex farms while not ignoring smaller farms, particularly in emerging enterprises, such as indoor and urban farming. The emphasis in this report is on understanding farm structures at a conceptual level, with guidance reflecting the diversity of agriculture, the need to reduce the burden for respondents, the urgency of addressing root causes of declining response rates, and the need to create operational efficiencies and better processes within the agencies. The report should prove timely, as the sponsors are aiming to have a revised system of data and statistical programs by 2022.

Box 1-1 Statement of Task

An ad hoc panel will review, assess, and make recommendations for the National Agricultural Statistics Service (NASS) and the Economic Research Service (ERS), USDA, on effective methods for collecting data and reporting information about American agriculture given the changes and increased complexity in farm business structure. Although the vast majority of today's farms continue to be run by a single operator or by spousal partners, the large farms that produce a substantial percentage of the nation's food tend to have more complex business structures. The panel will take into consideration the effect any changes in concepts and data collection practices could have on estimates of the number of farms and the reliability of sector finance and performance indicators. Specifically, the panel will

- Review existing information about the structure of U.S. farms, and how the information is collected, reported, and used for policy, research, and other purposes;
- Seek to identify best practices for accounting for multi-unit operations and operations that are vertically integrated, both on the farm register and in data collection and estimation, while ensuring sufficient coverage and reliable estimates;
- Seek to identify best practices for identifying and collecting information about ancillary or value-added economic activities that may be associated with a farm; and
- Examine the concept of the "farm operator" under different business structures (particularly the practice of attempting to identify one primary operator of a farm household) and the effects of a change in concept on the subsequent estimates of farm household finances and existing data series.

The panel will produce a final report with findings and recommendations at the conclusion of the study.

Guidance is also provided in a context of rapidly changing data collection opportunities. While a survey-centric approach is still very much emphasized within the U.S. statistical system, the recommendations made here echo those made in recent reports—such as those released by the *Commission on Evidence-Based Policymaking*[18] and the *Panel on Improving Federal Statistics for Policy and Social Science Research Using Multiple Data Sources and State-of-the-Art Estimation Methods* (NASEM, 2017)—which envision a much more substantial role for administrative, commercial, and other non-survey sources.

1.3. KEY THEMES AND REPORT STRUCTURE

This report addresses key challenges facing the statistical agencies of USDA, whose missions are to provide timely, accurate, and useful statistics in service to U.S. agriculture,[19] and to

> anticipate trends and emerging issues in agriculture, food, the environment, and rural America and to conduct high-quality, objective economic research to inform and enhance public and private decision making.[20]

[18]See https://www.cep.gov/.
[19]From the mission statement of NASS, https://www.nass.usda.gov/.
[20]From the mission statement of ERS, https://www.ers.usda.gov/.

Among the challenges faced by NASS and ERS in their quest to modernize their statistical programs are: (i) accounting for the increasing complexity of farm organization through data collection and reporting; (ii) accounting for the increased concentration of farming firms and production; and (iii) addressing the implications of more complex farming operations for estimating farm household financial indicators. In all three areas, maintaining coverage and statistical reliability of farm financial and production statistics are key considerations.

This report is organized into three parts to address these measurement issues: first, the current data infrastructure maintained by NASS and ERS—which includes among its key components the Census of Agriculture and the Agricultural Resource Management Survey (ARMS)—is described. Next, constructs that underlie the definitions of key production units are explored. This provides a basis for identifying the operational complexities of modern farm businesses, as well as the resulting measurement challenges. In the final chapters of the report, guidance is offered for addressing the measurement of complex operations in a way that would help modernize data collection by USDA statistical agencies.

In addressing the charge to the panel (see Box 1-1), Chapter 2 takes on the preliminary task of reviewing existing information collected and published by NASS and ERS about the structure of U.S. farms and how information is collected, reported, and used. The centerpiece data collection programs for these statistical agencies are the Census of Agriculture and the ARMS. This chapter also identifies the missions and mandates of the statistical agencies along with other user needs that dictate the scope of data collection.

Chapter 3 explores the factors that contribute to a farm operation's complexity, which include the following:

- Farm size,
- Geographic dispersion of operations,
- Multifarm, multibusiness operations,
- Farm-connected nonfarm output,
- Employment/labor structure,
- Ownership/legal structure, and
- Management and decision making structure.

Here, the factors that make farms more or less complex as businesses (in an operational and management sense) are identified, and how such complexity affects data collection is also discussed.

The purpose of Chapter 4 is to clarify definitions and concepts used in the measurement of the farm economy, particularly for complex operations within the sector. The key units underlying production—the *farm,* the *farmer*, the *farm household,* the *farm business*—are disentangled, and currently used definitions, which are summarized in Chapter 2, are compared with alternatives that could help sharpen the measurement of farming and broader agricultural production within the statistical system. The fact that the scope of modern agriculture is considerably more far-reaching than what takes place on the farm is a driver of complexity in data collection, and this is a theme that arises at several points throughout the report. The discussion in Chapter 4 provides a foundation for understanding proposed changes in measurement concepts and practices discussed later in the report.

Chapter 5 examines the implications for data collection of the increasing complexity of farm business structures. Shaped by the question, "what information is needed to estimate farm

(or, more broadly, agricultural) activity in the U.S. economy?" consideration is given to identifying the appropriate statistical units for measuring complex farm operations. Here, there are three productive inputs that can come into play: the farm operation (the business), the people (individuals and households), and the land. Ideally, these measurement units would be defined in a way that is as consistent as possible with the broader system of business statistics, including the definitions developed by the Census Bureau and the Bureau of Labor Statistics. Identifying best practices for measuring the activities of complex farm businesses—such as multi-unit operations and operations that are vertically integrated—and their associated households requires appropriately specifying the Farm Register, due to its central role in collecting data to ensure sufficient population coverage.

Chapter 5 further addresses a series of methodological questions: (i) Is the universe of "principal farm operator" households,[21] which is the one currently used, the most appropriate one for representing complex farm operator arrangements? (ii) Which alternative concepts and measures can feasibly be incorporated into existing statistical programs? and (iii) Would farm classifications of sales and operator types different from those used by ERS better characterize the heterogeneity of the farm population, and would other classifications for reporting farm household income and wealth provide a richer or more accurate picture of this population?

Alternative approaches to accounting for farm production and finances are explored, and recommendations are offered for sharpening the conceptual bases of agricultural statistics, particularly for complex farm operations. Practical guidance is provided for improving key data collection instruments such as the ARMS and the Census of Agriculture. One important consideration is how changes in data collection or statistical methodology might affect estimates, such as of the number of farms. In some cases, it may make sense to maintain more than one estimate for a period, so that a bridge can be constructed to transition from an old to a new statistical series. And any changes will have to be justified by their contribution to increased accuracy, reduced cost or burden, or both.

In Chapter 6, data collection options for measuring complex farm operations are discussed, and the potential roles of alternative survey approaches, as well as nonsurvey data sources, are considered. Expanded use of administrative data sources and effective linking across data sources (including nonsurvey data sources) are key elements of the strategy to modernize the data infrastructure within the statistical system. The essential perspective of data providers is considered in recommendations addressing respondent burden, response rate, and data accuracy concerns.

In some states, a relatively small number of large, often organizationally complex farms now account for substantial shares of the production of some commodities and some vegetable crops. These large farms often bear a substantial respondent burden from being contacted regularly to participate in surveys, and partly as a result their overall response rates have fallen. New survey approaches are explored with the hope not only of reversing the decline in response rates but also of collecting data that is more accurate from those farms that do participate in NASS/ERS surveys. The recommended strategies for increasing the accuracy of information and reducing the respondent burden are made with a cognizance of both USDA's mandates and users' needs. For example, any survey methodology reform that delinks household information from farm information would preclude much current reporting and analysis by ERS, such as the

[21]For definitions of this and other terms used by NASS and ERS, as well as terms used in this report, see Appendix 2.1.

reporting on the distribution of farm payments, which bases household income information on farm operations commodity payments.

2
Background: USDA's National Agricultural Statistics Service and Economic Research Service

The National Agricultural Statistics Service (NASS) and Economic Research Service (ERS) of the U.S. Department of Agriculture (USDA) publish statistics and reports that regularly and extensively detail the number of farms in the United States, the quantities and types of outputs they produce, the incomes of both farm businesses and the farm households that run them, and the status and conditions of the agricultural economy. The USDA's data collection programs entail significant investments of the agency's staff time, staff talent, and budget resources.[1] Equally important, data collection is costly in terms of time and burden to survey respondents. To justify these investments, surveys and other data collection instruments must succeed at fulfilling a range of demands, from legislative and programmatic requirements to research, policy, and general user-community needs. In this chapter, we attempt to provide some appreciation of the complexity of the measurement tasks faced by NASS and ERS by describing in detail their current statistical programs and data infrastructure. A major component of this complexity involves conceptualizing in a consistent way the key productive units in farming activities that need to be measured.

As described in Chapter 1, agricultural production has been shifting to large farms in recent decades. And as this production has become concentrated among a relatively small number of large farms, the characteristics of farms and farming have become more complex in ownership structure and operational and management norms as well as the way farms are integrated with other sectors in the economy involved in food production and delivery. An example of this complexity is the way farms have shifted toward employing firms to provide services—for land preparation, planting, spraying, and post-harvest transporting—as inputs in production. Agricultural production therefore embodies not just what farmers do, but also the

[1]NASS and ERS budgets for FY2017 were $171 million and $87 million, respectively. Of course, the agencies engage in many other activities beyond data collection.

activities that they may hire out to separate businesses.[2] Farming activities are also contracted to nonfarm firms. In addition, large corporate nonfarm businesses may also be engaged in farming.

These trends have heightened the challenge of accurately characterizing the nation's farms and farmers, and their productive economic activities, and carry implications for data collection and statistical reporting by NASS and ERS. The way farms are defined in the data collection apparatus shapes the way information in the sampling frame for USDA surveys is updated and maintained and the way key individuals involved in farms and households are designated, including the determination of who the appropriate survey respondents should be. The capacity of NASS and ERS to accurately account for complex operations engaged in large-scale and often diverse activities directly affects the reliability of agricultural statistics.

This chapter reviews the information currently collected and published by NASS and ERS and details how and why data are collected and reported. The ways in which statistical products are used by researchers, policy makers, and farm owners and operators are also discussed.

2.1. MISSIONS AND MANDATES

NASS and ERS are two of the 13 principal statistical agencies of the federal government. With the stated mission to "provide timely, accurate, and useful statistics in service to U.S. agriculture," NASS

> conducts hundreds of surveys every year and prepares reports covering virtually every aspect of U.S. agriculture [including] production and supplies of food and fiber, prices paid and received by farmers, farm labor and wages, farm finances, chemical use, and changes in the demographics of U.S. producers [. . .]."[3]

Among the publications produced by NASS are six of USDA's seven leading economic indicator reports,[4] which are broadly used in agribusiness and market analyses, including for decision making by buyers and sellers of agricultural commodities.

The mission of ERS is to "anticipate trends and emerging issues in agriculture, food, the environment, and rural America and to conduct high-quality, objective economic research to inform and enhance public and private decision making."[5] The top-level subject areas covered by ERS include:[6]

- *Agricultural Economy* – farm-sector performance and farm households' well-being; farm size and concentration; market analysis, data, and projections on commodity supply, demand, and prices; and federal farm policies;

[2]Later in the chapter, we move beyond lay definitions and provide statistical definitions of terms such as "farm" and "agricultural production." Technically, agricultural production includes a list of activities/sectors specified in the NIPAs. See also the Glossary in Appendix 2.1.

[3]See https://www.nass.usda.gov/About_NASS/.

[4]Some of the most important of these are the *Grain Stocks, Cattle on Feed*, and *Hogs and Pigs* reports (Pruitt et al., 2014).

[5]See https://www.ers.usda.gov/about-ers/.

[6]These categories are reproduced as described by the agency: https://www.ers.usda.gov/about-ers/.

- *Food and Nutrition* – food security; food and nutrition assistance programs; food choices and health outcomes; food access and store proximity; food retailing and marketing; and food prices;
- *Food Safety* – societal benefits associated with reducing food safety risks; global trade implications and economic impacts of food hazards; and potential results of regulation versus industry decisions;
- *Global Markets and Trade* – domestic and international markets; trade; and the U.S. food and agriculture sector's performance in increasingly globalized markets;
- *Resources and Environment* – economic impacts of alternative conservation programs; efficacy of policies designed to protect the environment; challenges of climate change and water scarcity; and enhancing agricultural competitiveness through technology; and
- *Rural Economy* – investments in rural communities and the capacity of rural economies to prosper in a changing global marketplace; demographic change and its impact on rural communities; and drivers of rural economic performance.

As described in Chapter 1, the key users of the information produced by NASS and ERS are diverse, ranging from Congress, the White House, and federal, state, and local government agencies to agribusiness and other businesses (e.g., secondary food-related businesses), participants in commodity market transactions, researchers, industry groups, and the farmers and ranchers themselves.

Legal Mandates to NASS[7]

NASS collects data to meet multiple demands: to fulfill legislative mandates, to generate key inputs for principal economic indicators, to support administrative programs and strategic goals and projects, and to inform research and policy making. Most NASS surveys are not mandated, strictly speaking; rather, they are conducted under a delegation of authority from the President to fulfill USDA's mission to provide "leadership on food, agriculture, natural resources, rural development, nutrition, and related issues based on public policy, the best available science, and effective management."[8] Finally, although mandated information—such as data on income, finances, farm production, and households' well-being—drives only a small

[7]Much of the material in this subsection is distilled from a memo (*Information about NASS Mandatory Data Items*, June 2017), prepared by NASS staff for this study.

[8]The Code of Federal Regulations, Title 7, Subtitle A, Part 2, Subpart K, Section 2.68 describes the delegation of authority made by the Under Secretary for Research, Education, and Economics to the Administrator of NASS. Delegations of authority are in place to: (i) prepare crop and livestock estimates and administer reporting programs, including estimates of production, supply, price, and other aspects of the U.S. agricultural economy; (ii) collect statistics, conduct enumerative and objective measurement surveys, and construct and maintain sampling frames and related activities; and (iii) prepare reports of the Agricultural Statistics Board covering official state and national estimates. Additional provisions are in place to: ensure data security precautions to prevent disclosure of crop or livestock report information prior to the scheduled issuance time and to avoid disclosure of confidential data or information supplied by any person, firm, partnership, corporation, or association. Language in the delegation of authority provides further guidance for: improving statistics, maintaining coordination with OMB and other federal agencies on statistical methods and techniques; cooperating and working with national and international institutions and other persons throughout the world in the performance of agricultural research; and carrying out a number of administrative functions. For details of the NASS delegation of authority, consult https://www.law.cornell.edu/cfr/text/7/2.68.

portion of USDA's data collection, obtaining this information is nonetheless a crucial aspect of the NASS (and ERS) missions.[9]

The centerpiece of NASS's mandated responsibilities is to administer the Census of Agriculture, discussed in detail in the next section. This action is required by law under the Census of Agriculture Act of 1997,[10] which directed the Secretary of Agriculture to conduct a census of agriculture in 1998 and in every fifth year thereafter, covering the prior year.[11] In turn, anyone who receives the Census questionnaire is mandated by Title 7 of the U.S. Code to respond to the Census of Agriculture, even if they did not operate a farm in that year. In connection with the Census of Agriculture, the secretary may conduct any survey or other information collection and employ any sampling or other statistical method that he or she deems appropriate. The intent of the Census of Agriculture is to provide a complete count of U.S. farms and ranches and of the people who operate them at various levels of aggregation. The content of this census, however, is not specified in the law.

In addition to the Census of Agriculture, NASS is mandated to produce a series of reports for the following:[12]

- *Cold Storage*, which includes mandatory data collection items such as stocks of butter and stocks of cheddar cheese (Public Laws No. 106-532 and 107-171);[13]
- *Dairy Products*, which includes mandatory data on stocks of dry whey and stocks of nonfat dry milk (Public Laws No. 106-532 and 107-171);[14]
- *Cotton Ginnings*, which includes mandatory statistics and estimates of grades and staple length of cotton (13 U.S. Code § 41);[15]
- *Cotton Supply & Price Data*, which includes mandated information on market supply, demand, condition, and prices (Title 7 U.S. Code § 473b);[16]

[9]"Non-mandated" but still crucial aspects of data collection and statistical production, in terms of the agency fulfilling its mission, are detailed later in this chapter.

[10]Description of the law, Public Law 105-113, can be found at: https://www.gpo.gov/fdsys/pkg/PLAW-105publ113/pdf/PLAW-105publ113.pdf.

[11]Prior to 1997, the Census of Agriculture was conducted by the U.S. Census Bureau, often in conjunction with the Decennial Population Census—as was the case with the first Census of Agriculture in 1840. Later, the timing was adjusted so that the reference year would coincide with the economic censuses covering other sectors of the nation's economy. NASS publishes a detailed procedural history of the Census of Agriculture, which can be found at: https://www.agcensus.usda.gov/Publications/2012/Online_Resources/History/2012%20History%20Final%203.14.17.pdf.

[12]While NASS must produce these estimates, it is not mandatory for a selected operator to respond to most of these surveys.

[13]Monthly and annual *Cold Storage* reports can be found at: http://usda.mannlib.cornell.edu/usda/current/ColdStor/ColdStor-12-23-2016.pdf; and http://usda.mannlib.cornell.edu/usda/current/ColdStorSu/ColdStorSu-02-23-2016.pdf

[14]Monthly and annual *Dairy Products* reports can be found at: http://usda.mannlib.cornell.edu/usda/current/DairProd/DairProd-12-05-2016.pdf; http://usda.mannlib.cornell.edu/usda/current/DairProdSu/DairProdSu-04-28-2016.pdf.

[15]Monthly and annual publications can be found at: http://usda.mannlib.cornell.edu/usda/current/CottGinn/CottGinn-01-12-2017.pdf; and http://usda.mannlib.cornell.edu/usda/current/CottGinnSu/CottGinnSu-05-10-2016.pdf.

[16]Supply data are compiled in the Crop Production Report (http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-12-2017.pdf); price data are compiled in the Agricultural Prices Report (http://usda.mannlib.cornell.edu/usda/current/AgriPric/AgriPric-12-29-2016.pdf).

- *Cotton Acreage*, with mandated information on the total estimated acreage of cotton planted and on estimated harvested acreage (Title 7 U.S. Code § 476);[17]
- *Peanut Processing*, with mandated information on a wide range of statistics pertaining to peanuts and peanut-based products (Title 7 U.S. Code § 951);[18]
- *Prices Received*, with mandatory information on corn, wheat, and cotton prices (Title 7, Chapter 35A, Subchapter II, Section 1441);[19] and
- *Cash Rents*, with a mandated survey—conducted "not less frequently than once every other year" (Agricultural Act of 2014, Title II, Section 2005)—of per-acre estimates of county average market dryland and cash rental rates for irrigated cropland and pastureland in all counties or equivalent subdivisions within each state that have 20,000 acres or more of cropland and pastureland (Food Security Act of 1985, Public Law 99-198, 99 Stat. 1504, amended through Public Law 113-75; Section 1234(C)5b).[20]

Additionally, the Food, Agriculture, Conservation, and Trade Act of 1990 and the Food Quality Protection Act of 1996 require NASS to collect and publish annual data on field crop chemical use.[21]

Legal Mandates to ERS and for Its Agricultural Resource Management Survey

USDA is required by Congress, through both authorizing and appropriations legislation, to produce statistics on a range of topics, many of which are estimated using data collected through the Agricultural Resource Management Survey (ARMS), which is jointly conducted by NASS and ERS. ARMS is the "primary source of information on the financial condition, production practices, and resource use of America's farm businesses and the economic well-being of America's farm households."[22]

The Food and Agriculture Act of 1977 mandated that ERS, through the Secretary of Agriculture, report annually on trends in family farms—which ERS met primarily by collecting data from *principal farm operator households* in ARMS—and produce comprehensive national and state-by-state data on nonfamily farm operations. Although that mandate expired, recent Family Farm Reports are still produced on a periodic basis (they have been published annually since 2014).[23] The reports draw on ARMS data to illuminate a host of relationships, including: (i) farm participation in agricultural programs and the distribution of farm program payments; (ii) the structure and organization of farms, including family and nonfamily ownership; (iii) the use of new production technologies and other management practices; (iv) farm use of credit; (v) farmers' participation in off-farm employment; and (vi) the characteristics of producers

[17]Publications are *Prospective Plantings* (http://usda.mannlib.cornell.edu/usda/current/ProsPlan/ProsPlan-03-31-2016.pdf); *Acreage* (http://usda.mannlib.cornell.edu/usda/current/Acre/Acre-06-30-2016.pdf), and the *Crop Production Repor*t (http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-12-2017.pdf).

[18]See: http://usda.mannlib.cornell.edu/usda/current/PeanStocPr/PeanStocPr-12-30-2016.pdf.

[19]See: http://usda.mannlib.cornell.edu/usda/current/AgriPric/AgriPric-05-31-2017.pdf.

[20]See: https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Cash_Rents_by_County/.

[21]The Agricultural Chemical Usage Field Crops Summary (http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1560), and The Agricultural Chemical Usage Restricted Use Pesticide Summary (http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1572).

[22]See https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/.

[23]See https://www.ers.usda.gov/topics/farm-economy/farm-structure-and-organization/readings/.

purchasing crop insurance (NRC, 2008, p. 17). Income estimates are designed to be consistent with the household income definitions used in the reporting of all U.S. households for most Census Bureau data series.

USDA's annual sector estimates and forecasts of net cash farm income and net farm income are in turn used by the Bureau of Economic Analysis (BEA) in the construction of its national, regional, and industry economic accounts. The USDA income statistics include detailed data on value added, cash receipts and value of production, government payments, and farm production expenses. The statistics are based on responses to ARMS as well as sector-level information provided by NASS, the Farm Service Agency, the Risk Management Agency, and other administrative data sources. Data for more than 200 components of farm income are provided to BEA and used in deriving farm sector GDP and personal income both for the United States and by state. In addition, the primary source for BEA's estimates of farm output—as used in their input-output accounts—is "cash receipts from farm marketings" by commodity as produced by USDA.

ERS is also mandated to publish cost-of-production information for a number of commodities. U.S. Code states that the

> Secretary of Agriculture, in cooperation with the land grant colleges, commodity organizations, general farm organizations, and individual farmers, shall conduct a cost of production study of the wheat, feed grain, cotton, and dairy commodities under the various production practices and establish a current national weighted average cost of production. This study shall be updated annually and shall include all typical variable costs, including interest costs, a return on fixed costs, and a return for management (U.S. Code, Title 7).

ARMS data are the key input into the annual cost-of-production estimates, and also provide baseline estimates for the years in which specific commodities are targeted.

Production input data collected through ARMS are also used to generate annual weights for the Prices Paid by Farmers Index, computed by NASS. This index, which indicates the average costs of inputs purchased by farmers and ranchers to produce agricultural commodities, is mandated by the 1933 Agricultural Adjustment Act. "Parity prices," generally calculated as national averages, are used in administering federal marketing orders for 45 categories of fruits, vegetables, and nuts. The 1978 Public Range Improvement Act stipulates that these price indexes are also to be used by the U.S. Bureau of Land Management and the U.S. Forest Service in the calculation of annual federal grazing fees on the nation's western public lands (NRC, 2008, p. 18).

How Legislative Mandates Shape Data Collection Strategies at NASS and ERS

As described above, most data collections are conducted by NASS and ERS under a delegation of authority from the President—which stipulates broadly the kinds of information needed for USDA missions—or to meet important research, policy, or general stakeholder demands for information. Relatively little data collection is driven directly by and with specific instructions from legislative mandates.

The legislative mandates that NASS and ERS are required to fulfill typically specify the types of information required but not how the information should be generated.[24] This means the

[24]This is the case for most of the principal statistical agencies in the federal government.

agencies have considerable latitude in how they fulfill these mandates. In many cases, USDA is required to collect a particular datum, such as on pesticide use, and to publish statistics about it, but there is no stipulation that it be collected through a particular kind of survey or by using a particular source of administrative information. For example, the mandate to "report on the financial health of the farm sector" could be handled in a number of different ways. Currently, ARMS data, including the production expense and the "farm-related income" line items, are used. However, administrative data are also used—from programs involving government payments (Farm Service Agency), commodity insurance indemnity payments (Risk Management Agency), commodity loans (Commodity Credit Corporation), and cash receipts (NASS production and price data sourced from Quick Stats).

Likewise, Congress mandates that data be produced for national income accounts for the sector, such as on household incomes and commodity costs, but the mandates do not specify that this information must be obtained using surveys. USDA's interpretation is that ARMS is the best mechanism for fulfilling the core of this mandate. Although ERS relies on a single year of ARMS data to set the baseline cost and returns measure, subsequent annual updates are made using non-ARMS data on input prices by commodity. And, because ERS is mandated to report income received by the entire sector, data must be collected from all types of operations, including complex operations.

In other cases, Congressional mandates directed to USDA do not specify the "type of information" to be produced, but only that the collection take place. Most notably, the content of the Census of Agriculture is not specified in the law; the mandate is only that this census be conducted.

> **Conclusion 2-1**: The mandated responsibilities faced by NASS and ERS allow for considerable latitude in how data are collected from farms and how statistics on their activities and finances are produced. For this reason, the panel's recommendations for improving or streamlining the Census of Agriculture and ARMS (see Chapter 5)—in terms of content, questionnaire structure, burden reduction, and design elements to facilitate easier and more accurate linkages with nonsurvey sources—would not hinder NASS and ERS from fulfilling their mandates. Meanwhile, the recommended actions are intended to improve the agencies' capacities to fulfill their missions to provide the data and statistics needed for policy, research, and other stakeholder requirements.

USDA also has considerable flexibility to explore nonsurvey sources of data, an approach that has improved a number of its data products. Mandates to NASS and ERS generally do not constrain the use of multiple kinds of data (administrative, commercial, web-based, etc.) that could complement or, possibly, substitute for elements of the current survey-centric approach. Indeed, given the types of information about the sector that have value, expanding the breadth and diversity of the data sources drawn from represents a natural evolution for statistical agencies. Across these agencies, the trend in fulfilling information demands is toward greater use of administrative data. Canadian and European statistical offices are leaders in this regard (Prewitt, 2010, pp. 11-12; NASEM, 2017a). At the same time, agencies within USDA have also demonstrated the value of using administrative data for statistical purposes, while also documenting the difficulties of gaining access to such data (Prell et al., 2009), as is discussed in detail in Chapter 6.

Finally, although the legislative mandates to NASS and ERS are not especially constraining, ideally their content would be revisited and evaluated for relevance to the contemporary agriculture sector. In some cases, the underlying rationale for certain mandates may no longer apply, which means it is possible that resources are being unnecessarily diverted away from more pressing information needs. For example, some mandates have their origins in the Great Depression, when most rural households earned their livelihoods through agriculture and when there was a large gap between rural and urban household incomes. Neither of those conditions holds true today for the majority of farm households.

2.2. DATA NEEDS AND CURRENT REPORTING

Any consideration of changes to data collection should take stock of the reporting requirements faced by the USDA. Articulating data needs can be challenging: almost any piece of data could be described by a user or group of users as "needed" or "essential," and many kinds of data will have value to some individual or organization. However, data collection carries with it a public cost, and it should therefore create public benefits. Here, and throughout this report, we highlight the types of USDA data for which there is a strong justification that they be publicly provided. Key data uses include the following:

1. *Program administration*: Data are needed to enable agencies to run programs as effectively and efficiently as possible. Example: Understanding how crop insurance premiums affect crop insurance enrollment.[25]
2. *Policy analysis*: Data are needed to evaluate whether policies and programs are affecting the right people and having the desired effects. Data are needed to evaluate how policies and programs operate in practice, including how the distribution of impacts across different groups is affected. Example: Identifying what types of farms receive the most farm payments.[26]
3. *Research on agriculture, health, food, and environmental concerns*: Data can be used to simulate how the promotion of particular farm practices can affect environmental quality in a region. Example: Managing the costs of reducing agriculture's footprint in the Chesapeake Bay.[27]
4. *Informing markets*: Information is essential for improving the workings of markets. Example: Reports on planting intentions for particular crops.[28]

Some kinds of data—such as price data—may be best collected from markets themselves rather than from producers. Even in these cases, USDA can play an important role as an aggregator or clearing house for such data.

Reports and data products from NASS and ERS often serve one or more of the above purposes, and this report does not attempt to classify them. Instead, here, we simply provide a brief description of what and how the two agencies report data or analysis.[29]

[25]See https://www.ers.usda.gov/webdocs/publications/45227/48299_err169.pdf?v=41827.

[26]See https://www.ers.usda.gov/publications/pub-details/?pubid=44653.

[27]See https://www.ers.usda.gov/amber-waves/2014/july/managing-the-costs-of-reducing-agriculture-s-footprint-on-the-chesapeake-bay/.

[28]See http://usda.mannlib.cornell.edu/usda/current/ProsPlan/ProsPlan-03-31-2017.pdf.

[29]Detailed documentation of data products and reports is available from both NASS. (https://www.nass.usda.gov/Publications/catalog.pdf) and ERS (https://www.ers.usda.gov/data-products/).

NASS reports typically present totals (such as total chemical expenditures) and ratios of totals (such as total chemical expenditures / total number of farms) for the United States and for regions, states, and counties, as well as across types of farms. Its reports generally do not include distributional statistics, such as percentiles, or relationships between variables.

Once every five years, NASS publishes tables of statistics based on the Census of Agriculture. Annual surveys provide the basis for annual reports on topics such as farm production expenses (from the ARMS), agricultural chemical usage (from the Chemical Use Survey/ARMS), and cash rents and agricultural land values (from the June Area Survey). Other reports occur more frequently. The quarterly Farm Labor Survey provides estimates of employment and wages for workers employed directly by farms and ranches. It serves as the basis for the semiannual *Farm Labor* report.

Monthly or quarterly series provide timely information on production and prices, and some of these are considered principal economic indicators for various programs within the agency. A principal federal economic indicator is defined as a major statistical series that describes the current condition of the economy.[30] Examples include the monthly *Cattle on Feed* report and the monthly *Crop Production* report. Two of the principal economic indicators reports only occur annually but provide prospective information on acreage planted or harvested, namely the *Prospective Plantings* report and the *Acreage* report.

ERS provides information and analysis that often goes beyond aggregate statistics such as production or expense totals. These include statistics to understand central tendencies, such as the typical number of hours worked off-farm, and diversity across groups, such as commodity program participation by demographic group. Analysis often involves farm-level relationships between variables, such as how participation in crop insurance relates to access to credit. New questions are added to surveys in the effort to anticipate new trends, technologies, and policies.

ERS does much of its reporting by publishing data products, often as a time series. Examples include statistics on farm household characteristics and finances, on farm business finances, on commodity costs and returns estimates, and on the adoption of genetically engineered crops. Context for these statistics is provided through topic web pages associated with particular data products.

Another form of reporting is through agency reports that provide analyses of data, including tables and statistical results. Reports such as the *Family Farm Report* are repeated annually or every few years, but most reports are unique and emerge from a combination of stakeholder and researcher interest, data availability, and analysis of new policies, technology, or trends. Recent reports include the following:

- *Changing Farm Structure and the Distribution of Farm Payments and Federal Crop Insurance* (2012)
- *USDA Microloans for Farmers: Participation Patterns and Effects of Outreach* (2016)
- *Farm Household Income Volatility: An Analysis Using Panel Data from a National Survey* (2017)

[30]Economic indicators are compiled, released, and periodically evaluated in accordance with procedures established in OMB Statistical Policy Directive No. 3. NASS provides OMB with its schedule of Principal Economic Indicator releases for the upcoming calendar year. If unforeseen circumstances make it necessary to change any scheduled release date after OMB issues the schedule, the agency must announce and explain the change as soon as it is known.

- *Federal Crop Insurance Options for Upland Cotton Farmers and Their Revenue Effects* (2016)
- *Changing Structure, Financial Risks, and Government Policy for the U.S. Dairy Industry* (2016)

In addition to providing statistics through reports, both NASS and ERS allow web users to query data for particular geographies or subject types. The NASS Quick Stats tool[31] provides totals or ratios of totals from the Census of Agriculture or its various surveys. To give an example, users can use the online tool to access the corn acreage for a particular county and year or download corn acreage for all counties and years. ERS has a similar query tool for the ARMS data on farm structure and finances or crop production practices.[32] Through these web tools, users can find results for particular states or types of farms and view trends over time.

ERS and NASS provide an additional service to the research community by granting access (to individuals who clear an approval process) to ARMS microdata through a data enclave system, and to the Census of Agriculture and other surveys through NASS Data Labs. Through data-use agreements, researchers outside the USDA can access farm-level data (that clears disclosure and confidentiality requirements) to perform analyses for reports or academic articles. This access has made possible hundreds of peer-reviewed academic publications on diverse topics ranging from agro-environmental issues to farm finances to issues for beginning farmers.[33]

2.3. CURRENT DATA INFRASTRUCTURE AT NASS AND ERS

The primary data collection from farms is administered by NASS on behalf of USDA. NASS collaborates with ERS on the content and other design features for some surveys, including ARMS. As discussed above, the collection of farm data occurs both through the Census of Agriculture, which is conducted every five years, and the annual ARMS. The two instruments provide information widely used in reports by both NASS and ERS.

Census of Agriculture

The Census of Agriculture is a complete count of U.S. farms and ranches and the people who operate them. It is conducted once every five years by NASS to collect information on land use and ownership, operator characteristics, production practices, income, and expenditures. It provides the only source of uniform, comprehensive agricultural data for every county in the nation through time, showing historical changes in U.S. agriculture and long-term trends. The first agricultural Census was taken in 1840 as part of the Sixth Decennial Census of Population. For 156 years (1840–1996), the Bureau of the Census (and its predecessor, the Census Office) was responsible for collecting data for the Census of Agriculture. In 1997, responsibility for conducting this census was transferred to USDA.

The Census of Agriculture remained part of the Decennial Census through 1950, with separate mid-decade Censuses of Agriculture taken in 1925, 1935, and 1945. As time passed, the

[31]https://www.nass.usda.gov/Quick_Stats/

[32]https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/arms-data/

[33]Moss et al. (2012) found that a Google Scholar search (May 17, 2012) produced 1,290 documents for the period 2011 to 2012 using the terms "USDA ARMS data." Using "anytime" as the period of time for the search, 18,200 documents appeared.

Census of Agriculture years were adjusted until the reference year coincided with the Economic Censuses covering other sectors of the nation's economy. Currently, the Census of Agriculture is conducted for years ending in "2" and "7."

The Census of Agriculture also collects information on the agricultural industry that may not be gathered elsewhere in the annual survey programs conducted by NASS, covering topics such as agritourism, organic production, farmer demographics, specialized agricultural production, Internet access, and more. However, the Census of Agriculture does not provide balance sheet information or information on household well-being. Several surveys that NASS has conducted following the Census of Agriculture cover on-farm energy production, farm and ranch irrigation, organic production, horticultural specialties, and local foods.

The Census of Agriculture defines a farm as a place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the census year. Discovering new farms and properly accounting for existing and continuing farms and operators is an ongoing challenge. NASS accomplishes this through use of a list frame that covers the population of all farms and farm operators known to NASS and through an area frame based on land segments. It collects data from units in both sampling frames, targeting the full population, for the Census of Agriculture. For surveys, the sample depends on the target population, the commodity of interest, and so on.

In the list sampling frame that NASS uses, the sampling units are operations. Operators, defined as those who run farms—that is, make day-to-day management decisions—may receive a questionnaire for each operation they are involved in. Since the Census of Agriculture produces county estimates, such as for livestock or crop production, it also needs to attribute the agricultural production and the land of an operation to each county.

ARMS and Other Key Surveys[34]

ARMS is an annual cross-sectional survey that collects information on farms and farm households. The national survey[35] is unique in that it collects, in a representative sample, (i) observations of field-level farm practices, (ii) information on the farm business, and (iii) characteristics of the household operating the farm. Responses to the survey are meant to provide estimates that are representative and reliable both at the national level and, for key states with the highest value of agricultural production, at the state level. Every year, farms producing a commodity (or commodities) of interest are oversampled and targeted with a commodity-specific version of the questionnaire.

ARMS occurs in phases, with initial screening of sampled farms occurring in Phase I. In Phase II, which samples roughly 4,000 to 10,000 farms, the survey collects information about a particular field, only from farms producing the crop being targeted in the survey year. The phase is dedicated to a detailed look at the production practices associated with the targeted commodity.

In Phase III, for which roughly 35,000 farms are sampled each year, the survey collects farm financial and household information, as well as additional information on production practices, from *all* sampled farms. Sampled farms that are producing the targeted commodity receive a commodity-specific Phase III questionnaire, which asks questions unique to the

[34]The discussion here is based on the ARMS documentation page: https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/documentation/

[35]Although Alaska and Hawaii are not covered.

commodity and permits a detailed assessment of the costs and returns associated with it. Sampled farms not producing the targeted commodity receive a more general version of the questionnaire.

The sampling unit for ARMS is the operation-operator pair from the sampling frame. Two operators of the same farm will not each receive an ARMS survey in the same year. Rather, only one operation-operator pair will be selected for a given operation. In addition, NASS employs techniques intended to help ensure that each operator only receives one ARMS questionnaire, even if she (or he) operates multiple farms. This contrasts with the design of the Census of Agriculture, in which a single operator may receive an additional questionnaire for each operation with which she (or he) is associated.

In years when the Census of Agriculture is carried out, ERS uses an integrated ARMS/Census questionnaire so that operators selected for ARMS fulfill their Census obligation by completing ARMS alone. In comparison to the Census of Agriculture, ARMS has questions that are more comprehensive, including balance-sheet information and farm-household income information.

NASS uses a stratified sample design for ARMS, in which the strata are defined by various farm characteristics such as commodity, farm sales class, and state. Larger farms are generally oversampled and small farms are undersampled. The oversampling of large farms can result in the same farm being sampled for ARMS two, three, or more times over a decade. Between 2000 and 2013, 16 percent of all sampled farms received the ARMS questionnaire more than once (Weber, Key, and O'Donoghue, 2016).

Other major annual surveys conducted by NASS include the June Area Survey and the Chemical Use Survey (there are many others as well). The June Area Survey collects information about agricultural land use, value, and rental rates in sampled land units, with the sampling unit typically being one square mile of land. Information is collected for all the land area in the sampled unit, with questionnaires filled out for each farm operating land in the sampled unit. For field crops, the Chemical Use Survey is embedded in the commodity-specific ARMS Phase II survey (fruit and vegetable chemical use are surveyed outside of ARMS II); questions appear as part of "Production Practices, Costs and Returns" (PPCR) field crop questionnaires administered by ERS and NASS. Information is collected regarding on-farm chemical use and pest management, including the area treated and rates of application of fertilizers and pesticides.

Key Concepts and Definitions Guiding Data Collection[36]

To produce meaningful information about the farming sector, there must be a common conceptualization of the basic measurement units among survey respondents and within statistical agencies. To this end, USDA has developed a set of related definitions for a *farm,* a *family farm,* a *farm operator,* and a *farm household*. In addition, the sampling unit can be the operation, operator, field, or a combination of them, depending on the survey. Finally, the farm operation may also be involved with a variety of on-the-farm, off-the-farm, and value-added activities.

[36]Definitions in this section reflect current usage by NASS and ERS. Officially used definitions for many of the terms referenced here can be found in the USDA glossary: https://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/glossary.aspx.] Additional discussion of key terms continues in chapters 4 and 5 where, when noted, usage may differ.

"Farm"

As mentioned earlier, the USDA defines a farm as "any place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the year." According to O'Donoghue et al. (2009), this definition was established by the Office of Management and Budget (OMB) and the Department of Commerce in 1975, suggesting that it is an administrative decision, not a definition set in statute. Note that the $1,000 value has not been adjusted for inflation over the subsequent period. Places with less than $1,000 in sales may also in some cases be considered farms based on their potential for sales, estimated using a points system that accounts for cropland or livestock assets. The idea is to continue to count these operations even if actual sales temporarily failed to meet the sales threshold due to bad weather, death of a household member, change in marketing strategies, and so on. Payments from government programs like the Conservation Reserve Program are considered sales when determining whether a place constitutes a farm.

There are practical (and political) reasons for maintaining the current definition of a farm in USDA publications. One of many examples is that 1890 universities (universities designated with land-grant status under the Morrill Act of 1890) receive funding partly based on the number of farms in a particular state. However, for collecting data with the purpose of measuring the activity and output of the nation's farms, there are compelling reasons to identify farm businesses in alternative ways as well, based on how those businesses are organized. These alternatives are discussed in detail in chapters 4 and 5. Questionnaires employed by the USDA use the word "operation" instead of "farm"—NASS generally treats the two terms as synonymous. The number of operations enumerated in the Census of Agriculture, for example, determines the USDA's estimate of the total number of farms in the United States and is highly dependent on the $1,000 sales (or potential sales) definition. ARMS and the Census of Agriculture collect information from the same universe of operations. However, a key difference is that ARMS focuses on an operation-operator pair, because it seeks information about a specific household associated with the operation. In practice, this means that a person managing three farms would receive three Census of Agriculture questionnaires (in the Census year) but only one ARMS questionnaire (if selected for ARMS in a given year), which would ask about the person's household and about one of the farms that he or she operates.

How one operation is delineated from another is largely determined by the respondent. If a respondent's notion of an operation meets the definition of a farm, it is considered a farm and enumerated as such. For example, a respondent could own and operate several poultry houses and also plant corn and soybeans. That respondent might choose to report the crop and livestock activities as one operation, or else as two. If he (or she) considers them two operations, the respondent would only be required to report one of them for ARMS. The Census of Agriculture is stronger in its guidance. The Report Form Guide for the 2012 Census of Agriculture (as cited by MacDonald et al., 2018) instructs respondents to complete a separate report form for each distinct agricultural operation (farm, ranch, feedlot, greenhouse, etc.) for which separate records of operating expenses and sales, livestock, and crop acreage and production are normally maintained.

Certain farm-related activities and their economic quantities (that is, their assets or income) may be reported as part of the operation if they are not part of a separate business. In some cases—such as when activities underlying the income generation use farm assets and create farm costs, such that they are joint products—this is at least partly justifiable. In practice,

activities are treated as separate businesses if they are separate from an accounting perspective—that is, if they maintain separate financial records.

"Family farm"

A family farm is defined by the USDA as a farm in which "the majority of the business is owned by the operator and individuals related to the operator by blood, marriage, or adoption, including relatives that do not live in the operator household."[37] The business referenced is the same as the "operation" as defined in the questionnaire; the operator is the principal operator of the farm (as defined below). Thus, a farm is a family farm if more than 50 percent of the farm's assets are owned by the principal operator and any people related to him or her by blood, marriage, or adoption. This definition is similar to the definition of a family-owned business used by the Census Bureau in its Economic Census.[38]

"Farm operator(s)" and "principal farm operator"

The farm operator is defined by the USDA as the "person who runs the farm, making the day-to-day management decisions."[39] The respondent determines who is an operator, and a farm may have multiple operators, including a hired manager or a partner(s). The only restriction on who is considered an operator is that the person must make day-to-day decisions for the operation. As such, anyone who considers themselves to be consistently involved in management of the farm could report themselves as operators. Operators are not required, by definition, to have an ownership stake in the farm or to spend a certain number of hours working on the farm.

If the operation has multiple operators, the respondent is asked to identify a *principal* operator; in other words, it is left to the respondent to define the principal operator. For surveys such as ARMS, ERS uses information on the principal operator to identify the household about which it will collect household demographic and financial information. In contrast, the 2017 Census of Agriculture has moved away from the principal operator concept, instead identifying up to four operators, now called "persons" involved in decisions. The 2017 ARMS also identifies up to four persons involved in specific decisions in the operation, but continues to identify a person who is most responsible for decisions as the principal operator and collects data for the associated household. A recent report recommended folding farm owners and decision makers into one group under the term *producer*—a recommendation that was adopted for the 2017 Census of Agriculture—to indicate any person involved in the business's governance structure.[40]

[37]https://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/glossary/.

[38]The question used in the Economic Census to define a family business was: "In 2012, did two or more members of the one family own the majority of this business? *(Family refers to spouses, parents/guardians, children, siblings, or close relatives.)*"

[39]See https://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/glossary/.

[40]See *Publication of Agriculture Census Data on Farm Operator Demographics* (a report by the National Institute of Statistical Sciences Technical Expert Panel, October 12, 2017). There it was recommended to replace the label "Operator" with "Producer" in all publications. The 2017 Census of Agriculture and future censuses use these terms: "All producers," "Principal Producers," "Non-Principal Producers." These terms span the breadth of agriculture and are seen as consistent with current terminology used by producers and by professional agricultural organizations.

"Farm household"

USDA collects and reports information on the household of the principal operator of a family farm. It defines farm operator households as those who share dwelling units with the principal farm operators of family farms. According to ERS, multiple operators who do not share the same household operate less than 10 percent of family farms. Using this definition, the farm operator household population would include the households of the principal farm operators, but not the households of the other operators.

Household income for the principal operator refers to all income earned by members of the principal operator's household, including both farm and off-farm income. Farm income includes net income from the farm operation that accrues to the principal operator household as well as farm-related income, such as income from renting out land to other farms. Nonfarm income is all other income, include that from wages, nonfarm businesses, and interest and dividends received by any member of the principal operator household.

"Farm business"

ERS defines farm businesses as farms with annual gross cash farm income greater than $350,000, along with smaller operations where the principal operator has farming as his or her primary occupation. This categorization is mostly used in reports and statistics as opposed to survey instruments—that is, it is not a cut-off line used to guide data collection in surveys. This distinction is descriptively important given the skewed size distribution of farms, as discussed in Chapter 1. In its reporting of farm business income, ERS estimates that "farm businesses" represent less than half of U.S. farms but "contribute over 90 percent of the farm sector's value of production and hold the majority of its assets and debt."[41] The farm business concept provides an example of how ERS has latitude in what it reports to fulfill mandates.

"Field"

NASS defines a field as a continuous area of land devoted to one crop or a single land use, such as farmstead, pastureland, woods, or wasteland. For some data that are collected, the sampling unit within the operation can be a field. In ARMS Phase II, the respondent is asked to list all fields with the target commodity, and then the enumerator selects one field randomly for the respondent to complete the questionnaire. Because of the design of ARMS, information collected about the field is linked to an operation and operator.

"**Value-added activities and products"**

This phrase refers to the manufacturing processes that change the physical state or form of the product(s), increasing the value of the primary agricultural commodities produced on the farm, and to the final product(s) of those processes. As examples, USDA defines value-added products to include beef jerky, fruit jams, jelly, preserves, and floral arrangements. In recent years, both the Census of Agriculture and ARMS have identified farms participating in value-added activities and collected information on the associated income or value of sales.

If a value-added activity, such as cheese making or grape processing, is not a distinct business, meaning that it is inseparable from the farm from an accounting perspective, then its economic values (assets, debt, costs, revenues) are typically included in those of the farm being

[41]See https://www.ers.usda.gov/topics/farm-economy/farm-sector-income-finances/farm-business-income/.

enumerated.[42] If the value-added activity is a separate business, then only the income from it would be recorded in ARMS, where it would be recorded in the household section as income that the household earned from a nonfarm business. USDA reports how the operator has decided to run the value-added activities and whether the operator reports them as part of the farm or not. Typically, when these manufacturing or transportation value-added activities become sufficiently large they would be reported as a separate business, allowing for proper classification of value-added products.

From the National Income Product Accounts (NIPA) perspective, where the goal is to appropriately measure GDP as the value of final products consumed, it may not matter whether value-added activities are classified as part of the farm or not, as long as there is no double- (or triple-, etc.) counting in the intermediate stages. For example, when a farmer produces wheat that is used by a miller to produce flour that is used by a baker to produce bread that is sold to a consumer, it is important to not double-count the value of the wheat and the value of the flour if milling is not part of the farm. It is worth noting that ERS also uses the Value Added Component Series, where the concept of value added has a different use, to denote the contribution of farm production (as opposed to food processing, packaging, food service, etc.) in the overall value added of all establishments that contribute to total food dollar purchases.

However, proper classification of raw farm product versus value-added product is still important in the North American Industry Classification System (NAICS). For example, grape vineyards (NAICS code 111332) are classified as agriculture (NAICS code 11), whereas wineries (NAICS code 31213) are classified as manufacturing (NAICS codes 31-33). To the extent that a farm does not report its wine-making activities and income as a separate winery business, those activities will be comingled with the farm, which may then be classified as a grape vineyard. Arbitrary reporting of value-added products as farm production is a central measurement problem, whether the goal is to measure the size of the agricultural sector, broadly defined, or to specifically measure "farm activities." Note that this boundary issue is dealt with throughout the remainder of this report.

Challenges Raised by Current Data Collection Practices

The concepts and definitions that USDA uses, described in the previous section, are helpful in providing guidance to survey respondents. Yet some definitions give considerable flexibility to the respondent on how to report data, and this has implications for data accuracy, interpretability, and respondent burden.

Several key concepts on which data collection are based are vague and left to the interpretation of respondents, including "the operation," "the principal operator," and a "separate business." Whether respondents consider (and report) their agricultural activities to belong to one farm or more than one is currently at their discretion. Respondents are also given little guidance on how to identify the principal operator, particularly in cases where multiple people each have primary responsibility over distinct aspects of the farm, such as management of marketing and management of crop production, as well as in cases of farms operated with spouses or in

[42]In some cases, an effort is made to distinguish between the farm output and the value added from nonfarm production. For example, NASS's 2015 Certified Organic Survey asks respondents reporting sales of roasted soy nuts to also estimate the value if they had instead sold the product as raw, unprocessed soybeans. In Section 11 (Item 4) of the survey, respondents are asked to record this value along with the Gross Certified Organic Value-Added Sales.

partnerships. Although guidance is provided to the portion of the sample visited by enumerators, this vagueness could create confusion for respondents, particularly in cases where there are multiple operators, managers, and owners, as is increasingly true of modern farming businesses. Interpreting the data that respondents provide also depends on assumptions about how respondents have understood the terms used in a survey.

Vagueness in the definition of a separate business is particularly important, as it influences whether various economic values are properly reported as part of the farm sector or not. As mentioned above, once value-added activities are identified as comprising a separate business they are excluded from outputs reported for the farm (while still included for the household).

Among the measurement goals of NASS and ERS is to obtain an accurate measure of gross farm income, including income from value-added activities. It is challenging to get information on how farm and nonfarm activities are linked into a single business (this is discussed in greater detail in later chapters). In other cases, a firm may be coordinating economic activity on farms and realizing a share of the value-added from their agriculture, but without operating any of the farms. In these latter cases, the question is, "Should the aim be to survey and measure data from all farm firms?"

USDA only reports on households containing the principal operator of a family farm. ARMS asks about the number of households that share in the net farm income from the farm but does not collect data on these additional households involved with the farm. All reported statistics, therefore, refer to the population of households containing a *principal* operator, not the full population of households containing any farm operator. Because a growing number of farms involve ownership, investment, and management by people from multiple households, the method of including only households with a principal operator will increasingly omit many households involved with a farm.

Moreover, it is unclear whether current practice best fulfills the mandate to report on the well-being of households of family farms in general. For example, as detailed in Chapter 5, redefining the farm household population as *all households that include a person who is an operator of a family farm* would permit capturing information about the financial health of households in a position to succeed a principal operator in the coming years.

Conclusion 2-2: When respondents are given a choice to decide the unit of measurement, such as which activities are included as part of the farm, who are the operators, or who is the principal operator, statistics on the number of farms, the size and scope of the farm sector, and the farm population are affected. Improving the clarity of definitions and requiring respondents to follow them would produce more accurate and interpretable estimates of the farm sector.

2.4. THE ESSENTIAL PERSPECTIVE OF DATA PROVIDERS: RESPONDENT BURDEN, RESPONSE RATES, AND DATA ACCURACY

An important aspect of assessing current practices is to consider how alterations would impact burden to respondents and, in turn, the accuracy of data and statistics produced by USDA. In this subsection, we consider each of these questions.

What is the Level and Distribution of the Respondent Burden on Farm Operators?

Respondent burden can best be defined by the *length* of a survey questionnaire, the amount of *effort* required by the respondent, the amount of *stress* on the respondent, and the *frequency* with which the respondent is interviewed. Length is usually measured by the total time it takes to complete the questionnaire. Effort refers to the ease with which questions can be answered, particularly the need to consult records and the degree to which records are kept in categories that match those asked about in the questionnaire. Stress refers to the sensitivity of the questions and the degree to which they may evoke emotional reactions in the respondent. *Frequency* of interviewing is determined by the design of the survey (Bradburn, 1978).

The gatekeeper for the federal statistical system, the agency determining if and when data collection can proceed, is OMB. OMB measures the total burden of a given survey as the average time required for a respondent to answer the survey questions multiplied by the total number of respondents. The amount of time required to answer a survey is affected not only by the length of the interview or questionnaire but also by the difficulty for the respondent in reporting the requested data (McCarthy, Beckler, and Qualey, 2006, p. 97).

Appendix 2.2, prepared for the panel by Hancock and Ott (2017), summarizes the annual total burden and the burden per contact for various surveys administered by NASS. The 2017 Census of Agriculture was estimated to have a total annual burden of 2,763,085 hours, which was distributed across 2.1 million farms (and, due to multiple operators in some cases, 3.2 million farmers). The figures in appendix table 2A.2 indicate that NASS data collection accounts for only a small portion of the total public burden generated by USDA agencies.[43]

Relative to the Census of Agriculture, the total respondent burden from fielding ARMS is much smaller because it is imposed on a far smaller portion of the farm operator population. Its survey sample consists of approximately 30,000 farms and ranches selected from NASS's list frame and area frame. However, the questionnaire contains more than 800 items for the respondent to potentially complete and for NASS to process after data collection; it is a very detailed survey.

In 2010, there were two primary versions of ARMS, known as the Core survey and the Cost and Returns Report. The Core survey was 16 pages long and took the average principal operator one hour and seven minutes to complete. The even more detailed version of the survey, the Cost and Returns Report, was 32 pages long that year and required an average of one hour and 36 minutes to complete (Weber and Clay, 2013, p. 757). These surveys create high levels of burden on selected individuals, not only due to their length but also as a result of the effort (including the process of checking records to identify a response) and the stress involved, such as when sensitive questions are asked.

Average burden figures, such as those emphasized in OMB evaluations, mask substantial variation across farm operations in the amount of effort required to comply with data requests. As described in Chapter 1, a relatively small number of large farms now account for the overwhelming majority of production for some commodities, including eggs, fed cattle, and some vegetable crops. Being involved in multiple activities, as many of these large farms are, leads to their experiencing an increased respondent burden. Also, certain categories of farmers may be contacted on numerous occasions, even on a yearly basis. For some surveys, the

[43]The USDA agency that imposes by far the most burden hours is the Food and Nutrition Service (FNS), which must collect information on a number of major programs, such as WIC and SNAP, and produce statistics on food distribution and child nutrition. Appendix 2.2 to this chapter itemizes the burden figures for FNS.

probability of a large operation being sampled is 1.0, or very close to it, for recurring surveys, especially with establishment surveys (McCarthy, Beckler, and Qualey 2006).

A complex farm includes many working entities, information on each of which has to be reported. Additionally, large, complex operations may face an increased burden if the definition of the responding unit—that is, who should answer the questions—is unclear. With added mandatory reporting requirements for such things as pesticide use, fertilizers, and water quality, the total respondent burden of compliance requires significant resources to manage, especially for large operations. This burden does not originate solely from the USDA but stems from the requirements of other federal, state, and county departments and agencies as well. Indeed, relative to many state and county requirements, the mandatory Census of Agriculture represents only a small fraction, as a percentage of total reporting burden, in part because it is only required once every five years.[44]

How Does Burden Affect Respondent Cooperation and Data Accuracy?

Minimizing the burden placed on survey respondents is a matter of deep concern at statistical agencies, for several reasons. The most obvious reason is that people's time is an economic good in its own right. As with all productive members of society, farmers' time has high value. Thus, reducing respondent burden brings down the full cost of data collection. Another important motivation for reducing burden is provided by survey research[45] which, while not definitive, suggests that increased burden can lessen survey cooperation. Reduced cooperation, in turn, can affect the robustness of findings and conclusions based on analyses of the resulting data. This is an especially pressing issue in today's climate of declining survey response rates and increasing survey costs.

Declining response rates are a problem for surveys throughout the U.S. statistical system. Recently, even the mandatory Census of Agriculture has achieved only around an 80 percent rate.[46] And roughly one-third of sampled farm operators ignore the ARMS survey entirely, an occurrence known as "unit nonresponse." This leaves a unit response rate for ARMS that is well below the 80 percent level, which triggers the OMB, which monitors all federal information collection, to require the administering agency to conduct a nonresponse bias analysis (U.S. Office of Management and Budget, 2006, p. 8). Although low unit response rates do not always create significant nonresponse bias in the resulting statistics, the lower the rate the greater the effect on any derived estimates. That is, the lower the unit response rate, the greater the effect that any differences in answers between respondents and nonrespondents will have on estimates based on respondent-only data (Groves, 2006). Besides unit response, among responding units

[44]For illustrative purposes, a farm in California may be asked to fill out the Agricultural Commissioners' crop, production, and income reports for several counties; the California Water Resources Control Board, Irrigated Lands Report for crops, nitrogen planned to apply, nitrogen actually applied, and erosion control plan; Air Quality Management District report for tractor hours used, truck mileage used, and stationary and mobile engine hours used; the State of California's pesticide-use report for every field; the State of California Environmental Reporting System's hazardous materials inventory for each facility; the commodity purchaser's report for pesticide use, water use, and production; the California Certified Organic Farmers' Production report on inputs and income; the State of California Organic Program report on income; the Federal Food Safety Modernization Report; and numerous federal and state market news surveys.

[45]For examples, see Hansen (2007); Galesic and M. Bosnjak (2009); and Beckett et al. (2016).

[46]The response rates for the Census of Agriculture in 2012, 2007, and 2002 were 80.1 percent, 85.2 percent, and 88.0 percent, respectively. (https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1-Chapter_1_US/usappxa.pdf).

nonresponse to certain questions can be greater than 50 percent on the ARMS questionnaire (Miller and O'Connor, 2012).[47] Such high item nonresponse may indicate that that the questions are too difficult to answer, either because they require information to which respondents do not have access (for example, property taxes on rented land or contractor expenses), or because they are too complex or require detailed record checking, or because they are sensitive. Again, whether it is measured as length, effort, or stress, burden may be leading to missing data.

Reasons Underlying Nonresponse

If greater cooperation with data requests is to be achieved, it is essential to understand why farm operators sometimes choose to not respond to agriculture surveys. Here, we identify several of the key factors.

Length

Based on a limited research literature, one reason for nonresponse is a resistance to committing the time necessary to complete questionnaires.[48] A study by O'Connor (1992) of the 1991 Farm Cost and Returns Survey found that the most common reason for noncooperation—accounting for one-quarter of all refusals—was that respondents "would not take the time / were too busy." A study by Gerling, Tran, and Earp (2008) of the 2006 ARMS in the state of Washington reported the same top-level finding. Changes in questionnaire design and data collection modes have occurred after some of these studies, so evidence on the relationship between survey length and nonresponse cannot be assumed to apply uniformly over time or from survey to survey.

One factor affecting survey length is the ability to use information residing in other sources effectively. Often, respondents are called upon to provide the same information on multiple questionnaires. This problem sometimes surfaces because enumerators are not always able to tap into information from previous interviews. In part because data are confidential, situations arise in which information previously collected must be confirmed (re-collected) by enumerators. Information on the number of acres, for example, does not always prepopulate in NASS data sets, but possibly it could be configured to do so.[49] Of course, confidentiality protections are needed, but statistical agencies are increasingly finding ways to automate surveys that prepopulate previously collected data. The Census Bureau's management of the annual American Community Survey, which includes such capacity, serves as something of a model.

Effort

The reasons underlying respondent nonresponse, particularly item nonresponse, go beyond time burden. Some questions are problematic to answer for conceptual reasons, such as when it is not clear which entities, processes, or activities to include in responses; others may be ignored for practical, logistical, or cost reasons.

As part of its information gathering, the panel met with a number of large farm producers to understand their farms and practices and the challenges they may have reporting on them. One conceptual hurdle highlighted by producers is how to match the structure of their farms with the

[47]NRC (2008) includes a detailed discussion of item nonresponse in the ARMS.

[48]Herzog and Bachman (1981) and Peytchev and Peytcheva (2017) find that survey length can lead to reduced quality reporting.

[49]NASS is piloting the idea of doing just this for its Acreage and Production Survey.

categories and concepts used by the survey. Despite NASS's well established and professional approach to collecting data, producers reported cases where answering questions was hard because of ambiguity over which entities, processes, or activities to include. More clarity about the definition of production units to be used for reporting would alleviate some of these difficulties. And, as outlined in Chapter 5, a Farm Register that followed the accounting structure used by the farm business could also help to demarcate entities that an individual respondent could better understand.

Farm operators also reported to the panel that surveys sometimes ask for data formatted in a way that does not align with the way information is recorded in their accounting or other record keeping systems. For example, one operator reported that its accounting software recorded combined fuel expenditures. The ARMS, however, asked respondents to differentiate gasoline from diesel fuel, so obtaining that level of detail would require a review of all of the fuel receipts. As another example, a question on the 2017 integrated Census of Agriculture/ARMS, asked for data on tractors based on horsepower. One operator interviewed by the panel has more than 800 tractors, so the question would require him to go through the farm's entire inventory to provide accurate numbers.

Flows of income for a given field may be another example requiring a difficult effort because, among other reasons, income from a crop cycle is not constrained within a single year; payment for a seed crop, for example, may not come in until 18 months after planting and therefore involve guesses about value. Another example is when there are two different products from a field. In such cases, the dividing lines for sorting inputs and linking them to outputs make reporting difficult and do not capture the economic processes, which are integrated.[50] In contrast, questions about acreage are examples of information that is relatively straightforward to report.

In the above situations, operators may have little choice but to skip a question, provide a guess, or spend a large amount of effort preparing an answer—that is, to convert the records into the format requested by the survey. Ideally, the format of information requested by the survey would match the format of records as they are typically kept by farmers.

The trend of production moving to large operations composed of multiple enterprises, which require legal and contractual complexity, exacerbates these challenges and, by extension, increases respondent burden. When a farm with multiple owners or managers overseeing a range of product lines is surveyed, this may necessitate the involvement of several people in information reporting tasks, which may be difficult to coordinate (Weber and Clay, 2013, p. 758).

In general, the cost accounting required appears to generate information that, for complex entities, is at best difficult to interpret and at worst simply inaccurate. Indeed, an earlier expert panel convened by the Committee on National Statistics (CNSTAT) (NRC, 2008, p. 80) recommended improving the understanding of respondents' record-keeping practices in order to assess their effect on the quality of survey-collected data.

Question Sensitivity/Stress

There is anecdotal evidence, echoed in the panel's meetings with farmers, that certain types of questions are perceived as sensitive. Previous studies cited by McCarthy, Beckler, and Qualey (2006) of operators in North and South Dakota found that privacy concerns were a

[50]There are also issues in the timing of questions (i.e., how to allocate income when the crop overlaps the reporting year). In some instances, such as with crop futures, the actual income related to crops may not occur until some future year.

reported reason for refusals to participate, even though they were not the reason that was mentioned most frequently. Financial information can be seen as sensitive or proprietary and releasing it may be seen as risky.

Understanding which questions might be perceived as sensitive or difficult is an important task for statistical agencies. Identifying these questions during pretesting is an approach that allows agencies to consider whether they should be modified or asked at all. Such identifications are generally made through cognitive interviews, focus groups, and other pretesting methods. A previous CNSTAT panel (NRC, 2008, p. 80) recommended that NASS and ERS "test revised instruments before they are put into production, and use experimental control groups to evaluate the differences between the old and new questionnaires." NASS responded, instituting a variety of testing methods for both the ARMS and the Census of Agriculture, including cognitive testing, focus groups, and split-sample field tests.[51] It is important that testing continue and cover diverse types of farms, especially complex ones.

Among the reasons researchers found for nonresponse among farmers is that farmers are willing to complete other kinds of surveys "but not financial surveys," and that "information [requested is] too personal / none of your business" (Gerling, Tran, and Earp, 2008). One large-operation producer, who described to the panel his own complex farm operation, cited several reasons why a respondent may be reluctant to participant in NASS surveys: (i) not wanting to reveal information about net worth; (ii) the enormous amount of time commitment required; (iii) the need to create and maintain a separate database; (iv) perceived political agendas showing up in the surveys; and (v) failure to see what is in it for the respondent.

Frequency

When combined with other reporting requirements—such as county pesticide reports, regional water quality reports, and all the other county, state, and federal forms—the Census of Agriculture and ARMS are perceived by respondents to generate high levels of burden to farm businesses. As reported by one of the operators visited by the panel, compliance with data requests requires not hours but many days of labor each year.

Concerning household surveys, there is speculation that one reason for survey nonresponse is survey fatigue, that is, persons receiving too many requests to complete surveys. This hypothesis is difficult to assess for household surveys, given the relatively low sampling rates for virtually all such surveys. Therefore, there is scant evidence to confirm or reject this explanation. One study of students found that the number of prior survey requests did predict nonresponse to a new survey (Porter and Whitcomb, 2005).

McCarthy, Beckler, and Qualey (2006) assess the impact of frequency on the success of subsequent requests to complete a new survey and note that 73 percent of farm operations in the NASS frame were never sampled for the surveys during the four-year period examined. Further, among the 27 percent of operations that were sampled, 72 percent were sampled for four or fewer surveys. The authors evaluate five measures of burden: the number of contacts; the number of surveys completed; the total length of all completed surveys; the number of days since previous contact; and previous participation in ARMS. They look at previous participants in ARMS separately from other surveys, since it is a particularly long survey. What they find is uneven evidence: for some survey requests, but not all, each measure was associated with nonresponse.

[51] See https://www.nass.usda.gov/Surveys/ARMS_Progress_Report.pdf.

The Link Between Farm Size/Complexity, Burden, and Nonresponse

Large farm operations exhibit lower response rates to ARMS and Census questionnaires than do small farms. Whether the higher burden of reporting on large, complex operations is a factor in response rates is a key question of interest to statistical agencies. Weber and Clay (2013) address the size variable by examining the motivations and characteristics associated with unit nonresponse for the ARMS. As represented in Figure 2-1, they find that

> response rates decrease monotonically with farm size . . . [and] even after controlling for other farm and household characteristics, farm operators who do not respond have substantially greater sales than respondent operators, part of which reflects that completing the survey takes longer for operators of larger farms (p. 756).

Descriptive statistics calculated for the study indicate that "refusal" farms have higher average sales ($902,327) than responding farms ($518,934) and own more land (906 acres compared to 627 acres). Because of the skewed distribution of farm output, with a small percentage of farms producing a high percentage of the total output, it is also useful to examine median figures. The median refusal operator harvested more than twice as many acres as the median respondent operator (307 acres compared to 132 acres), and the difference in sales was almost as large ($229,130 compared to $120,454). In "probability proportional to size" sampling, size would be associated with the number of times sampled, which makes it difficult to assess precisely why these associations occur. For example, it could be driven by the frequency of sampling, or it could be driven by the survey in question being more onerous to complete for large operations, due to the reasons noted above.

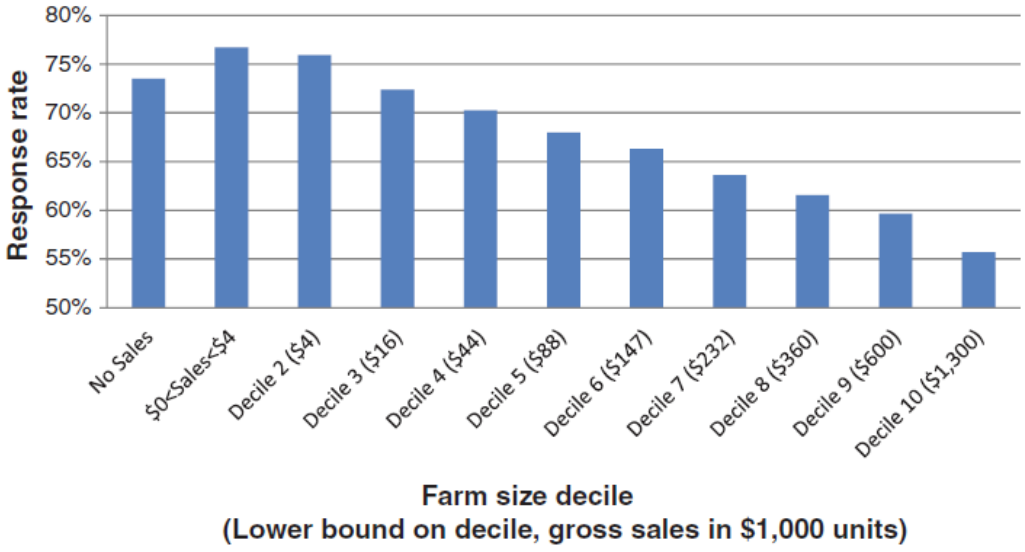


Figure 2-1 Response rates to ARMS, by farm size, 2012
SOURCE: Weber and Clay (2013). Reprinted with permission.

The results from Weber and Clay (2013) corroborate earlier findings from Earp et al. (2008a,b). By comparing means of variables for ARMS respondents and nonrespondents who also responded to the 2002 Census of Agriculture, which is mandatory and captures higher response rates, Weber and Clay (2013) find that "the matched sample means for sales, production expenses, and acres operated all exceeded those from the respondent-only sample, implying that nonrespondents have larger farms than respondents" (p. 757).

Even after conditioning on-farm and operator characteristics using a multivariate analysis, Weber and Clay (2013) find that "response propensities decrease in a perfectly monotonic fashion when going further out in the distribution of farm size" (p. 763). In their model, which accounts for the relationship between farm size (sales class) and complexity, respondents who received the least burdensome version of the survey, the Core survey, were indeed most likely to respond:

> Conditional on many other variables, receiving the shorter Core version increases the propensity to respond by 0.035 [3.5 percent] relative to receiving the Version 1 Cost and Returns Report, which over the last three ARMS took an average of 30 minutes longer to complete than the Core (p. 763).

If the differences in response rates are attributable to the length of the survey, the estimates "imply that decreasing the time required to complete the survey by one hour would increase the propensity to respond by 7 percent (or 3.5 percent per half hour)" (p. 763). Under this scenario, response burden accounts for 21 percent of the different response propensities between the smallest and largest farms; however, this would be an underestimate if the nonrespondent group consists disproportionately of large farm operators who anticipated that completing the questionnaire would take a long time (Weber and Clay, 2013, p. 763). As noted above, however, the time spent completing the questionnaire is not the only variable affecting response burden, so shortening the questionnaire without dealing with other issues of questionnaire clarity would not necessarily increase the response rate. Weber and Clay (2013) argue that

> The time and disutility of responding will tend to increase with the size and complexity of the farm. Looking at ARMS response times over the last three years reveals that farm operators in the largest sales decile took about 55 percent longer (36 minutes) to complete the survey than operators of farms with no sales (p. 758).

At the same time, it is unclear what aspects of complexity matter most for burden: Weber and Clay find that sole proprietorships were slightly less likely to respond to the ARMS, while farms using production contracts were more likely to respond. It is also unclear if all types of burden reduce respondents' willingness to participate. Completing two shorter surveys six months apart may have a different effect from completing one long survey. For example, McCarthy, Beckler, and Qualey (2006) produce results indicating that "burden (for example, the number of other NASS surveys operations were contacted for, the length of time since they were last contacted, and the type of information they were contacted for in the past) does not uniformly have a negative effect on survey response" (p. 97).

Finally, the form of the contact may play a role in response. NASS has already identified highly tailored strategies to recruit large operations to surveys like the ARMS. These strategies often include personal contact from NASS field staff who build relationships over time with the

large operations. NASS might consider further tailoring contact strategies for small and medium-sized operations. Thompson and Kaputa (2017) present the results of experiments with small and medium-size manufacturing establishments. They find that different contact strategies are more effective for operations of different sizes.

Impact of Burden on Data Quality

Greater respondent burden can reduce data quality by reducing the willingness of farmers to respond to any questions (resulting in unit nonresponse), or to particular questions (resulting in item nonresponse), or to give careful and accurate responses. Aside from reducing the effective sample size, an increase in the percentage of farmers refusing to answer any questions has several potential implications for data quality, including the following:

1. *Introducing bias into estimates of totals and ratios of totals (mean values).* Such a bias is introduced if respondents differ from nonrespondents and there is no recalibration of sample weights. NASS researchers have used Census of Agriculture data on ARMS respondents and nonrespondents to assess bias in unconditional means of variables of interest. They find that NASS recalibration methods generally correct for nonresponse bias in the estimates of totals for many variables (Earp et al., 2008a; McCarthy et al., 2017).
2. *Introducing bias into estimates of conditional means.* This bias appears if, for example, the relationship between operator education and crop yields is different for respondents and nonrespondents. In two examples that they examined, Weber and Clay (2013) did not find evidence of differing conditional means: "Despite the observed differences between respondents and nonrespondents, we find minimal nonresponse bias in the two econometric models estimated. The coefficients estimated from the respondent sample always fall inside the confidence intervals generated by repeatedly drawing from the full sample of respondents [provided by the Census of Agriculture] and nonrespondents" (p. 756). They conclude that nonresponse bias in ARMS is unlikely to undermine conclusions based on analysis of conditional means, but they also note that bias can vary from application to application.
3. *Introducing bias into estimates of percentiles.* This too is a bias that can appear if respondents differ from nonrespondents in ways that are not addressed by re-weighting methods. NASS only reports totals and ratios of totals, but ERS regularly reports percentiles such as median farm household income. Neither agency has explored how current re-weighting methods affect estimates of percentiles. However, Robbins and White (2011) find that NASS imputation methods for direct and counter-cyclical payments dramatically understated payments to farms at the lower end (10th and 25th percentile) and upper end (90th percentile) of the payment distribution.

Farmers refusing to respond to particular questions can introduce similar issues, but responses to other questions can allow for imputation of missing values. At the same time, to the extent that an imputed value differs from the real value, imputation introduces a source of nonsampling error (National Research Council, 2008, p. 107). The error could increase dramatically as more respondents refuse to answer questions: greater nonresponse increases the need for imputation and, at the same time, reduces the accuracy of imputed values because less

data is available to establish the statistical relationships on which imputation is based. Moreover, care must be taken to ensure that imputation methods do not lead to an understatement of variance or bias in distributional statistics (such as the percentage of farms with off-farm income) (Ahearn et al., 2011; Robbins and White, 2011).

Burden can also affect the quality of the data that farmers themselves provide. A farmer suffering from survey fatigue may begin to provide very rough numbers just to finish a survey sooner. This aspect of data quality is perhaps even more pernicious than explicit refusals of questions or of the survey itself, because it can be difficult, if not impossible, to distinguish between careful responses and rough guesses. Operators of complex farms in particular may find many questions inappropriate or simply unclear and become frustrated. For example, should the hired manager or the farm owner be listed as "the principal operator?" One panel member enumerated the ARMS with a farmer friend and found that mounting confusion over how to apply questions to his operation led to less concern for providing specific and accurate information.

APPENDIX 2.1
GLOSSARY OF TERMS, WITH USDA DEFINITIONS AND ALTERNATIVE DEFINITIONS

This appendix lists some of the key terms used in this report and in the measurement of the farm sector and broader agricultural sectors. Some of these terms are used in multiple ways, which can create confusion for individuals providing information or interpreting it. In the following table, the middle column reproduces definitions supplied by USDA—such as in the USDA/ERS glossary[52], and as generally described in Chapter 2 of this report. The right-hand column includes some alternative definitions that are used in later chapters of this report to help describe more precisely some of the measurement approaches proposed by the panel.

Terms	Official (USDA) definition, used in Chapter 2 of this report	Alternative definitions, as used in Chapters 3–6 of this report
Business-related terms		
Farm	A place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the census year.	An establishment (single unit with a legal or informal management structure) that (1) has its principal or secondary activity in farming, with the production of agricultural products and biological assets as seeds and animals; and (2) for which full economic data on key business variables, such as costs and revenues, can be collected and made available.
farm business	A farm with an annual gross cash farm income of more than $350,000, or a smaller operation where the principal operator has farming as his primary occupation.	A collection of business establishments with at least one farm establishment linked by common ownership or control; this includes cases in which one business owns and operates one establishment (a simple farm business) or in which one business owns and operates a group of establishments (a complex farm business).
farm operation	(USDA treats terms “farm” and “farm operation” as synonymous)	Same as for “farm.”
farm establishment		A business establishment engaged in farming.

[52]See https://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/glossary.aspx

Terms	Official (USDA) definition, used in Chapter 2 of this report	Alternative definitions, as used in Chapters 3–6 of this report
family farm	A farm in which "the majority of the business is owned by the operator and individuals related to the operator by blood, marriage, or adoption, including relatives that do not live in the operator household."	
Farming	Not defined.	The management of biological processes in crops or livestock.
Agriculture		The sector of the economy that includes both farming and agricultural support activities, as defined in NAICS.
People-related terms		
Farm operator	The person who runs a farm, making the day-to-day management decisions.	The owner(s) of the business entity who are responsible for decisions made on the farm (by appointing managers if there are others), and who bear(s) all the financial risks.
principal operator	Determined (defined) by the respondent in the ARMS; not used in the Census of Agriculture.	
producer	Replaces the term "operator."	
Farm household	Those who share dwelling units with principal farm operators of family farms; determined by survey respondents.	
Land-related terms		
place	USDA uses "place" in a non-standard way. A "place" does not imply contiguous land parcels where ownership and management overlap.	
field	A continuous area of land devoted to one crop or land use, such as a farmstead, pastureland, woods, or wasteland.	
Terms for outputs		

Terms	Official (USDA) definition, used in Chapter 2 of this report	Alternative definitions, as used in Chapters 3–6 of this report
Farm outputs/products		Goods and services produced that fall under NAICS 111 (Crop production) and 112 (Animal production and aquaculture).
Agricultural outputs/production		A broader sector than farming that also includes agricultural services, many of which are found in NAICS 115 codes (Support Activities for Agriculture).

APPENDIX 2.2
TOTAL RESPONDENT BURDEN FOR NASS FOR 2017

Prepared for the panel by David Hancock and Kathleen Ott, NASS

For each data collection NASS conducts, a docket is submitted to OMB for approval to conduct the survey. The docket contains information on the survey sampling plan, sample size, data collection plans, questions asked, analysis plan, estimated number of burden hours that will be placed on the public, estimated number of contacts that will be made, and other relevant information about the survey. Table 2A.1 shows the current OMB approved survey dockets for 2017 (as of December 2016) and an estimate of the number of burden hours and number of contacts that would be made.

Notes for reading Table 2A.1. Each docket listed may contain multiple surveys. For example, the Agricultural Surveys Program docket contains the quarterly Crops Agricultural Production Surveys, the Quarterly Hogs Survey, the biannual Cattle Survey, the biannual Sheep and Goats Survey, and the June Area Survey. Surveys with no hours or contacts listed are currently inactive.

Neither column is the same as the number of operations sampled. Some of the surveys are conducted multiple times during the same year, so the number of contacts is much higher than the actual sample size for that survey (such as for the weekly crop weather and weekly broilers).

The number of operator contacts is an estimate of the number of times an operation will be contacted, given multiple contacts for a survey and estimates of nonresponse during each contact. For example, if a questionnaire is mailed out twice and nonrespondents are called by telephone, a respondent could be contacted one, two, or three times. Some of the contacts would not occur until 2017 (such as many of the Census of Agriculture contacts).

The number of burden hours is the estimated number of contacts multiplied by the average estimated time for each contact. The total number of estimated burden hours for all USDA surveys in 2017 was 211,851,887. The total number of estimated operator contacts for all USDA surveys in 2017 was 1,070,506,570.

Table 2A.1 Estimated total annual burden hours and operator contacts for selected NASS surveys, 2017

OMB #	Docket Title	Annual Burden Hours	Annual Number of Operator Contacts
0535-0001	Cold Storage	3,965	10,728
0535-0002	Field Crops Production	200,919	817,100
0535-0003	Agricultural Prices	30,583	167,320
0535-0004	Egg, Chicken and Turkey Surveys	2,493	15,904
0535-0005	Livestock Slaughter	2,504	11,800
0535-0007	Stocks Report	5,581	26,785
0535-0020	Milk and Milk Products	10,035	60,100
0535-0037	Vegetable Surveys	5,838	19,030
0535-0039	Fruits, Nuts and Specialty Crops	36,821	105,250
0535-0088	Field Crops Objective Yield	2,820	8,000
0535-0093	Floriculture Survey	4,950	18,285
0535-0109	Agricultural Labor Survey	12,634	53,000
0535-0140	List Sampling Frame Survey	40,219	354,400
0535-0150	Aquaculture	907	3,233
0535-0153	Honey Survey	8,937	51,660
0535-0209	Supplemental Qualifications Statement	-	-
0535-0212	Mink Survey	89	706
0535-0213	Agricultural Surveys Program	204,764	1,257,250
0535-0218	Agricultural Resource Management, Chemical Use, and Post Harvest Surveys	91,208	148,306
0535-0220	Cotton Ginnings	1,104	6,850
0535-0226	2017 Census of Agriculture	2,763,085	13,468,839
0535-0227	Equine Surveys	-	-
0535-0234	Farm and Ranch Irrigation Survey	-	-
0535-0235	Childhood Injury and Adult Occupational Injury Survey (NIOSH)	-	-
0535-0236	Census of Horticultural Specialties	-	-
0535-0237	Census of Aquaculture	-	-
0535-0243	Census of Ag - Content Testing	42,552	196,550
0535-0244	Nursery Production Survey and Nursery and Floriculture Chemical Use Survey	659	3,115
0535-0245	Conservation Effects Assessment Project (CEAP)	13,080	27,420
0535-0247	Distillers Grains Survey	-	-
0535-0248	Generic Clearance of Survey Improvement Projects	15,000	25,000
0535-0249	Organic Production Survey	13,004	44,032
0535-0251	Residue and Biomass Field Survey	64	220
0535-0252	Wheat and Barley Scab Survey	-	-
0535-0253	Pesticide Protection Equipment	-	-
0535-0254	Current Agricultural Industrial Reports	4,746	15,130
0535-0255	Colony Loss Survey	7,899	53,120
0535-0256	Feral Swine Survey	6,192	19,440
0535-0257	Organic Certifier Survey	885	55
0535-0258	Cost of Pollination Survey	14,987	78,000
0535-0259	Local Food Marketing Practices Survey	28,905	131,600
TOTAL		**3,577,429**	**17,198,228**

The USDA agency with by far the most burden hours is the Food and Nutrition Service (FNS). FNS and two other larger-burden agencies are shown in Table 2.A2 (not all USDA agencies are shown).

Table 2.A2 Total annual responses and burden hours for surveys by three USDA agencies, 2017

	Number of dockets	Total number of annual responses	Annual burden hours
Food and Nutrition Service (FNS)	79	847,066,971	106,736,781
Animal and Plant Health Inspection Service (APHIS)	135	147,424,507	7,530,873
Food Safety and Inspection Service (FSIS)	30	52,660,413	11,469,151

SOURCE: NASS

3
Dimensions of Farm Complexity

Farms that are complex, along many dimensions of their business operations, have existed for decades. In fact, it has always been common for families that own and operate farms to also be engaged in other businesses and occupations (Sumner, 1982; USDA/ERS, 2017). Nonetheless, over time, complex business organizations have become more commonplace in farming. Farming is not unique in this regard; the banking, retailing, and manufacturing sectors have likewise experienced greater business complexity over this same period. But what does it mean to be a complex business? In this chapter, we identify factors that make a farm more or less complex as a business, in an operational and management sense, and how such complexity affects the collection of data from farms.

This chapter focuses on the farm, rather than on the farm household or on demographic or other characteristics or activities of individuals and families that own or operate farms. It is worth keeping in mind, however, that some of the complexity that statistical agencies face when they collect data from complex farm operations is due to the types of demands placed on them to produce information about individuals, households, and families engaged in farming.

The majority of agricultural commodities today are produced by farms that employ more complex business models (Gardner, 2002; Sumner, 2014). The descriptive data from the Agricultural Resource Management Survey (ARMS) presented in Chapter 1 (see Figure 1-2) also indicate that large farm operations dominate market share; the midpoint of production—the point at which half of the sector's production takes place on larger farms and half on smaller farms—is produced on a farm with sales of $1,416,050. These data also show that consolidation of farm production has continued over the last two decades. In 1996, just one third (33.2 percent) of the value of total farm production was produced on farms with sales above $1 million (in 2014 constant dollars), but by 2014 that figure had risen to 57.3 percent.

The importance of larger and generally more complex farming units has significant public policy implications. Limitations on gross income regularly feature in Farm Bill debates.

As an example, current discussions[1] are underway to limit crop insurance subsidies to those farms that have lower gross incomes. This has implications for the stability of the crop insurance markets, which certainly could be affected if larger farms were no longer eligible for subsidies and they chose not to participate in the market. Because the current crop insurance system is backed by reinsurers located in Europe, reduced participation could increase the price of that reinsurance or ultimately lead those insuring entities to stop offering that service. Understanding the decisions made by large and complex farms is certainly important as legislators debate policy outcomes.

Bonnen et al. (1972) describe how effective data systems facilitate empirical work that, in turn, results in sound agricultural policy and private-industry decision making. The authors argue that the decision-making unit of the farm had become "a heterogeneous and functionally dissimilar set of activities and processes" (p. 868). This heterogeneity of farm operations has only become wider in the ensuing years, creating a need for a more detailed understanding of those farms that produce a majority of the food and fiber produced in the United States.

Public policy makers are concerned with distributional effects on individuals and businesses. To fully understand the differential effects of agricultural, environmental, tax, and macroeconomic policy requires robust data-collection measures that cover the full span of the agricultural production sector.

Although large farms, as measured by gross income or production volume, tend to be complex along multiple dimensions that affect the collection of information, size need not translate into complexity. The factors and causes of complexity extend to characteristics beyond production unit size. Indeed, small farms may also be complex, as in the case of small farms that retail their own production at farmers' markets or the case of Community Supported Agriculture (CSA) farms, which are vertically integrated to include retail packaging and delivery operations. In a study of 54 CSAs in California's Central Valley, Galt (2013) finds that farmers' profits and economic rents were very difficult to measure.

To understand the dimensions of complexity, it is useful to consider examples of farming operations that highlight a variety of farm characteristics. Farm complexity may arise from the interactions between business and family, for example when family priorities, the timing of business transitions to meet those priorities, and the economic and government policy context affecting those family priorities all come into play. Featherstone et al. (2012) consider examples that are useful in understanding this dimension of complexity. Figure 3-1 shows the structure of the owner-operators of a business with farm enterprises that began as a sole proprietorship. The priority was to allow for some of the next generation of the family to be involved in the operation, while facilitating the transfer of assets to others in that next generation who might choose not to be involved in farming. The result was the formation of a general farming partnership, four S corporations, and three individual ownership entities.

Featherstone et al. (2012) discuss the complexities encountered in measuring the overall profitability of this total operation, in monitoring its machinery costs, in increasing the efficiency of transferring the assets of the original owners to the next generation, in monitoring family living expenditures, and in positioning the operation for growth. The main objectives in adopting the business and ownership structure depicted in Figure 3-1 were to facilitate the transfer of assets and to manage machinery costs.

[1]See https://www.ewg.org/agmag/2018/06/senate-farm-bill-amendment-would-rein-crop-insurance-subsidies-rich#.W5BDoehKhPY.

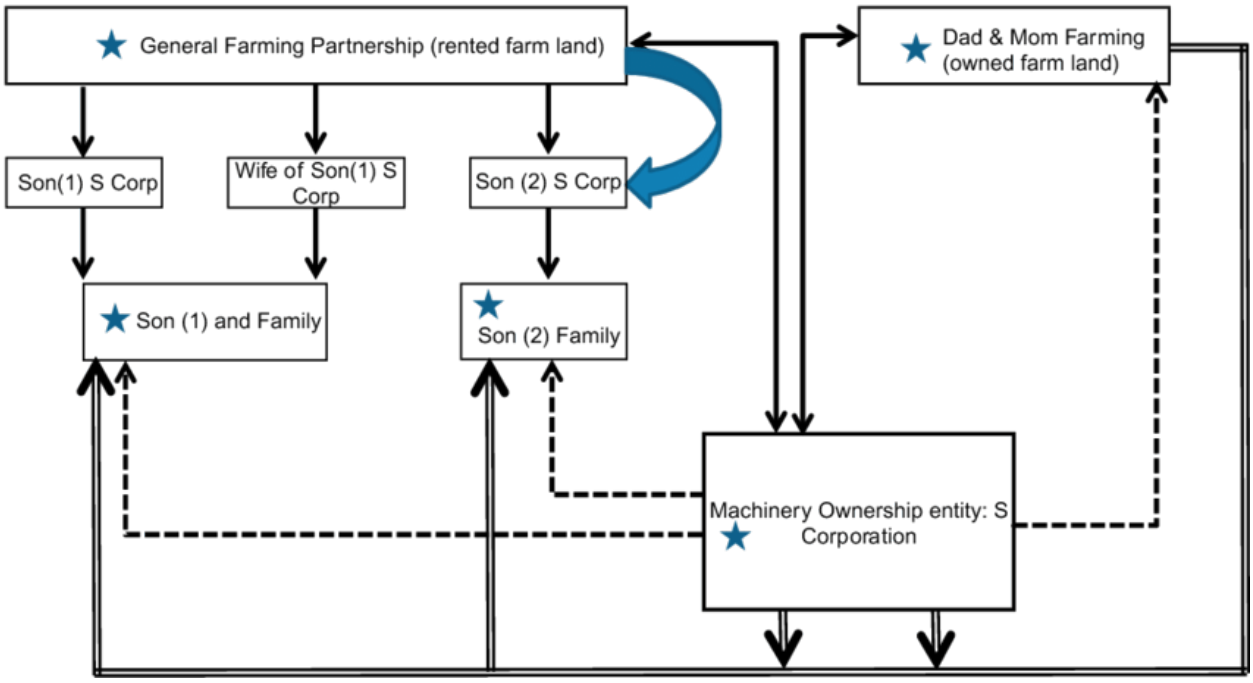


Figure 3-1 Organizational structure of complex farm #1
Source: Featherstone et al. (2012). Reprinted with permission.

Featherstone et al. (2012) also provide details of a second farm, for which measurement of its profitability was complicated by the fact that financial records for different entities involved in the business were kept separately (see Figure 3-2). This second farm began in the 1940s as a sole proprietorship and consisted of both livestock and cropping enterprises. Of the owner's three sons, two were involved in the farm operation while a third wanted to limit his participation to the farm's asset accumulation. To accommodate all concerns, they created a structure consisting of two general partnerships, one limited partnership, three trusts, two S-corporations, and three C-corporations, along with other land-holding entities. Reasons for the formation of the organizational structure included farm program payment limits and tax law advantages—including both increased deductions and facilitation of the transfer of assets from one generation to the next. To obtain a measure of the overall profit of such a farm business, data must be accumulated from different entities, including the personal records from each participant and the C-corporations. Nonetheless, because of the transfer pricing of feed from the cropping partnership to the livestock partnership, income can be attributed to either enterprise for tax purposes.

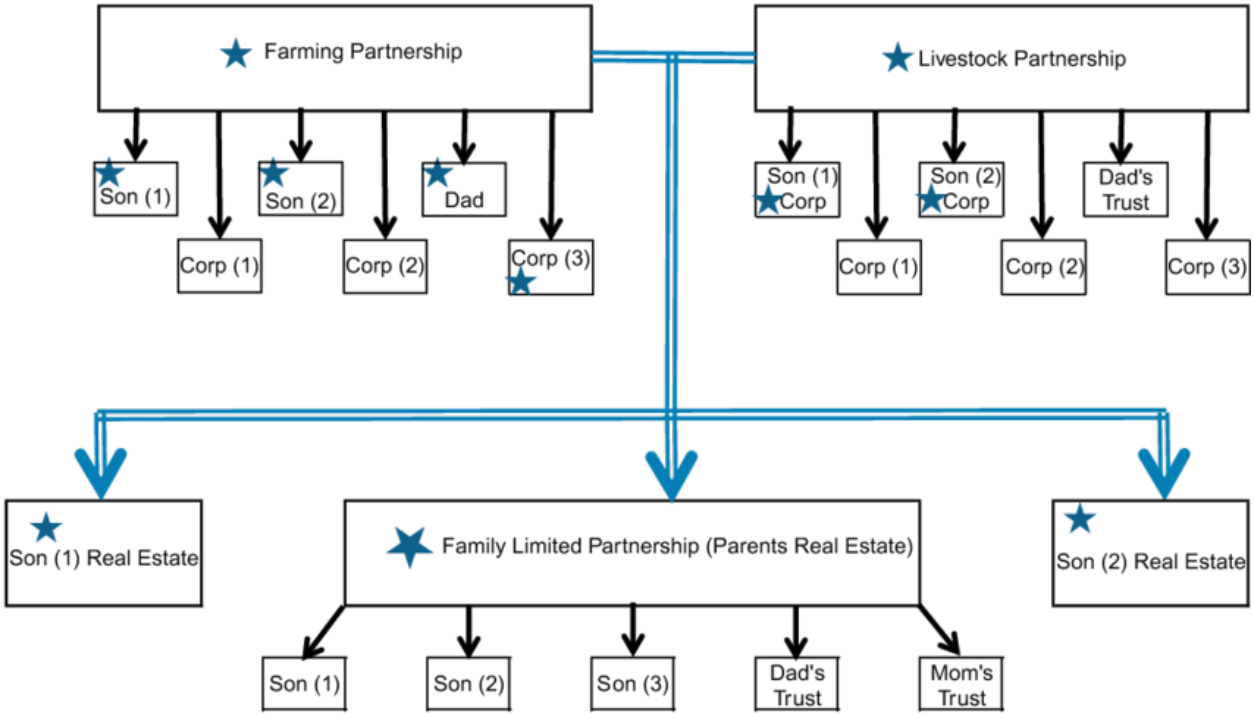


Figure 3-2 Organizational structure of complex farm #2
Source: Featherstone et al. (2012). Reprinted with permission.

Featherstone et al. (2012) discuss the burden on respondents in the collection of ARMS data (as does section 2.4, above) that is related to the business structure. Answers to questions about farm complexity, even about how many farms are counted, are dictated largely by definitions. Questions about size are measured in several dimensions, including quantity of output, value of output, and land area. Questions about ownership involve the extent of value-added enterprises. Questions about contracting arrangements among entities involve both marketing and production. And questions about relationships with nonfarm activities include related companies, employment relationships with labor used on the farm, and the purchase of input services that combine management, labor, and capital.

Each of the above-described dimensions of farm complexity is discussed below. The recommendations presented later in the report are intended to help guide data collection and facilitate accurate measurement in the presence of these complexities.

3.1. FARM SIZE

As the examples just presented show, farm size is often related to farm complexity, but the measurement of farm size itself is complex. The most useful measures of size differ by enterprise and purpose. Often, in comparisons across farms with different commodity enterprises, size is measured by farm value of production. Area of land harvested, number of livestock, and quantity of production or sales are all useful metrics for comparing farms with the

same enterprise or mix of enterprises. However, they are less useful for our current purpose (in this chapter) in the sense that more of a single crop or livestock enterprise does not add to measurement complexity.

Similarly, neither land nor even gross revenue is a particularly useful measure of farm business size when comparing farms that produce multiple commodities, because the mix of productive activities can be so different. For example, 100 acres of strawberries may generate $4 million in cash farm income and, by most categorizations, would be considered a large farm. But a farm with 100 acres of wheat is considered tiny among its same-crop peers. A beef feedlot with $1 million in revenue would market fewer than 1,000 head per year, and is small for that industry, but $1 million in sales for a corn farm is relatively large for grain farms. Farm returns to capital, land, and labor (rather than purchased inputs) are a unifying measure that allows comparing across commodities and enterprises and vertical integration on farms. An example of vertical integration would be the production of livestock feed or carrying of livestock from farrow to finish or from the cow-calf stage through a feedlot.

A value-added measure of a farm—which figures directly into GDP measurement, for example—computes the contributions of farm labor, management, and capital to income generation. The larger a farming operation becomes, the more likely it is to be split into stand-alone operations considered separate from the farming activity and, in turn, to be correctly accounted for in the measurement of GDP. Since the decision of when to separate these activities is made by individual producers, it is difficult to know what percentage of the economic activity is being counted as part of a farm and what is being counted as a separate business. Unfortunately, value added is seldom reported widely for farms, because it demands more data and accounting expertise to construct than the simpler measures of gross revenue and physical outputs and inputs.

One reason larger farms may be complex, as alluded to in Figures 3-1 and 3-2, is that they often have several owners, more complicated legal entities and relationships, issues of transition across generations, and nonoperating owners. Larger farms are more likely to farm land areas that are geographically dispersed and to do so with several distinct commodity enterprises. Larger farms also may be more likely to maintain ownership of nonfarm businesses that are linked to the farm enterprises by vertical integration. For example, when a certain scale is reached, the size of the operation may cause a producer to haul grain further with his or her own fleet of trucks to obtain more favorable prices. In turn, to mitigate liability, a farm may create a separate trucking enterprise to limit legal liability for road accidents through appropriate ownership. Hauling farm products to a local market is a typical farm operation; however, an independent trucking company performing a similar function is not considered a “farm activity” in most USDA measurement programs. If, over time, there is a transition from trucking as a typical farm activity to trucking as a separately owned and operated nonfarm business, complexities and ambiguities arise about where to draw the line between the farm and the trucking enterprise. As explored in the next chapter, collection of data about farm activities requires clarity about these definitional issues.

Understanding the returns to the farm versus the nonfarm parts of the business (such as the trucking just described) requires that each component be tracked. In addition, comparing an operation that does not separate different activities in the organization with operations that may result in noncomparable measures of economic performance, depending on how the data are collected and reported.

In summary, farm size itself is not a dimension of complexity for the farm operation or for the data collection, but farm size is often correlated with complexity in other dimensions, and the collection of accurate data from large farms is especially crucial for industry and sectoral measurements.

3.2. GEOGRAPHIC DISPERSION OF OPERATIONS

The presence of multiple locations for farming activities or multiple addresses for farm management sites creates complexities in farm operation and management and, with those complexities, the potential for significant mistakes in data collection. Results from a survey of NASS field office staff showed that "operations across states/counties" was the third-most frequently selected determinant of farm operation complexity for data collection purposes (Parsons, 2011). When farms operate in multiple counties, data collection and reporting become more challenging for USDA because of the difficulty of assigning production by a single farm to individual counties, which is required for widely used county-level statistics. Biased estimates of statistics such as average yield can lead to skewed outcomes for programs such as crop insurance. Even simple statistics, such as acres by crop by county, may be affected by errors in assigning production correctly to the county or counties of an operation.

Geographic dispersion may also increase survey burden. Respondents in charge of multi-county operations might be surveyed multiple times for the same data fields, leading to their frustration and their lower willingness to participate or to their providing less accurate responses.[2] When separate records are kept for the different locations of a farm, county-based estimates are more reliable. However, to ensure that farm size and other indicators are accurate, it is also vital that such records be associated with the correct farm.

3.3. BUSINESSES THAT OPERATE MULTIPLE FARMS OR OTHER BUSINESSES

Some farm businesses encompass several operations that may each be properly considered separate farms, even if they are managed by a single entity. The examples provided in Figures 3-1 and 3-2 show how defining the separate farms can be complicated and even ambiguous unless data collectors use specific, detailed definitions and concepts. Data collection becomes difficult when there is a sharing of capital and related inputs, or when the transfer of an output from one entity is used as an input into another entity when accounting is done for the entire operation, unless there is a clear understanding as to whether those items are transferred on a cost basis, a market value basis, or by some other valuation process. The returns to the farm can change dramatically depending on the value attributed to the sharing or transfer of items between entities. The issues related to geographic definitions are even more troublesome when overarching organizations include operations that do not farm but may provide services to the farms operated by the organization and sometimes provide services that many farms provide themselves. Such services could be as simple as hired farm labor contracting as a separated business unit rather than being incorporated within the farm business unit.

Separate functional operations or legal entities may be established to optimize government program payments, better implement management strategies, limit liability, reduce tax liabilities, or better align risks and returns when ownership shares differ across operations

[2]See the discussion of this topic in Chapter 2, section 2.4.

and entities. Succession planning is also important in determining the definition of operations and entities. For certain types of analysis, it may be useful to think of different parts of an operation as different companies; for other purposes it may be more useful to think of all the activities as being part of the same company, including the same farm. For these reasons, measurement purpose must first be well understood in order to structure data collection instruments accordingly.

3.4. FARM-CONNECTED NONFARM OUTPUT

The presence of farm-connected nonfarm inputs and outputs, together with the costs and revenues that they generate, adds to farm complexity. Nonfarm business activities are sometimes closely linked to a farm enterprise and, as discussed above, this arrangement complicates the question of where to draw the line between them, both for the business and for the data collector.

This attribution issue is not complex when income is clearly not farm-related, as in the case of a portfolio of income generating stocks and bonds or of income from a spouse's nonfarm wage employment. Complications arise when activities are related to a farm but may or may not count as farm activities in the specific context at hand. For example, a grain farm ("farm A") may provide custom harvesting services for a farm operated by neighbors ("farm B") using the same personnel and equipment used when harvesting on farm A. When a company specializing in harvesting provides such custom services to farm B, the revenue it generates is not part of farm income. The concern arises when the operator of farm A keeps no separate records and simply includes the costs and returns from the custom harvesting business within his or her own farm accounts. It is conceptually clear that services provided on other farms are distinct from farming but, as a practical matter, this sort of farm-related income raises complexities. It is also conceptually clear that the capital and labor used in harvesting on both farm A and farm B are properly counted as inputs into farming in productivity measurements and other indicators, no matter which firm does the harvesting on farm B.

In nearly all commodity industries—among them grain and oilseeds, livestock, and fruits and vegetables—when considering the processes one step upstream or downstream from the farm it becomes difficult to decide whether or not those activities should be classified as farming. If the statistical methods allow different farms to report differently, it will be difficult to correctly interpret the data. In addition, when these activities occur in different entities it also becomes difficult to understand the records, depending on how inputs or outputs are priced as they move from one entity to another entity.

Activities downstream from farms that are part of businesses engaged in processing or marketing are sometimes referred to as "value added" activities. This usage of *value added* should not be confused with the standard economic use of the term—for example, as used in national income accounting—which reflects the contribution of labor and capital to production.

For data collection, the most important characteristic is clarity about the way respondents should report their activities. In some cases, the current reporting relationships are clear. For example, it is clear that cheese processing is not to be reported as a farm activity when it is done by a farm-owned cooperative that is managed separately from the milk production on the farm. In other cases, what farms do or are supposed to do is less clear. For example, when seed cleaning is done by a business operated by a farm producer that also grows the seed crop, and no seeds from other farms are cleaned by the cleaning business, and the seed cleaning is physically located on property connected to the farm, then it may be correctly reported as a farm activity.

This is so even though most farms ship seeds to businesses distinctly set up for cleaning. In the case described, farm accounting may be integrated with seed production. The challenge is that farms differ from one another in the amount of post-production processing or services that are incorporated before they "sell" their farm output.

The current approach NASS uses in its data collection on value-added activities can be improved. Currently, those who supply data are not given clear guidance about how to categorize production activities or income. Additionally, information is gathered for several purposes, such as (i) measuring farm income, (ii) measuring farm production quantity, or (iii) measuring how farm and nonfarm activities are linked in a single business. Thus, guidance needs to be provided regarding the data collection so that users are able understand the purpose for which it is being collected.

Farmer-owned cooperatives are the most common example of separate farm-owned nonfarm businesses. In a cooperative, a group of farms jointly produces or purchases inputs, processes commodities, or markets products through a single firm they jointly own. Farms receive revenue for commodities they deliver to processing and marketing cooperatives, and they gain additional benefits from their ownership proportional to the commodities delivered, which may be received in subsequent years.

Also, many farms are involved in processing inputs that are otherwise conducted by nonfarm firms. For example, animal feeding businesses sometimes operate an on-site feed mill. In some cases, a firm that has an animal feeding farm also operates a feed mill and sells feed to other operations. The measurement challenge is in clearly delineating when the operation of a feed mill is not a farm activity and therefore when its revenue is not farm revenue. Gathering information on the feed mill business is important, whether that business is operated by a farm or not. The general point is that, in some cases, input processing is integrated within a farming operation and no clear market-based transfer prices for milled feed are available to report. In other cases, operating a feed mill is distinctly a nonfarm activity and clear market prices are identifiable for milled feed.

The Penobscot McCrum organization in Maine is an example of a complex farming operation (County Farms, Sunday River Farms) that provides value-added potato specialty products as outputs (Penobscot McCrum) and transport services (JDR Transport) for its farming operations and for other outside companies. In addition, Penobscot McCrum brokers grain (County Grain Merchants), along with marketing and selling both its own potatoes and potatoes for other independent farmers throughout Maine (County Super Spuds).[3] The value-added nature of these activities illustrates the potential complexity of data collection, here depending on the valuation of raw material such as potatoes when transferred from the farm to further processing into baked skins or other value-added specialty products.

3.5. FARM EMPLOYMENT AND DATA ON HOURS AND WAGES

Farms use directly hired labor, family member labor, and labor that is employed by a separate firm. When family labor is employed for an explicit wage or salary, it can be treated in the data in the same way as any employee's labor. Often, however, family members are not paid direct wages or salary, and few clear records of their labor contribution are maintained. This can make it difficult to accurately measure the full labor inputs associated with a farm.

[3]See http://penobscotmccrum.com/family-of-companies/index.php.

While family labor should generally be treated as farm employment, contract labor is not; such laborers are employed by the contracting firm. Because contract workers are not farm employees, associating their labor inputs with specific farm output is complex. The complexity increases when capital and labor are both engaged in the contract services because, in such cases, records that clearly separate costs between labor and equipment are rarely kept.

To accurately measure productivity and understand labor or capital intensity, it is crucial to collect accurate data on the farm labor used for particular enterprises. For example, if custom labor use rises relative to hired labor in fruit production and the data do not allow one to associate the custom work with specific crops, it becomes impossible to track and understand the farm's productivity and related measures on a commodity by commodity basis. Current data collection practices that the USDA's ERS uses to measure agricultural productivity seem to be deficient in this regard, and they will become more so as labor is hired at higher rates by nonfarm businesses and those workers perform an increasing share of the nation's (or state's or county's) farm activities.

3.6. FARM AND BUSINESS OWNERSHIP: LEGAL RELATIONSHIPS BETWEEN OWNERS AND FARMS

Producers sometimes organize their operations using several companies: one company may own the land, another may own the equipment, and still others may own the livestock; and each may be separate from the company that operates the farms. The owners of these companies may be individuals, trusts, partnerships, or other organizations, and the individuals involved may include family members or others, including unrelated shareholders. On some farms, management is separate from company or asset ownership, and several individuals are involved in each area. In these cases, specific contracts are often used to coordinate activities and to specify responsibilities and rewards.

There are many drivers motivating these complex business and legal organizations. These include incentives to reduce tax liabilities, compliance with rules of government commodity subsidy programs, and incentives to reduce taxes or other concerns in transition of assets and business to new ownership (say, across generations). In Figures 3-1 and 3-2, discussed earlier, both operations began as sole proprietorships. The goals of their reorganizations were similar, but the laws creating tax and other incentives varied, resulting in different legal arrangements across the companies. When farms separate into multiple entities in this way, complexity in both data collection and analysis increases.

3.7. MANAGEMENT AND DECISION-MAKING RELATIONSHIPS

Farms are commonly characterized by management that is spread across several individuals. In this way, farm businesses are not unlike businesses in many other industries. Data programs must be designed in a way that ensures that collection draws on the most knowledgeable person in the organization for each information request. Employees of a large farm business may be able to accurately report farm quantities, prices, revenues, and costs, but the same employees may not be able to respond knowledgeably about management responsibilities or about the relationships among owners and hired managers.

These concerns raise practical complexities when it is unclear who can best respond on behalf of the farm. For a large and complex organization, more than one respondent may be

required, depending on the set of questions or the survey. Such a situation with multiple respondents is routine for other industries, and it must become routine in agriculture.[4]

With the goal of improving data collection from farms that display the characteristics described above, we next turn (in Chapter 4) to the definitional questions that must be untangled as a prerequisite to better measurement of farming and agriculture. After that, it will be possible to provide guidance on the data collection infrastructure and on specific data collection instruments, such as the Census of Agriculture and the ARMS, which we undertake in Chapter 5.

[4]A major farm commodity producer in California recently reported that a Census of Agriculture form mailed to his family home in the hills overlooking Los Angeles was tossed out as junk mail, whereas if it had been sent to the farm offices one of the accounting clerks would have duly completed as much as possible and submitted the forms.

4
Conceptual Issues: Defining Farming, Farms, Farmers, and Agriculture

In this chapter we identify and clarify definitions and concepts used in the measurement of the farm economy and propose alternatives that may be more useful for measuring the activities of complex operations. The material here bridges the discussions of complexity in the previous chapter and proposals for the statistical framework in the next chapter.

4.1. FARMING

Observing the complexity of modern farms raises the basic question: What makes a farm a farm? It is essential to address this question before turning to others, such as what a complex farm structure is and how such operations should be dealt with in statistics, research, and policy making.

In the United States, even children who have never seen a farming operation firsthand develop an intuitive notion of what a farm is. Their toys are likely to include plastic cows, sheep, horses, and tractors. They are taught songs like *Old MacDonald*. And with the advent of electronic gaming devices and smartphones, it is now possible for anyone to "run" a virtual farm. These experiences tend to reinforce the idea that the typical farm is an entity that perfectly overlaps the activities of a farmer who centrally manages a diverse set of production activities.

In the contemporary economy, however, this heuristic of the farm has become outdated as more complex organizational forms have emerged and become commonplace. For example, some traditional farming activities, like harvesting crops, are now often carried out by hiring specialized companies or using workers who are employed not by the farm but rather by an employment agency. Some farms are parts of vertically integrated businesses, such as feedlots attached to slaughterhouses or vineyards attached to wineries. And some farms derive income from related activities that are not necessarily farming, such as agri-tourism, food processing, or energy production.

All of these examples create problems for statisticians the moment they want to classify a company (or operation, business, firm, or holding—for now, we use these terms interchangeably) as either a farm or another type of business. Before turning to that classification question, we therefore address an even more basic question: What is farming?

The Merriam-Webster Dictionary defines farming as *the practice of agriculture or aquaculture*, providing a synonym to be discussed later, in Section 4.5. The Oxford English Dictionary (second edition, 1989) defines farming as "*the activity or business of growing crops and raising livestock.*" This latter day-to-day definition seems fairly close to what social scientists use in some of the disciplines relevant to this study. For instance, the International Accounting Standards Board (IASB)[1] describes agriculture (an imperfect synonym of farming) as follows: "*Agricultural activity is the management by an entity of the biological transformation and harvest of biological assets for sale or for conversion into agricultural produce or into additional biological assets*" (IFRS, 2017). The essence of this definition is of a business activity that manages a biological process that leads to either products (such as milk, potatoes, or oranges) or biological means of production (such as animals or seeds).

Definitions Current at Major Statistical Agencies

International statisticians rely on a definition similar to that used by the IASB. Both the Statistical Classification of Economic Activities in the European Community (or NACE)[2] and the international integrated system of economic classifications (ISIC) managed by the UN Statistical Commission (UNSTAT) group agriculture (again, used as a synonym of farming) together with fisheries and forestry in one category for national accounting purposes. This broad category is defined as "*the exploitation of vegetal and animal natural resources, comprising the activities of growing of crops, raising and breeding of animals, harvesting of timber and other plants, animals or animal products from a farm or their natural habitats*."[3] Note that the addition of natural habitats includes fisheries, hunting, and forestry in the category.

Likewise, the North American Industrial Classification System (NAICS) groups farming together with fisheries and forestry activities:

> The Agriculture, Forestry, Fishing and Hunting sector comprises establishments primarily engaged in growing crops, raising animals, harvesting timber, and harvesting fish and other animals from a farm, ranch, or their natural habitats.
>
> The establishments in this sector are often described as farms, ranches, dairies, greenhouses, nurseries, orchards, or hatcheries. A farm may consist of a single tract of land or a number of separate tracts which may be held under different tenures. For example, one tract may be owned by the farm operator and another rented. It may be operated by the operator alone or with the assistance of members of the household or hired employees, or it

[1]The IASB is an independent, private-sector body that develops and approves International Financial Reporting Standards (IFRSs). The IASB operates under the oversight of the IFRS Foundation. The IFRS Foundation is a not-for-profit public interest organization established to develop a single set of high-quality, understandable, enforceable and globally accepted accounting standards—IFRS Standards—and to promote and facilitate adoption of the standards.

[2]Commonly referred to as NACE, based in the French term, *"nomenclature statistique des activités économiques dans la Communauté européenne.*"

[3]See http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=MSG_PRINT_NOHDR&StrLanguageCode=EN&ml=NACE_REV2__18493724__0.

> may be operated by a partnership, corporation, or other type of organization. When a landowner has one or more tenants, renters, croppers, or managers, the land operated by each is considered a farm.[4]
>
> The sector distinguishes two basic activities: agricultural production and agricultural support activities. Agricultural production includes establishments performing the complete farm or ranch operation, such as farm owner-operators and tenant farm operators. Agricultural support activities include establishments that perform one or more activities associated with farm operation, such as soil preparation, planting, harvesting, and management, on a contract or fee basis.
>
> Excluded from the Agriculture, Forestry, Fishing and Hunting sector are establishments primarily engaged in agricultural research and establishments primarily engaged in administering programs for regulating and conserving land, mineral, wildlife, and forest use. These establishments are classified in Industry 54171, Research and Development in the Physical, Engineering, and Life Sciences; and Industry 92412, Administration of Conservation Programs, respectively (NAICS, 2017).

Agricultural activities as classified by NAICS are listed in Appendix 4.1, together with forestry, fishing, and hunting activities. Within this activities listing, Codes 111 (Crop production) and 112 (Animal production and aquaculture) are farming activities. Codes 1151 and 1152 include support activities for agriculture, covering the important set of nonfarm businesses that serve as potential substitutes for direct management and operation of production activities by farmers. Indeed, these businesses introduce two possible sources of statistical complexity: some farms purchase agricultural production services from other entities and some farms sell their production services to other farms.

Neither the National Agricultural Statistics Service (NASS) nor the Economic Research Service (ERS) offers explicit definitions of farming.[5] However, the Census of Agriculture questionnaire follows an enumeration that is similar to NAICS' list of "agricultural activities." The census includes questions on field crops (Section 6), hay and forage crops (Section 7), cut Christmas trees and maple syrup (Section 8), nursery, greenhouse, and floriculture (Section 9), vegetables and melons (Section 10), fruits and nuts (Section 11), berries (Section 12), cattle and calves (Section 13), hogs and pigs (Section 14), equine (Section 15), sheep and goats (Section 16), aquaculture (Section 17), poultry (Section 18), apiculture (Section 19), and other livestock (Section 20) (NASS Census of Agriculture, 2012).

This emphasis on the management of biological processes in crops and livestock makes it clear that farming is not defined on the basis of the purpose to which products are put: it is not necessary to produce food or animal feed. Biological processes are also managed to produce fiber (flax, cotton, and wool), flowers, fur, pharmaceutical products, and fuel, to give just a few examples. However, there are some activities based on biological processes that are classified not as farming but as industrial—such as the industrial production of yeast or the use of microbes in sewage sludge.

---

[4]Conceptually, this panel agrees with the NAICS position that households that own land and rent to operators should not be considered farm businesses. If a household (or any other sector of the economy, such as government or a business) is engaged only in renting the land, it should be classified as part of the Rental and Leasing industry. The farming activity undertaken on such rented land would be measured by the establishment/enterprise using the land, that is, the lessor, not the lessee. Practically speaking, if both the lessor and lessee of the land were both considered to be farms, then there would be a risk of double counting.

[5]See Appendix 2.1 in Chapter 2 for a table with references to definitions used in USDA documents.

Rather than describing activities on the basis of the purpose of the products or even on the basis of the products themselves, the NAICS objective is to describe a certain production function associated with the *way* a process creates outputs (products or services) from inputs. It is not a product classification and so differs from the National Income and Product Accounts, which is. When biomass is produced and transformed into ethanol or electricity, this includes a farming activity, very often as a primary activity, as well as a manufacturing activity, in this case energy production that turns biomass into ethanol.

In some cases, the production activity is very different from farming, even if plants or animals are used. Examples include petting farms, dude ranches, or farms offering yoga with goats. Although these operations may have been conventional farms in the past, focused strictly on crop or livestock production, they do not necessarily fit farm definitions any longer, even if their animals are registered for animal health reasons.

A side remark is that statisticians employ a variety of approaches for dealing with "gray" or illegal activities, for example when the goal is to assess the total size of an economy. The United Kingdom has added income from illegal activities, such as sex work and illegal drug sales, to its GDP calculations to conform with reporting rules from the European Union. In Canada, now that recreational marijuana use will soon become legal, Statistics Canada is preparing to add estimates of the plant's production and sale to assess its economic impact.[6] The U.S. Bureau of Economic Analysis, which produces GDP estimates, only includes legal activities, so in the United States marijuana growing would only be counted as a farming activity in states such as Washington and Colorado, where marijuana production, sale, and use are currently legal.

These examples also show that to operationalize the concept of farming as the management of a biological process, an agreed-upon list of activities is needed that can determine the exact borders where farming stops and nonfarm production and services begin (or, as we argue in Section 4.5, where it may make sense to introduce a new category of agriculture that lies between farming and industry). Such a list of activities could make clear that a tree nursery and an energy plantation (where miscanthus or willows are grown, for example) are still farming, while a forest is not, even if it undergoes some pruning. Fish farming (aquaculture) is considered an agricultural activity, as is the growing of seaweed (algae),[7] but the capture of wild fish from the ocean and harvesting of kelp from the open ocean are not.

Defining Agricultural Support Activities

Essential to this definition of farming is that biological products must be managed. Simply harvesting (as a contractor) is not farming if it is done as a service but is rather an agricultural support activity.[8] The less technical OECD definition is "*the activity or business of growing crops and raising livestock*"; one could add "and their preparation for the primary market" to make clear that harvesting, storing, and packaging are included.

---

[6]See: https://unstats.un.org/unsd/nationalaccount/aeg/2017/M11_8_2_Satellite_Accounting_at_Statistics_Canada.pdf.

[7]NAICS defines algae harvesting as an agricultural activity (112519).

[8]However, custom work performed by a farmer is a category of income on the Schedule F federal tax form that farmers complete. There are many farms that engage in such work, and often they do not have a separate business for it; it is just lumped under their farming income, as the equipment that they utilize is first and foremost for their farming purposes.

Another example of agricultural-related activities that has posed problems for classification purposes is the making of wine and cheese. Are the biological processes in cheese making and wine making characteristic enough to call them farming activities instead of industrial activities? NACE has tried to solve this dilemma by introducing the idea of a primary market of a farm:

> Agricultural activities exclude any subsequent processing of the agricultural products classified under divisions Manufacture of food products and beverages and Manufacture of tobacco products, beyond that needed to prepare them for the primary markets. The preparation of products for the primary markets is included here (that is in farming). The division excludes field construction (e.g. agricultural land terracing, drainage, preparing rice paddies etc.) classified in the section Construction, and buyers and cooperative associations engaged in the marketing of farm products, classified in the section Wholesale and retail. Also excluded is landscape care and maintenance, which is classified in the class Landscape service activities [*text slightly altered for readability by deleting codes*].[9]

And:

> In agriculture, one frequent situation where the breakdown of the value added presents difficulties is when the unit produces grapes and manufactures wine from the own-produced grapes, or when it produces olives and manufactures oil from the own-produced olives. In these cases [. . .] these vertically integrated activities would generally lead to classification of the units under agriculture (FSO, 2008, p. 18).

In other words, by NACE standards, wine and cheese making are not part of farming but of manufacturing. However, because these activities are sometimes not fully separable from the farming activities, for a given entity they may in practice be counted as farming. As discussed in Chapter 3, this is one characteristic that makes farms more complex, because they have the management skill and technical capabilities to run a significant value added component as part of their farming business.

Under code 115, NACE also includes in its definition of agriculture activities incidental to agricultural production and activities similar to agriculture but not undertaken for production purposes (in the sense of harvesting agricultural products) that are done on a fee or contract basis. Also included are post-harvest crop activities aimed at preparing agricultural products for the primary market, such as contract work for sorting or grading products or packaging them. Such contract work may be viewed as an agricultural support activity, and this is one of the differences between farming (this section) and agriculture (see Section 4.5). Companies specializing in such activities are agricultural support firms that perform farming activities on farms. They are engaged in a farming *activity*—that is, farming (managing the whole biological process from inputs to outputs)—which is different from *being* a farm.

To summarize, farming is the characteristic activity that takes place on a farm, and typically it involves the management of a biological process, such as growing crops or raising livestock, for the purpose of harvesting products or reproducing a biological means of production. A list of activities and products such as those included in NAICS industry codes 111

[9]See http://ec.europa.eu/eurostat/documents/1965800/1978839/NACE_rev2_explanatory_notes_EN.pdf/b09f2cb4-5dac-4118-9164-bcc39b791ef5 (p. 2).

and 112 is useful for precisely delineating between farming, agriculture, and manufacturing activities. Conceptualizing farming activities in this way does not imply a change in data collection by NASS and ERS, but it may help both agencies clarify and separate discussions about complex farms as businesses (farms), locations (farms, fields), and activities.

4.2. THE FARM

Having defined farming, or at least described it, the next step is to define what a *farm* is. In principle, this is the entity that carries out farming, whether it is a firm, business, holding, or operation. Applying this principle to the modern economy, however, turns out to be less straightforward than one might perhaps expect.

The Merriam-Webster Dictionary defines a farm as "*a tract of land devoted to agricultural purposes;*" the Oxford English Dictionary widens this to *"an area of land and its buildings used for growing crops and rearing animals.*" However, as argued above, farming does not necessarily demand either land or buildings: pigs and poultry farms may consist only of buildings with no crop production present, but they are still farms; as are greenhouses and buildings used to grow chicory or flowers from roots and bulbs, sometimes cultivating their crops in a substrate instead of soil. And the vertical farms[10] that are the latest development in urban farming are certainly farms. The more scientific definitions, such as those used by statistical agencies, use terms such as "entities," "businesses," "holdings," and "units" that engage in farming activities.

Examining the Principal Activities of a Farm Business

More importantly, the organizational forms that farming activities can take create definitional problems. Statistical agencies that count and describe farms must classify organizations instead of activities; businesses receive a census form and report multiple activities whose operating costs often cannot be disentangled. This is not a problem when businesses are fully specialized in farming, narrowly speaking, and have no other activities, whether agricultural, industrial, retail, or service.

Even in these simpler cases, however, the results of such a classification may raise questions if farms have moved production activities that were once commonly carried out on the farm to other producers downstream in the production pipeline. Historically, for example, cheese and butter making were farm activities but, beginning during the industrial revolution of the 19th century, such activities were increasingly transferred to the food industry. On the input side, support tasks involving contract work or machines, advice from risk-management firms, and work now done by other specialized companies either on the farm or elsewhere are all examples of activities that have displaced comparable ones that were once commonly handled "on the farm." This shift toward specialization means that functions previously classified as farming are now more typically classified in industrial sectors. Therefore, as captured by economic statistics,

[10]Vertical farming is the practice of growing food or medicine in vertically stacked layers, vertically inclined surfaces, or integrated in other structures. The modern idea of vertical farming uses Controlled Environment Agriculture (CEA) technology, where all environmental factors can be controlled. These facilities utilize artificial control of light, environmental control (humidity, temperature, gases) and fertigation. Some vertical farms make use of techniques similar to greenhouses, where natural sunlight can be augmented with artificial lighting.

The Association of Vertical Farming has its own typology; see https://vertical-farming.net/vertical-farming/integration-typology/

the farm sector has contracted, relatively, while industry and service sectors have grown. This evolution suggests that, for some purposes, it may make sense to view agriculture as a sector that is broader than farming (see Section 4.5).

More problematic from a classification perspective is the low level or even lack of specialization in mixed enterprises, such as mixed farms and integrated companies. For entities engaged in several activities, some of them farming and some of them not, there is a need for a rule to classify them as belonging to a certain sector or industry (a class of activities). The horticulturalist who occasionally sells Christmas trees at the roadside is still a farm and not a retailer, but conceptually we enter a fuzzy area where, at the opposite end of the spectrum, there might be a retail garden center that also supplies a few percent of its sales in December from self-grown Christmas trees (a farming activity). Such examples indicate the need for a criterion and a threshold whereby entities may be classified by sector.

In macroeconomic statistics, guided by such frameworks as the NAICS or ISIC / NACE, it is recognized that a unit may perform one or more economic activities described in one or more categories. According to the *OECD Manual on Business Demography Statistics*:

> In such cases the principal activity of a statistical unit is the activity which contributes most to the total value added of that unit. The principle activity is identified according a top-down method and does not necessarily account for 50% or more of the unit's total value added. A secondary activity is any other activity of the unit, whose outputs are goods or services which are suitable for delivery to third parties. The value added of a secondary activity must be less than that of the principal activity. A distinction should be made between principal and secondary activities, on the one hand, and ancillary activities, on the other. Principal and secondary activities are generally carried out with the support of a number of ancillary activities, such as accounting, transportation, storage, purchasing, sales promotion, repair and maintenance, etc. Thus, ancillary activities are those that exist solely to support the principal or secondary economic activities of a unit, by providing goods or services for the use of that unit only.[11]

By applying such a rule and threshold, mixed enterprises such as the Christmas tree sellers described above can be classified as farms (if the tree nursery is the principal activity) or as retail (if the tree nursery is a secondary activity). Adoption of such a method would bring agricultural statistics in line with statistics covering other sectors of the economy, which the panel finds to be an important consideration guiding decisions on the agricultural statistical framework. The value of statistical agencies taking this approach would be generated because the modern agricultural sector is an integral element of the economy, with many medium-sized businesses and much linkage with the rest of the economy. A *status aparte* in statistical methods is therefore to be prevented as much as possible.

> **Recommendation 4.1** In line with statistics for other parts of the economy, NASS and ERS should apply clear rules based on the nature of its principal productive activities for classifying a business as a farm or as an entity operating in a nonfarming sector with secondary activities in farming.

[11]Eurostat, *OECD Manual on Business Demography Statistics* (available at https://www.oecd.org/sdd/39974599.pdf), p. 67.

This recommendation does not imply that only entities classified as farms with farming as a primary activity are of interest to NASS and ERS. On the contrary, the agencies should be interested in all businesses engaging in farm activities, even those for which it is a minority activity. Counting farming activities in businesses that are classified in sectors other than farming is necessary to accurately and consistently estimate totals for the sector when there is a change in the role of large companies in farming. There are good reasons to try to survey agricultural production that takes place within these companies, especially if a large share of the production takes place in them. This may be appropriate because, among other reasons, such mixed production businesses are affected by agricultural policies (see Chapter 2). For reporting on entities engaged in farming as a secondary or tertiary activity, specific classification categories may be needed, such as *part-time farms* and *multi-functional farms*. At the sector- or macro level, the size of agricultural business complexes can be calculated with input-output tables (see section 4.5).

Applying this method of classification requires careful specification of how entities or units are defined. For a company like Smithfield Foods Inc.,[12] which owns both slaughterhouses and farms, should the farms be classed as different units, or should the full multinational company be classified under its principal activity as a slaughterhouse and therefore part of the food industry? Are Walmart stores individual units, or is all corporate activity one business? The answers to such questions—presented in detail in Chapter 5—determine the number of farms counted in an agricultural census.

Distinguishing Between Firms and Establishments

In economic statistics, the answer to the above-posed question is dictated by the difference between firms and establishments. A firm is "*an organization conducting a business . . . . A firm may operate one place of business or more."* An establishment is *"a single physical location where a firm's business is conducted"* (NRC, 2007).[13] For our example, this means that Walmart is a firm with many establishments; and that the farms of Smithfield Foods can be counted as farms, if we define a farm as an establishment and Smithfield Foods is organized in such a way that data from the farms can be separated from other activities.

Eurostat has followed this accounting approach, stating that: "*A farm is a single unit, both technically and economically, which has single management and which produces agricultural products … either as its primary or secondary activity."*[14]

[12]Smithfield Foods, Inc., is a meat processing company and wholly owned subsidiary of the WH Group.

[13]In line with BEA's Nov. 2017 update to its handbook on "Concepts and Methods of the U.S. National Income and Product Accounts": "Companies consist of one or more establishments owned by the same legal entity or group of affiliated entities. Establishments are economic units, generally at a single physical location, where business is conducted or where services or industrial operations are performed (for example a factory, mill, store, hotel, movie theater, mine, farm, airline terminal, sales office, warehouse, or central administrative office). Establishments are classified into an industry on the basis of their principal production method, and companies are classified into an industry on the basis of the principal industry of all their establishments."

[14]From Eurostat's online *Glossary*, under "Agricultural Holding" (http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Agricultural_holding). The online text further states: "An agricultural holding, or holding or farm is a single unit, both technically and economically, operating under a single management and which undertakes agricultural activities within the economic territory of the European Union, either as its primary or secondary activity. Other supplementary (non-agricultural) products and services may also be provided by the holding."

As covered in Chapter 2, the official definition of a farm used by NASS is as "*any place*[15] *from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the year.*"[16] In line with the day-to-day use of the word "farm" and the current practice in NASS and Eurostat, a farm should be defined in a way that focuses on the productive entity as a business engaged in clearly specified types of activities:

> **Recommendation 4.2** For conceptual purposes, NASS and ERS should define a farm as an establishment (single unit with a legal or informal management structure) that (1) has its principal or secondary activity in farming with the production of agricultural products and biological assets such as seeds and animals; and (2) for which full economic data on key business variables, such as costs and revenues, can be collected.

This recommendation is intended to help NASS and ERS unravel the structure of some common types of complex holdings. When a person or businesses has two clearly distinguishable farms with separate accounts, such as one farm in Wyoming and one in California, or one on either side of the same county, the farms should be treated as two farms that happen to have the same owner. This definition is quite close to the USDA's approach, which defines a farm as a management unit.

Regarding the practicalities of implementing recommendation 4.2, the panel recognizes that there may be scant political will to change the definition of a farm, regardless of how antiquated the dollar threshold may be, due to the implications for federal funding and other reasons. Indeed, the transfer in 1997 of the data collection for and publication of the Census of Agriculture from the Census Bureau to NASS resulted, in part, from congressional concerns about a proposal to modify the definition of a farm.

Discussions of the politicized aspect of the farm definition often take place in the context of proposals to increase the farm-size threshold (e.g., O'Donoghue et al., 2011), which would invariably reduce the number of farms. The recommendation above, which is expanded on in Chapter 5, is quite different: it recommends that NASS and ERS use a farm enterprise concept and a farm establishment concept when collecting data in order to (i) make the concepts as consistent as possible with those used by other statistical agencies, (ii) provide clarity to survey respondents, and (iii) identify business units that are likely to correspond better with the way respondents organize their own data.

The establishment concept is not fully equivalent to the current farm concept used by USDA to estimate the number of farms, but it probably comes close. The majority of farms, especially the smaller ones, have a simple structure and will report their current operations as establishments, resulting in the same number of farms. For complex farms, it is currently unclear whether they would report one of their establishments or report a complex farm corresponding to a statistical enterprise. If they were to report the latter, the number of farms (farm establishments) would increase; otherwise it would stay unchanged. Recommendation 4.2 should therefore be interpreted as a recommendation to help obtain a clearer picture of the farm

[15]USDA uses "place" in a nonstandard way. A "place" in its usage does *not* imply contiguous land parcels where ownership and management overlap.

[16]See the USDA/ERS Glossary (https://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/glossary/) which covers the following terms: Farm; Farm operator and principal farm operator; Family farm; Farm operator household; Farm operator household income; Farm operator household wealth; Farm typology; Commodity specialization; Disposable personal income of farm and nonfarm residents.

structure and reduce the administrative burden for complex farm structures, and not as a recommendation that intends to reduce the number of farms.

However, if administrative or political reasons required that there be a check with the traditional farm-count number, ERS/NASS could develop a mapping from enterprises and establishments to farms. It could then provide counts at the county, state, and national levels that would be consistent with those that have historically been produced. Such a mapping could be done to preserve the distribution of farms across space or in total in a given base year.

Counting and Measuring Farms and Fields for a Complex Farm Business

The definition of a farm as an establishment means that there are businesses that may be involved in two or more farms. (The logic of this is elaborated on in Chapter 5, where it is also clarified why in statistical terms businesses are equated with firms.) Even defined in this comparatively granular way, a single farm may engage in several different activities, such as growing different crops or raising different types of animals or doing both. A cropping farm will have different fields that are not necessarily adjacent to each other, and so, as noted above, the "place" that defines a farm need not be contiguous. Likewise, an operation that raises livestock or fish does not necessarily have all of its barns, ponds, and so forth on the same site. A farm may engage in relatively minor activities in other industries, such as providing support services to agriculture or being involved in retail. Farmers sometimes use the terminology, "different enterprises," if they produce more than one type of crop or livestock within the farming operation; for example, corn, soybeans, and a herd of beef cows or calves could each be classified as a separate enterprise.[17]

Another issue is that a farm, as defined above, is an economic concept. That is, a farm might be organized as a complex legal structure for reasons of tax, inheritance, or otherwise. For instance, a father and son might run a farm wherein the former owns most of the land while both own the machinery together. Or some of the land might be part of a family trust. In any case, the use of these different legal entities does not necessarily change the status of a farm from a single farm to more than one farm.

Defining the Farm as an Establishment

Defining a farm as an establishment means that, as measured by a statistical agency, the number of farms as well as the size of the average farm both depend on *how* farmers organize their activities. The farmer who buys out his neighbor, adds that land to his own property, and then rents out the purchased farm building or uses it for his contract work operations or for an agri-tourism business is enlarging his establishment. He (or she) still has one farm, though he also may now have other income sources or may own some other establishments, such as the agri-tourism business. But, if a similar farmer also buys out his neighbor but instead uses the farm buildings for another farm activity—such as when pig breeding continues on his original farm and hog finishing occurs on the recently acquired farm—proper categorization then starts to depend on the way the business is organized.

In the second scenario, if the two farms are or can be separated in a technical and economic sense, such as by having different management accounts each with its own profit-and-

[17]As in general statistics, the term "enterprise" is used for the level of units (firms) that include several establishments, this report does not use the term.

loss account and balance sheet, then there are two establishments. This conceptualization is in line with the day-to-day language: the farmer owns two farms. This can of course easily become a source of complexity in data collection if statistical agencies view such a situation as two establishments even though the farmer manages them as one—or the other way around. The solution is to take the farmer's reality as a point of departure, even if this reality is mainly shaped by the history of the farm or by the legal, fiscal, or risk-management considerations of the farm.

Giving farmers clear guidance concerning the fact that a farm is an establishment and that the data collection aligns with how they see and organize their own farms is important for reducing confusion among respondents about how to report. This implies that NASS needs to be clear to farmers about what "a place" means in its definition of a farm. Where the farm business's establishment and legal entity structures align, opportunities for exploiting linkages with administrative (e.g., tax) data will exist.

The Census Bureau faces a similar challenge, though on a much larger scale, when it measures the number of "businesses" in the U.S. economy overall. Although it never formally defines what a business is in its technical documentation, the bureau nevertheless reports statistics on businesses in the U.S. economy. This is possible only because it uses the word "business" (synonymous with "company" and "firm") to simplify its presentation of the more nuanced concept of an "institutional unit," such as an establishment or firm, which it does describe in its technical documentation: an institutional unit exercises decision-making control and has rights to residual profit; it can be a single-location corporation or LLC, or a multi-location (perhaps multinational) conglomerate with many subsidiary units, each operating under a distinctive legal form of ownership.

The Census Bureau attempts to measure some of this organizational complexity using a mix of administrative and survey data, but of course it cannot capture it all. Two elements of the Census Bureau's approach that may offer helpful guidance for NASS are: (i) the use of a "company organization" instrument to identify and link related reporting units prior to conducting its census enumeration and mailing; and (ii) the sampling of smaller units to free resources for more intensive focus on larger units. We elaborate on this in Chapter 5 in the context of data collection, where it becomes crucial to distinguish between entities from which data are being collected and entities in the statistical structure that we want to measure. The latter should be well defined and set in stone; in contrast, the collection entity may have to conform more to the availability of data as maintained by the businesses.

Refining terms in this way can have an influence on the number of farms that are reported and therefore also on the average size of farms, although it is not clear in which direction. As argued above, most likely the number of farms would increase and the average size would decrease, since this approach invites complex holdings to report separately for each of their component establishments. In judging this advice, the agencies should consider the fact that statistics will also have to report on the number of farming businesses that own more than one farm (see below, and Chapter 5).

Defining the farm as an establishment would be an important step toward improving data collection on complex holdings. It would imply that some farmers do own, operate, or are the residual claimants of more than one farm, instances that indicate a type of complex farm holding, that is, management structures that extend over a combination of establishments. Grouping multiple establishments under a single management unit for the purposes of data collection and reporting is a common feature of measuring business activity outside the agricultural sector: as

mentioned earlier, Walmart is a single firm (business), but the different stores it owns are separate establishments.

If the farm is defined as an establishment, research into the organization of farming, agriculture, and the food chain would benefit from the use of a consistent definition of the firm in the agricultural context. This could be one business that owns several farms (establishments); a business that owns one farm as well other nonfarm establishments; or a business for which the farmer is involved in several farms that have complex ownership and managerial structures. Such situations can increase the complexity of a farming business, at least relative to the classical situation in which the establishment level and the business level are equivalent.

In Chapter 5, we discuss the value of implementing in the measurement infrastructure a definition of business entities based on shared management structures. This includes both cases in which one business owns and operates one establishment (a simple farm business) and cases in which one business owns and operates a group of establishments (a complex farm business). Using this type of definition would facilitate collecting and reporting statistics at both the "establishment" level and the "farm business" level. Chapter 5 discusses how this can be implemented.

Sharecropping and Noncommercial Farming

Having defined the terms *farm* and *farm business*, we end this section by discussing two phenomena that also complicate measurement and reporting. One is the existence of sharecropping, and the other is farming for purposes other than selling agricultural products.

Sharecropping is an institutional form of renting out land that has engaged economists since Adam Smith. If a landlord leases land to a farm owner, that activity is an investment activity, which does not make that landlord (or a company like an LLC acting as a landlord) either a farmer or a farm business (simple or complex). However, if a share-contract is in place, the rent is not charged as a fixed-dollar value but rather is charged in the form of the products produced, that is, as a share of the crop. In this case, the landowner becomes a risk-bearing managing partner, and one could argue that the operation is a (separate) farm and part of a complex farming business.

However, in line with the Farm Service Agency's payment eligibility criteria and the IRS definition,[18] if someone is receiving rental income or farm income, sharecropping arrangements should not been seen as separate farms unless the landowner is "actively engaged in farming." *Actively engaged* here means that all participants, whether individuals or legal entities (such as partnerships or corporations), must provide significant contributions to the farming operation. These contributions may consist of capital, land, equipment, active personal labor, or active personal management, or some combination of these. If it is a management contribution, it must be critical to the profitability of the farming operation, and the contributions must be at risk.

Another phenomenon that requires attention is the question whether a business has to be actively engaged in the sale of agricultural products to be considered a farm. For subsistence farms in developing countries, the answer is "no," but in developed countries it is unusual to include hobby-activities for leisure, like gardening or unpaid housekeeping work, in official economic statistics. From an agricultural policy perspective, activities that are outside markets and not sensitive to agricultural policies are also not very interesting, even if some types of registration can be useful, such as registering all horses in a country for animal health reasons.

[18]See IRS Publication 225, *Farmer's Tax Guide*.

Therefore, for most purposes, it makes sense to define a farm as an establishment that at least has the objective to market its produce. In Europe, Eurostat uses a lower threshold of at least 1 hectare (2.5 acres), and includes farms with less than 1 hectare provided they market a certain proportion of their output or produce more than a specified amount (regionally specified in Euros); this latter threshold would normally exclude household gardens providing incidental sales. The Eurostat method guarantees that farms that do not sell products in a certain calendar year, for example due to extreme weather events or harvest loss or if they store their harvest to sell later, are still considered farms.

4.3. THE FARMER

Now that the terms *farming* and *farm* have been defined and elements of their complexity revealed, we turn to the question, who is a *farmer*? This is a relevant question because researchers and policy makers are not only interested in the farm itself but also in the farmers behind that business. Understanding their decision making, for example regarding trade-offs between food production and environmental or other public aspects, is crucial, and monitoring their well-being and any income problems associated with farming is equally crucial.

Not all persons that work on a farm are farmers. Many farm workers are employees or are hired through an agricultural work agency or contractor, as in the horticulture sector. On some farms, family members provide a helping hand in busy periods or on weekends. In most cases, while all of these persons are active in farming, they are not typically treated as "the farmer" either in commonplace discussion or from a statistical agency perspective. According to the Merriam-Webster Dictionary, a farmer is "*a person who cultivates land or crops or raises animals (such as livestock or fish)*." The Oxford English Dictionary defines farmers as the persons "*who own or manage farms*."

Owners or Managers?

The Oxford English Dictionary definition exposes a quandary for statisticians: Are farmers owners *or* managers? Both categories are relevant to various research and policy questions. While they overlap and in many cases are the same person, on some farms they are different persons, creating additional complexity in data collection and interpretation.

Defining the farmer as the owner of the business entity *and* the person who makes management decisions (or appoints managers to do so) signals the importance of those who are responsible for the operation and who bear all the financial risks. They are the entrepreneurs, the residual claimants, in a view that sees the farm as a bundle of contracts (Allan and Lueck, 2002). Risk-management is an important aspect of American farm policy, implying that this is an important dimension when identifying farmers.

Defining the farmer as the owner of a farm means that there will be some farmers (as owners of a farm) that do not work there and are absentee owners. It also means that there may be family members working occasionally on a farm who will be designated as owners for reasons of income tax, marriage, or with an eye to intergenerational transfer and inheritance taxes. This may be broad but at least it is a jurisdictionally precise criterion, and the distribution of profits would make clear who is and who is not an owner.

In principle, both categories—owners and decision makers—are of interest. They could be folded into one group under the term *producer*, a term that NASS introduced for 2017[19] to indicate any person involved in farm decision making (governance structure), from day-to-day decision making to the work of absentee owners who may only make investment decisions once a year. The term also includes hired managers.

Compared with ownership, there is less clarity with the term *decision maker*: Are these the persons who are involved in day-to-day operational decision making concerning, for example, spraying and environmental management?[20] The term *farm operator* also hints at a level of active working and decision making. Or are these the persons that make the important strategic decisions on the choice of marketing channels or about investments and finances? Or are they the owners, while it is a *farm manager* who "runs the show"? From a user point of view, the appropriate designation of the farmer will depend very much on the type of (policy) research to be informed.

If a farm is a corporation, it could be argued that, from a legal perspective, the corporation is the farmer, and the officers of the corporation (the CEO and others) are the persons who express decisions on behalf of the business. In several ways, however, this seems not to be a very informative approach. A farmer or producer is always a person, even in the case of a corporation that does farming, where most likely at least the CEO is the farmer/producer.

What If a Farm Has Multiple Farmers?

It follows from the analysis above that a farm may be associated with more than one farmer, and similarly with more than one manager and more than one owner. It makes sense that NASS and ERS are now asking respondents to the Census of Agricultural to identify multiple producers at a farm and farming businesses, where the farm has more than one producer. In the 2017 Census of Agriculture and ARMS, respondents may list multiple men and women who have been engaged in making decisions for the operation; they are then asked to provide more detailed information on up to four people per farm. Of those four people, the respondent may choose to identify one or more of them as a *principal operator* or *senior partner* on the census, which is different from previous census surveys, which only allowed one principal operator to be identified. However, the term *principal operator* is being used in the 2017 census for bridging purposes only and will cease to exist in the future. The ARMS survey continues to ask for a single principal operator to be identified for the farm household.

This raises the question whether or not a single person can be designated as most important in operating the farm or in share of ownership. In larger, more complex farm businesses it can be difficult to rank the "most important manager" as more or less important than the most important owner. But even within one of these subgroups of producers, persons can have more or less equal decision-making or ownership rights. In more traditional family farms, there have often been cultural gender issues and generational issues that have led to

[19]See *Publication of Agriculture Census Data on Farm Operator Demographics*, a report by the National Institute of Statistical Sciences Technical Expert Panel, October 12, 2017. There it was recommended to replace the label "Operator" with "Producer" in all publications. The 2017 Census of Agriculture and future censuses use these terms: "All producers," "Principal Producers," "non-Principal Producers." These terms span the breadth of agriculture and are seen as consistent with current terminology used by producers and by professional agriculture organizations.

[20]The 2017 Census of Agriculture and ARMS were broadened to capture farm decision making beyond the day-to-day.

misrepresenting the reality of who is contributing to operating the farm, and in this regard data could very much be influenced by who responds to a survey in any given year.[21] The maximum that can be done is to ask, in ARMS or special surveys, for data on each farmer/producer that identifies how many hours each worked and who was involved in which decisions, but such questions are potentially time-consuming and thus costly.

In the case of multiple operators, during the data collection process currently it is the respondent for the farm who identifies the principal farm operator. This identifier also serves as the link to a household for the collection of data on income and other household variables. Such an approach, however, does not lead to an accurately defined household population. (We return to the issue of the relevant household population in the next section, and then again in Chapter 5.)

It is also clear that, based on the definitions above, a farmer may be involved in more than one farm and even in more than one farm business. In the case of a farmer who owns or co-owns a complex farm business that includes more than one farm, the farm business is a holding company. However, this firm is often a conceptual construct: the presence of a company that oversees the whole structure might be lacking, in practice, even for cases where a farmer owns or co-owns several farms or is involved in LLCs with family members or others. It is these constructions that are especially complex for data collection. For agricultural statistics, in cases like these the thinking should be clear, namely that the farm establishment is often the point of entry from which higher-level structures, like farm businesses, as well as households can be profiled.

Finally, it is worth remarking that the producer/farmer (whether owner or manager) identified in the Census of Agriculture or ARMS is not necessarily the respondent to the survey. Especially in complex farm businesses, the respondent could very well be an administrator, an outside accountant, or another staff member involved in farm management. Conceptually, the issue of who the respondent is should not influence the view of who the farmer is, but we will return to this issue in Chapter 5, where survey designs are discussed, since this issue also has consequences for what can be asked and to whom.

4.4. THE FAMILY FARM AND THE FARM HOUSEHOLD

Policy makers and researchers are not only interested in farms and farmers, but also in farm households, because the household situation can influence the behavior of the farmer and the activity of the farm. A well-known example of this influence is the way investment decisions by farmers over long time horizons, such as 30 years, are influenced by whether or not they have a successor, which affects supply responses to policies (Gasson and Errington, 1993; Calus, 2009). Investment decisions can also be influenced by income from sources other than the farm, because such incomes can reduce the cash flow needed for consumption or even be allocated to farm investments. Agricultural policies are also sometimes justified as needed to sustain the farm family in times of low production or income, making the interest in family households a legitimate question.

[21] In the past, as documented by NASS (https://www.usda.gov/media/blog/2016/01/19/counting-all-farmers-capturing-many-faces-agriculture-2017-census) there has been bias as to who has been reported as "operator," and even with the new questions and approach on the Census it is unlikely one will see a total change, due to the "culture" of agriculture. Although a woman may be the decision-maker on a farm, she may not be reported as such by some families because women are not viewed as "the farmer."

This interest in the total income and well-being of the farmer and the farm household—which factors into ERS mandates—is especially relevant for the family farm. On very large farms and in complex farm businesses, the family dimensions are relatively less important to the functioning and stability of the operation. Such organizations are more like big family firms in other sectors: it is hard to imagine that family income from other sources plays as big a role in the investment decisions made by Walmart as it does in decisions made by small retail businesses.

The *family farm* is a dominant concept in the public mind and in political debates on agriculture. Historically, family farms are what brought most of the United States into cultivation. In 1930, one of the most iconic American paintings (now in the Art Institute of Chicago), was popularly called "Iowa Farmer and Wife," although the painter Grant Wood originally named it "American Gothic," after the unusual window in the building pictured. In reality, the couple in the painting was not a couple, nor was either of them a farmer (the man was Wood's dentist).[22] Perceptions of the family farm in agricultural statistics create a similar distortion: already in 1977, James Bonnen, in his *Assessment of the Current Agricultural Data Base: An Information System Approach,* questioned the concept:

> The idea of the 'family farm,' with all its value and organizational assumptions, constitutes the central concept around which most of our food and fiber statistics are designed and collected. Yet, it has become an increasingly obsolete representation of the reality of the food and fiber sector. … The world has changed and the concept has not (p. 387).

Given that this observation is 40 years old, it seems even harder to adapt an ingrained image, such as "American Gothic," to current reality. Adapting the concept to reality would nevertheless follow a tradition. Reinhardt and Bartlett (1989) point out that the concept of the family farm was originally used for homesteading farms, which had no outside labor or capital nor used contractors, and over time it was broadened to keep up with changes in the organization of farming.

The interest in family farms and farm households raises the question of how to define them. The fact that a farm can have several farmers (producers) means that it can be associated with several different households, although this is not necessarily so: a man and his spouse can both be producers, as can one or more of the children, while they are living in the same household. But just as typically, children or brothers may set up their own households, which results in several households being associated with a single farm. Such households may also contain persons who are not necessarily direct relatives of the farmer: children from an earlier relationship of their spouse, interns, and so on. Meanwhile, some family members, such as children away at college, are only members of the household for part of the year.

All these factors make it useful to have a clear definition of both *family* and *household*. The Census Bureau defines a family as "*a group of two people or more (one of whom is the householder) related by birth, marriage, or adoption and residing together; all such people (including related subfamily members) are considered as members of one family.*" Beginning with the 1980 Current Population Survey, unrelated subfamilies (referred to in the past as secondary families) are no longer included in the count of families, nor are the members of unrelated subfamilies included in the count of family members. The number of families is equal to the number of family households, but the count of family members differs from the count of

[22]Source: http://mentalfloss.com/article/64853/15-things-you-might-not-know-about-american-gothic.

family household members because the latter count also includes any nonrelatives living in the household.

According to the Census Bureau, a *household* consists of all the people who occupy a housing unit. A house, an apartment, a group of rooms, or even a single room is regarded as a housing unit when it is occupied or intended for occupancy as separate living quarters; that is, when the occupants do not live with any other persons in the structure and there is direct access from the outside or through a common hall. A household includes both the related family members and all the unrelated people, if there are any, who share the housing unit, including lodgers, foster children, wards, or employees. A person living alone in a housing unit is counted as a household, and so is a group of unrelated people sharing a housing unit, such as partners or roomers. The count of households excludes group quarters. There are two major categories of households: "family" and "nonfamily."

NASS and ERS use these same definitions for household and family, which are used broadly across the statistical system and should continue to be so used. For NASS and ERS, a family is a group of two people or more related by birth, marriage, or adoption and residing together. A household consists of all of the people who occupy a housing unit that is a house, an apartment or other group of rooms, or a single room, occupied or intended for occupancy as separate living quarters.

Even with precise definitions for farm and family, however, linking the two in the term *family farm* is problematic. First, there are many cases in which a farmer lives alone, unmarried or as a widow(er), and such farms are nevertheless called family farms. More problematic is that an overwhelming number of farms, including those that are part of complex farm businesses, are owned and operated by one or two related families. Therefore, the terms *simple farm* and *family farm* are not synonyms. Indeed, in its publications, ERS highlights the fact that family farms can be very complex, with multiple households sharing in the farm's income.

In practice, data collection from household units can be restricted to the households of farmers who are owners, assuming that the nonfarm activities of salaried managers do not affect farm management. Data collection on those owner-households, for instance regarding nonfarm income, can also be restricted to those persons who are farmers/owners and their spouses, assuming that the nonfarm income of children who might live at home is used for their own personal expenses and savings and not for financing the farm or for reducing the amount of farm income needed for household expenditure. However, this last assumption may be questionable if the child is the potential successor on the family farm.

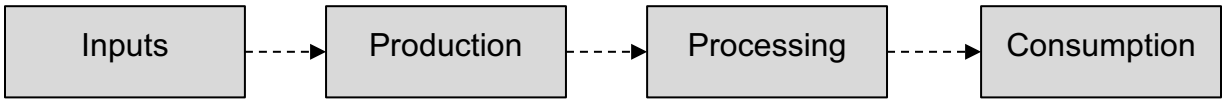
Off-farm income, under this approach, consists of all income from sources other than farming. This may include other agricultural (support) activities, self-employed income, wage income, and capital income, among other things. Neither is it particularly relevant if this income is earned at a location on the farm or elsewhere; whether it is income from a tractor repair shop or bed & breakfast on the farm premises or income earned at a job in the town does not matter.

From the above line of reasoning, it follows that the definition of a farmer is not influenced by his or her income streams from other, nonfarm sources, nor by his or her age. A 75-year-old farmer who receives 75 percent of his income from a retirement pension is still a farmer, and so is the 40-year-old who earns 80 percent of her income from working outside of agriculture or from owning another enterprise. Even if a farmer earns all of his or her income outside farming and makes a *loss* running a farm—perhaps to enjoy the living environment, or to benefit from certain social security payments or tax facilities—he or she is still a farmer (producer).

4.5. AGRICULTURE AND AGRIBUSINESS

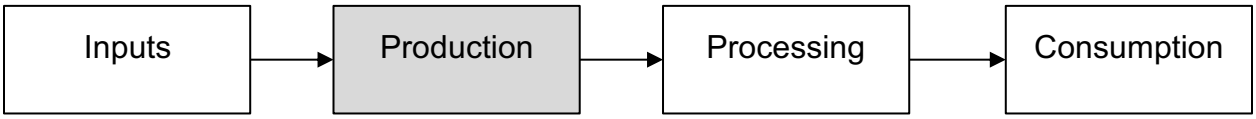
Farming and agriculture are often taken to mean the same thing. In section 4.1, we provided several examples of agencies that use agriculture as a term to describe farming, but we also argued, based on the NAICS classification, that it is useful to restrict the term *farming* to the management of biological processes that excludes support activities to farming (NAICS 115000)[23]—such as harvesting, cotton ginning, and farm management—if they are provided as services by other firms. In other words, if farmers harvest their own grain, it is an activity that is part of farming, but if these activities are carried out by another firm or farm as a service it is a support activity to farming (see Section 4.1).

The essence of the problem, and an important source of complexity in agricultural production, arises from the way farming activities are located within the larger food and agriculture supply chain. Compare the following stylized economies, where gray shading denotes activities undertaken by farmers:



In the most extreme case, farming is a matter of subsistence for household members, and the activities of such farmers span the entire supply chain. Households in such cases are responsible for all steps in the production of farm output, and there is no market-intermediated transfer of intermediate goods between input suppliers, farmers, and processors, prior to final consumption by consumers. If such a household were farming wheat, the members of the household would be responsible for collecting seed and manure, preparing the fields, planting, harvesting, threshing, winnowing, grinding, baking bread, and ultimately consuming their own product. This is a model of agricultural production devoid of any trade or specialization and is perhaps approximated by very poor societies in developing parts of the world.

Less extreme is the case where farmers specialize in one part of the production process. This supply chain (diagrammed below) represents the case in which farm businesses only undertake the agricultural production activities and transfers of intermediate goods that arise from inter-firm transactions. Returning to our wheat example, the farm purchases seed and fertilizer from elsewhere, and the harvested wheat is sold to a flour mill.[24] The rest—all of the planting, cultivation, and harvesting—is undertaken under the direct management of the farm.



Of course, there are definitional challenges in precisely demarcating the distinction

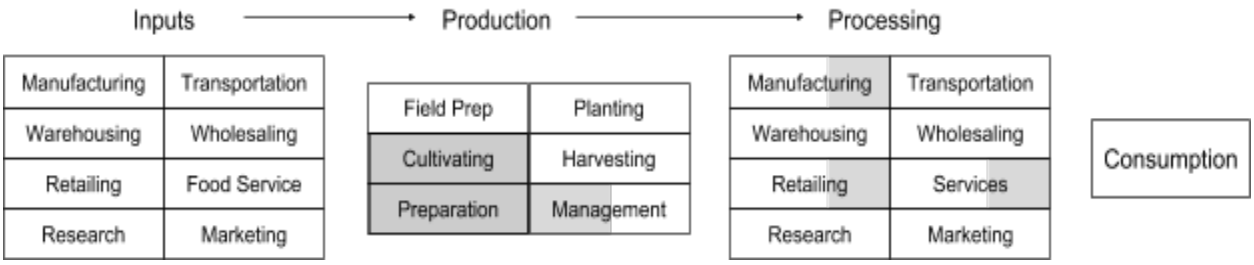
[23]NAICS list them as support activities to Agriculture and Forestry, but if this difference is made between agriculture and farming, they should logically be called support activities to farming.

[24]There may be additional steps: the farmer may sell to a cooperative or dealer who does not process the wheat but sells it to flour mills.

between production and processing. Are threshing and winnowing grain the former or the latter? What about the drying of tobacco leaves? Or the packing of citrus fruit? Nevertheless, for the purposes of discussing complexity in agricultural production, these distinctions may be taken at face value, and under the current NAICS system all the aforementioned examples are classified as agricultural production activities (see Section 4.1).

Agricultural production is constantly evolving, and it has generally increased in complexity over the course of human civilization. In particular, three phenomena stand out. First, trade across economies has existed for millennia, but the breadth and depth of such trade are greater today than ever before and occur at every stage of the supply chain. Second, the number of links within the supply chain has grown tremendously. Third, the boundary of the farm increasingly fails to align with the definition of agricultural production activities.

In the supply chain diagrammed below, the farm business is no longer responsible for all the agricultural production activities:



Instead, in this example the farm hires a specialized service provider who assumes responsibility for soil preparation and planting (NAICS 115112). When the crop is ready, a potentially different service provider is contracted to harvest (NAICS 115114). And while some management tasks are undertaken by the farm business itself, others are delegated to a specialist provider (NAICS 115116).

The increasing decentralization of production activities through the use of specialist service providers is of particular interest because, while these businesses are not farmers themselves, they engage in on-farm production, covering the whole range of activities from soil preparation, planting, and cultivating, to harvesting, packing, and management. More than 50 years ago, the USDA and the Census Bureau recognized the increasing importance of these businesses to the overall farm economy and conducted the first Census of Agricultural Services as a follow-on survey to the 1969 Census of Agriculture. The following paragraph from the 1974 Census of Agriculture succinctly summarizes the motivation for expanding its scope to include such service providers:

> Until the 1940s, agriculture in America was largely self-reliant in regard to many production and harvesting practices now available from off-farm sources in the form of agricultural services. During the last three decades agricultural services have become an increasingly specialized industry. The technological and scientific changes in American agriculture have been directly related to the development of the agricultural service industry. A census of this industry is essential to provide facts necessary for:
>
> a) A broader view of today's farm production.
> b) A better understanding and interpretation of long-term agricultural changes and trends.
> c) A more meaningful analysis of the interrelationships of agriculture and agricultural

services.[25]

Despite these well-intentioned exhortations, the Census of Agricultural Services was discontinued in 1978 and no current Census Bureau or NASS survey program specifically targets these businesses. While a dedicated data collection effort has long since disappeared, the importance of specialist service providers to farm production and their contribution to farm complexity has only continued to increase. Indeed, these businesses fall into a coverage void between NASS, which surveys farms, and the Census Bureau, which surveys nonfarm, nonagricultural businesses. A critical component of the modern agricultural supply chain has simply fallen between the statistical cracks.

This situation is problematic for two reasons. First, key measures of the overall agricultural production sector are increasingly being mismeasured. For example, we currently cannot construct a reliable estimate for the number of individuals employed on farms. From the Census of Agriculture, we know how many workers farms hire and how much they spend on contract labor, but the actual number of farm workers, how much they are paid, and their take-home wages at the national, state, or county level are all unavailable in any current economic survey. The increasing concentration of land and output is a commonly recognized source of intra-firm complexity, but the *vertical disintegration* of farm production activities is another source of complexity that also requires attention.

Second, service providers may themselves be farm operators. The Census of Agriculture currently asks farmers to report related income from the provision of specialist services, but only if these are not stand-alone businesses. Indeed, in the past, explicit dollar thresholds were employed to define farm-related versus stand-alone businesses. Documenting the relationships within farm businesses between the constituent establishments—farm establishments and nonfarm establishments—is key to overcoming the challenge of reporting on complex farms.

To the extent that USDA should be reporting on agricultural production activities in the United States, regardless of the business entity carrying out that activity, these agricultural production activities should, in future, be surveyed.

> **Recommendation 4.3** A program akin to the defunct Census of Agricultural Services, perhaps undertaken as a follow-on survey to the Census of Agriculture, should be developed to collect and report economic activity undertaken by establishments and firms engaged in agricultural production activities through the provision of support services to farms. To accomplish this task, such providers must be identified and included in a Farm Register (described in detail in Chapter 5).

Defining and Enumerating Secondary Activities

The other boundary issue, discussed earlier in this chapter, is that farms may engage in secondary activities that are not farming but rather food processing or retail. Cheese making and produce selling by the roadside or at a farmer's market stall are examples of activities closely linked to farming. Notice that, in the last supply chain diagrammed above, the farm business engages in economic activity beyond agricultural production. As an example, if the farm business is a berry farm, it might process some fruit on-site to produce jams or preserves

[25] *1974 Census of Agriculture, Volume III: Agricultural Services*, Appendix A, Page A-1.

(manufacturing). Some of the fruit may also be used in baking pies and muffins (manufacturing). These items may be sold at an on-site store (retailing) to patrons who visit the farm to pick-their-own berries and enjoy recreational activities (services). It follows from the recommendations in Section 4.2 (and as illustrated with the wine and cheese examples) that these are very often secondary activities of the farm, because they are not organized in a separate establishment; nor do they account for the majority of the work (or value added) in the farm. However, it also follows that, if activities of this kind are organized in a separate establishment or where farming is not a primary activity (such as a garden center that grows some Christmas trees itself), the establishment is not a farm and these are not agricultural support activities. Often, an establishment of this kind will be classified as a food processor or retailer.

These sectoral spillovers mean that a census of farms (farm establishments) is not a census of all farming or agricultural activities, because some of the included farms also engage in a subset of activities outside farming; meanwhile, some farming and agricultural activities are carried out by establishments that are not classified in a statistical framework as farms but instead as agricultural support firms, food processing companies, or retailers. This is the direct consequence of the fact that the Census of Agriculture and ARMS seek to survey organizations and their managers and are interested in the decision making in those organizations.

It is therefore often necessary to extend survey coverage to include these secondary or smaller activities. Farms could be asked how big (and what) these other activities are, and other types of firms could be asked if they engage in farm activities and farm support activities. Based on such an estimate of activities, statistics could be generated to identify the size of farming or agriculture in the total economy.

Two key features of agricultural production in the United States are missing in the supply-chain diagrams above depicting stylized economies. First, the diagrams ignore the central role of international trade. While long-distance trade in agricultural commodities in either their raw or manufactured state is nearly as old as human civilization itself, the world economy has never been more integrated that it is today. Second, the diagrams ignore the relative contribution of each link in the supply chain. The contribution of agricultural production activities to total economic value-added in the United States is now lower than at any point in history. Although food is no less important to human well-being, the sources of value-added have shifted to other links in the supply chain—toward companies like Monsanto, ADM, and John Deere on the input side and ConAgra, Coca-Cola, and McDonalds on the processing side.

In addition to agriculture, an agribusiness complex has been created in the United States and other large economies (Davis and Goldberg, 1957). That also means that agricultural policy, along with environmental policies that target farming, has effects on other sectors than farming and agriculture. To give policy makers and the public insight into these interdependencies, statistics on the agribusiness complex are needed. This can be done with a methodology based on input-output tables of the national accounts (and its satellite accounts) that link farming to activities in other sectors.

> **Recommendation 4.4:** NASS, the Census Bureau, and the Bureau of Economic Analysis should all report on the size of the agribusiness complex and its components in terms of income, employment, and environmental impacts and develop a program that harnesses existing data collection efforts to create a new satellite account for reporting on the food and agriculture industries.

4.6. FARM AND NONFARM INCOME

A farm is an establishment and should be economically identifiable. This implies that any farm either has a profit-and-loss account and a balance sheet or, at least, that the farmer should be able to create them for a survey. In the real world, terms like *income* and *net worth* are not always very clearly defined, and many farmers restrict their accounting to fiscal accounts that satisfy the tax authorities.

Accounting standards as issued by the International Financial Reporting Standards Foundation (IFRS) give clear guidance on how to create a balance sheet, a profit-and-loss account, and a flow of funds for the farm based on accrual accounting. Their accounting standard IAS41 gives clear definitions on valuations in agriculture that ARMS could use. For farm households, an income statement can be added. Off-farm activities generate substantial income and thus contribute to the well-being of farm households. Income at the household level thus may originate from different sources: farm income and nonfarm income. Nonfarm income can be defined as the net income from all nonfarm businesses, wage and nonwage categories.

An important issue in complex farm businesses is that of transfer pricing between different farms and other establishments that are part of the farm business. IFRS accounting standards provide guidance for transfer pricing, but the complexity of this accounting could be a reason to create consolidated accounts for complex farm businesses.

Providing information to policy makers and the public on the real structure of agriculture and the inequalities in the farm sector makes it necessary to provide data not only at the farm (establishment) level but, even more, on the financial situation of complex farm businesses. This means that it is preferable to select farms for ARMS from a register that includes both simple and complex farm businesses. Alternatively, if census data are only available at the establishment (farm) level, ARMS could use these data as a basis for selection, but whenever a farm is part of a complex farm business it should collect data for the complex business as a whole.

4.7. IMPLICATIONS

To find a solution to the issue of data collection on complex agricultural holdings, as specified in Chapter 3, this chapter analyzed concepts from general statistics and accounting. The analysis suggests that complex agricultural holdings should not be fitted into a definition of a farm (counting them all as one farm); nor should such holdings be pressed into a definition imposed on them that fails to recognize the juridical or fiscal organization or the informal management arrangements that are present in such complex holdings.

Adopting a generalized statistical framework and integrating agricultural statistics into it could help solve the problems in data collection and interpretation created by the presence of complex farm operations. Central to this framework is the recognition that a farm business may have either a simple organization, in which it is one establishment—the “classical” farm—or a more complex organization, in which it consists of several establishments.

For the simple farm businesses, not much needs to be changed in data collection. Such farms can be questioned through census forms that query all that NASS wants to know about them, their farm activities, their farmers, and their households. The current structure of ARMS measures land uses, farming practices, input use, and financial statements well. The use of administrative means or surveillance to collect data on these farms will enhance the quality of

the information collected and reduce the burden for the farmer. For data collection from complex farm businesses, the differences between farms and farm businesses and the decisions about whom to ask about farming activities, farmers, and households all have to be taken into account. Farms, farm businesses, and farm households—and even farm fields—each inhabit their own universes that can be represented in registers or list frames designed to capture them. This aspect of data collection is discussed in detail in Chapter 5. The presence of these distinct universes also implies that different data items should be associated with different reporting units.

The main objective of the Census of Agriculture is to report on the structure of farming in terms of the number of holdings, their activities, and their size. That distribution is fairly stable over time, which means that a yearly survey may not be needed for most purposes and that the current five-year interval may be sufficient. As argued in Section 4.2, complex farm businesses could benefit from more guidance in profiling their activities, revenues, land, and assets for the Census of Agriculture. In addition to the structure of farms, this census could also be adapted to report on farm businesses.

Information on yields and prices has to be reported comparatively more often to improve the functioning of markets. However, on this score, experts from upstream and downstream industries who visit multiple farms are often more knowledgeable than individual farmers; and remote sensing technology is another valuable source of information about fields, cropping activities, and yields. As a result, there is a diminishing need to include such variables in the Census of Agriculture.

By contrast, to report on revenues and income for the purpose of carrying out policy evaluations, farm businesses (both simple and complex) are the most important level of analysis and should be a major focus of reporting for ARMS. In addition, as argued above, attention must be paid to agricultural support activities and agricultural business complexes. This increased attention to complex farm businesses will require additional resources, which may be freed up by reducing the number of questions in the Census of Agriculture (especially concerning monetary aspects), by reduced attention to very small farms, and by the deployment of modern IT solutions. Chapter 5 provides additional guidance on these strategies.

APPENDIX 4.1
AGRICULTURAL ACTIVITIES LISTED IN NAICS SECTOR 11: AGRICULTURE, FORESTRY, FISHING AND HUNTING

NAICS code	Type of activity
111000	**Crop Production**
111100	Oilseed and grain farming
111200	Vegetable and melon farming
111300	Fruit and tree nut farming
111400	Greenhouse, nursery, and floriculture production
111900	Other crop farming
111910	Tobacco farming
111920	Cotton farming
111930	Sugarcane farming
111940	Hay farming
111990	All other crop farming
112000	**Animal Production and Aquaculture**
112100	Cattle ranching and farming
112110	Beef cattle ranching and farming, including feedlots
112120	Dairy cattle and milk production
112200	Hog and pig farming
112300	Poultry and egg production
112400	Sheep and goat production
112500	Aquaculture
112900	All other animal production
112910	Apiculture
112920	Horses and other equine production
112990	All other animal production
113000	**Forestry and Logging**
114000	**Fishing, Hunting and Trapping**
115000	**Support Activities for Agriculture and Forestry**
115110	Support activities for agriculture
115111	Cotton ginning
115112	Soil preparation, planting, and cultivating
115113	Crop harvesting, primarily by machine
115114	Post harvest crop activities (except cotton ginning)
115115	Farm labor contractors and crew leaders
115116	Farm management services
115210	Support activities for animal products

SOURCE: U.S. Census Bureau, https://www.census.gov/eos/www/naics/2012NAICS/2012_Definition_File.pdf

5
The Growing Complexity of Farm Business Structure: Implications for Data Collection

A better data collection strategy is needed to improve the measurement of complex farm operations, and that is what we seek to present in this chapter. Particular attention is given to the Census of Agriculture and ARMS, for which a critical aspect of data collection is construction of the sampling frames used to enumerate and survey farm entities at appropriate levels of disaggregation.

Defining the population of productive or income-producing units is a prerequisite to frame construction. Depending on the purpose of the survey instrument, the relevant population may be businesses, land units, or households. The presence of large complex farm operations creates challenges to defining statistical units for observation and reporting, challenges that do not exist to the same extent in the case of simple farms. The *operator dominant* methodology currently used in many USDA surveys and the *operation dominant* methodology used in the Census of Agriculture are compared here with alternative methods for organizing the reporting of farm production and finances.

A second issue addressed in this chapter is the relationship between data collection approaches and potential respondent burden, with the goal of improving both survey response rates and data accuracy. Several actions are recommended to reduce respondent burden, including strategic use of administrative and other nonsurvey data sources. Because USDA currently exploits nonsurvey data sources for other reporting efforts, it is a natural extension to treat complex farm operations within the context of an integrated system of data collection. In fact, doing so would be consistent with efforts across the federal statistical system to increase reporting capacity by exploiting linkages across survey, administrative, and private-sector data collection programs.

5.1. DEFINING THE STATISTICAL UNITS OF FARM, FARMER, AND LAND

A statistical unit is an identifiable element or group of elements that may be selected from a frame when drawing a sample for a survey or census. For a business list frame, the statistical unit is the operation or the operator, recorded in the sampling frame by their name or ID (or both). For an area frame, the statistical unit could be a segment, a tract, or a field. When considering the appropriate statistical unit for measuring complex farm operations, the motivating question should be "what do we want to measure?" At least conceptually, there are three types of statistical units that can come into play, each with a distinct emphasis:

1. The *farm operation*: Institutional unit (later in this chapter redefined as statistical enterprise/establishment)
2. The *people*: Individuals and households
3. The *land*: Farmland, subdivided into fields

As documented in Table 5-1, each type of statistical unit embodies different attributes. Key variables may be best collected using statistical units that are not necessarily the same across different situations. All of these units, however, should in principle be capable of being linked to one another and, in some cases, to additional information. The variables that can be collected from each type of statistical unit can overlap in some cases, such as for land use and production information, but in many other cases the variables are unique and refer only to that particular type of unit.

Table 5-1 Statistical sampling units in agriculture and their attributes

	Institutional unit of farm (enterprise/establishment)	Individuals and households	Land
Agriculture nomenclature	• Farm • Farm operation • Farm business	• Farmer • Farm operator • Farm family • Household of the farm operator	• Segment • Tract • Field
Attributes	• Industry/activity • Location • Ownership • Legal structure • Commodity output • Farm typology	• Individual – Demographic (age, sex, education) – On-farm role – Off-farm activity – Decision making role • Household – On-farm and off-farm activity	• Location (county, state) – Soil and climate – Commodity produced – Land use
Key variables	• Financial information: Income statements and balance sheets • Production information	• Farm and off-farm income, assets, and debt • Farm labor and off-farm employment	• Area/acres • Ownership and tenure • Land use

	Institutional unit of farm (enterprise/establishment)	Individuals and households	Land
Potential sampling frames	• Register of all farms	• List frame, area frame (e.g., Census Master Address File)	• Area frame, list frame of fields (e.g., Farm Service Agency)
Linkages to other statistical units	• Linked to individuals and households through ownership and decision making • Linked to land through geographical location	• Linked to institutional units through ownership and decision making • Linked to land through the geographic location of the farm	• Linked to the institutional unit through its geographic location

Linkages exist between each of these statistical units. For example, ownership, decision making, or employment linkages exist between the business unit and the individuals and household involved, and a geospatial link exists between the business unit and the land. Designing sample frames that maintain reliable linkages between statistical units should be a high priority in a data collection program, because such linkages can be used to indirectly generate representative samples of statistical units across different frames. Using these linkages can allow one to overcome the challenges in creating a complete frame for a single unit-type, which otherwise can be prohibitively expensive or simply infeasible. For instance, a probability sample of farms can be used to define an induced probability sample of households, with its probabilities determined by applying indirect sampling principles (Lavallée, 2007).

As farms become increasingly complex, the traditional perception of farms as self-contained, family-operated businesses no longer accurately captures the contemporary institutional units responsible for the majority of agricultural production in the United States. The statistical and policy-making communities are better served by recognizing that while there is a set of attributes that can be applied to the business part of the operation, there is also a nonoverlapping set of attributes that belong to the operators of those businesses. The remainder of this section examines in greater detail delineations between these different types of statistical units.

Applying alternative definitions of *farm* (an income-generating institutional unit), *farmers* (the individuals and households connected to the ownership and operation of farms), and *farmland* (the locations where production activities occur) may lead to changes in the conclusions reached about important contemporary agricultural policy issues. For instance, an alternative definition of farmland could encompass urban farms. Clearer alternative definitions should also yield more informative answers. Moreover, comparing how these answers change depending on the definition employed is itself an informative exercise.

Some of the challenges currently encountered by USDA exist because of the amorphous statistical unit that arises (in part) by allowing farmers to define the farm and its boundaries. When farms are complex—so that there is no longer perfect overlap between the institutional unit, the household, and the location—this ambiguity makes it difficult for NASS and ERS to accomplish their missions of providing policy makers, researchers, and producers with reliable estimates of agricultural production activity. For this reason, as discussed below, more structure should be imposed on respondents regarding the definition of their farms and operations (two terms that are replaced in later sections with *enterprise* and *establishment*).

To provide this definitional structure, one should build on three broad principles: definitional clarity; recognition of the sampling unit that is best suited to provide particular information; and, as explained above, continuing maintenance of a "crosswalk" that links each type of sampling unit with the others. These principles, summarized by example in Table 5-1, are briefly explained next.

The Farm as an Institutional Unit

Farms may be classified by their primary and secondary industries, using the North American Industry Classification System (NAICS). The farm activities themselves (and the resultant output) can also be classified, using the North American Product Classification System (NAPCS). Although a farm will typically be classified under one primary industry, such as crop production, it may also be involved in other productive activities—including those classified as agricultural production, like cattle ranching, or outside agriculture, like local trucking—that are not in that industry.[1] Other attributes of the farm include its ownership, legal structure, farm typology, and geographic location.

As institutional units, farms are best equipped to answer questions related to the operation of the farm, including its finances, management practices, input usage, and output.

Individuals and Households

A key part of USDA's mission is to provide information regarding the well-being of households involved in operating farms (as discussed in Chapter 2). However, individuals have different attributes from the farms they manage, and the type of information that can be obtained from individuals, as statistical units, is different from what can be collected from a farm as an institutional unit.

The link of an individual to a farm varies from case to case. Different individuals may be involved in different day-to-day decisions regarding the operation. Individuals can be categorized by age, education, and years of experience, and their relationship with the farm may be one of ownership, decision making, work as a paid manager, or as some other type of employee. Moreover, an individual may engage in both on-farm and off-farm business activity.

The key variables that can be collected from individuals or households include on-farm and off-farm income, assets, and debt. Households may also have on-farm and off-farm employment.

Land

Production activity can be tracked for the whole farm, but it can also be tracked by geographical location. A farm's land can be broken down into segments, tracts, and fields; the

[1]One outcome of the classification system is that some of the largest complex farms, in fruit and vegetable agriculture for example, are not listed as farmers; Dole Food Company falls under NAICS 424480 (fresh fruit and vegetable merchant wholesalers) and Premier Raspberries LLC under NAICS 424410 (general line grocery merchant wholesalers). These firms hire farm workers and have contracts with farm worker unions.

latter is the smallest meaningful unit of disaggregation for which information can be reported, for example on the commodity grown, fertilizer applied, or irrigation used.[2]

Attributes of a plot of land include physical location, soil and climate, and water and land use, as well as the commodity produced. In addition to production information, other information that can be collected from a plot of land includes ownership and tenure (if rented), size (acres), and use.

In complex farm operations, the linkages between the farm, the individuals and households involved, and the land are more difficult to track. Understanding these linkages is important for the USDA to deliver on its information commitments (as identified in Chapter 2).

5.2. AREA SAMPLING FRAME METHODS USED BY THE USDA RESEARCH, EDUCATION, AND ECONOMICS (REE) MISSION

The USDA/REE has a well-established sampling frame methodology. By using a combined frame of farm businesses and individuals, NASS and ERS track the linkages between individuals and farms. This structure works well for simple farms, but measurement issues arise as operation complexity increases. In this section, the current USDA/REE approach is described and some of its limitations are presented.

NASS List Frame

NASS uses a dual sampling frame approach for most large scale national surveys and the Census of Agriculture. It uses the list frame, which includes names for both persons and operations, to identify, stratify, and sample operators and operations of interest.

NASS maintains the Enhanced List Maintenance Operations relational database, which includes tables organized by *Person*, by *Operation*, and by *PersonOperation*. Every record in the database is associated with at least one row or entry from each of these three tables, where *PID* (person ID) uniquely identifies a person, *OID* (operation ID) uniquely identifies an operation, and *POID* (person-operation ID) is a unique combination of person and operation. For example, an operation (*OID*) can be associated with multiple persons (*PIDs*), leading to several person-operation pairs (*POIDs*). All of these IDs are used to select the statistical units for the NASS surveys and the Census of Agriculture. NASS uses both operator-dominant and operation-dominant statistical units, depending on the survey. Operator-dominant statistical units are operators reporting for all their operations, and this is used for all multiple frame surveys like ARMS. On the other hand, operation-dominant statistical units are based on operations, rather than individuals, and are used for the Census of Agriculture and the majority of surveys conducted by NASS.[3]

As a result of this structure, a "person" has two roles in the statistical framework. The first role is as a contact for the operation. The second role is as a statistical unit in their own

[2]As noted in Chapter 2, a *field* is defined as a continuous area of land devoted to one crop or land use (according to the ARMS Phase II manual). In the June Area Survey, the one-mile-by-one-mile segment is subdivided into tracts (a portion of the farm that is located in the sampled segment, so equal to or less than the full farm acres), with only one tract per farm, and then into fields. The June Area Survey Data Manual provides additional details on these definitions.

[3]Special codes simplify overlap by maintaining target name and operator dominant classification for individual/partnership, secondary decision maker, decision maker for multiple operations, and large/complex farms.

right. The first role makes sense for supporting a list of farms, but the second role is likely not the optimal approach to use when seeking to obtain a sample of households.[4]

Currently, the list frame is maintained mostly at NASS headquarters, where sample selection is also undertaken. Surveys are typically conducted by the NASS state and regional offices, whose staff modify list frame information based on feedback received from those surveys. Regional field offices are responsible for procuring new list sources and are involved daily in reviewing and updating their list frames. Recognizing that this allocation of responsibilities could create disconnects in sharing and updating list frame information, in 2012 NASS created the Frames Maintenance Group in St. Louis to better centralize maintenance of the list frame. This has led to more consistency and efficiency in the frame's maintenance.

More broadly, there are a range of ways in which NASS's list frame information can be updated: (i) through the Census of Agriculture, as a complete enumeration; (ii) using data from other surveys; and (iii) based on administrative records. NASS is well-equipped to systematically handle the first two sources. Ideally, as detailed in Chapter 6, NASS would have access to a broader set of administrative data, such as tax information, information collected for income-support programs administered elsewhere in USDA, and information from other federal statistical agencies.

Area Frame

The NASS *area frame* is intended to be an exhaustive collection of land use segments. Unlike the list frame, the area frame is, in principle, complete with respect to coverage of the population of farms. However, it is inefficient and expensive to maintain and to enumerate. It is difficult to link a specific plot of land to a farm or farm household through farm operators. One of the main purposes of the area frame has been to provide a way to estimate the undercoverage of the list frame through selecting and enumerating a sample of the land segments, matching the operations on the area frame sample to the list frame and determining those area frame farm operations not on the list. This information is used to create dual-frame estimates for planted acres and agricultural production in NASS surveys and capture-recapture estimates for the census of agriculture. To make these estimates, independence between the list and area frames is a necessary assumption.

Linkages among Farms, Individuals, and Households

The sampling frame for farm households in ARMS is the household attached to the principal operator, defined as the person most responsible for making decisions about the farm operation. The current ERS approach is to use the ARMS sampling weights to expand to the population of farm households from the sample of principal operator households. Even though it is not explicitly treated as such, this is an application of indirect sampling: using a sample of farm operators with known probabilities, the connected principal operators are surveyed and the sampling weights are "inherited" from the farm operator sample. However, because the current sampling frame consists of pairs of farm operators and operations, the implied population of

[4]ARMS is the primary example of when the agency focuses on households. This approach may be efficient for multiple frame surveys, and for ease of "overlapping" the Area Frame to the List Frame for measuring list incompleteness. It should again be noted that the number of entities with "multiple" operations on the list frame is a very small percentage.

farm operators that can be reached through this approach will not always not correspond to the true target population of "farm households."

Additionally, the current sampling approach may not always capture the right target population, because operators with multiple operations have a higher probability of being sampled than do operators with a single operation. For example, a person's household with two farms is twice as likely to appear in the sample as another person's household that only operates a simple business consisting of one farm, but the weights will be the same, assuming that the farms are otherwise identical. These issues can be remedied within the same sampling framework by developing a more detailed set of linkages between the operator or household and the operation units, as will be further described below.

The mismatch between the actual farm household population and the population reachable through the current operator-dominant approach has a number of notable disadvantages, which have already been identified by data users. Currently, information is collected only on the household of the principal operator. If such a designated operator discontinues being involved with an operation from one period to the next, the continuity of information for households associated with the operation may not be preserved.

In addition, the current approach focuses on the principal operator. When constructing statistics derived from these data, a bias is created whereby women and younger generations are underrepresented; principal operators are also older and their households are richer than those of non-principal households. The same information is collected regarding the operator and spouse, so this bias occurs in reporting, not in data collection (that is, ERS only reports on the characteristics of the principal operator).

5.3. ALTERNATIVE MODELS AND METHODS FOR DESCRIBING FARM PRODUCTION AND FINANCES

The challenge of accurately reporting economic activity within complex business structures is not unique to the American agricultural sector. Other statistical agencies, both domestically and internationally, are faced with the challenge of creating organizational systems that can accommodate complex business structures. In doing so, they must balance the need to create a complete, unduplicated list frame of all organizational entities in the population against the ability of respondents to see themselves represented within the statistical structure provided. The following is a review of current practices adopted by international and domestic statistical agencies to accommodate complex organizational structures, designed to lead to more specific recommendations for the U.S. farm sector.

International Guidance on Registers

The United Nations Economic Commission for Europe (UNECE) has published *Guidelines on Statistical Business Registers* (henceforth referred to as the *Guidelines*) with the aim of advising member countries on creating and implementing a standardized approach to the treatment of registers of business enterprises. The *Guidelines* calls for the inclusion of all institutional units engaged in productive activities.[5] These include government units, business

[5]Statistical units are similarly defined in the *System of National Accounts,* 2008 (2008 SNA) and the International Standard Industrial Classification of All Economic Activities, revision 4 (ISIC Rev.4).

corporations, and nonprofit institutions. The *Guidelines* also suggests including household enterprises (both employers and nonemployers) in statistical business registers if suitable administrative sources are available to do so (UNECE, 2015, pp. 27–28).

In addition, the *Guidelines* makes recommendations for a "statistical unit model" that presents a hierarchical structure for representing the activities of institutional units. These units—enterprise groups, enterprises, establishments, and local units—are defined as follows:

- An *enterprise group* is a collection of enterprises linked by ownership or control.
- An *enterprise* is "the level of the statistical unit at which all transactions, including financial and balance sheet accounts, are maintained, and from which international transactions, an international investment position (when applicable), consolidated financial position and net worth can be derived." (UNECE, pp. 40–41). An enterprise may be composed of multiple legal entities.
- An *establishment* is a part of an enterprise in which a single productive activity occurs, or in which the primary activity accounts for most of the value-added of the entity. The statistical unit, at this level, must be able to provide data on production, intermediate consumption, investment (that is, capital expenditures), and employment. An establishment may incorporate a geographic component, such as physical location, as well as a kind-of-activity dimension.
- A *local unit* (location), as described in the *Guidelines*, is a statistical unit in which an activity occurs. If it exists only as a cost or revenue center it is part of an establishment.

The *Guidelines* recommends that statistical business registers track the linkages among these units. Maintaining linkages between the enterprise group, enterprise, and establishment is important as it provides a connection between the location of the activity (establishment) and the legal structure of the entity (enterprise), which further downstream will permit the integration of data, including the potential use of administrative data.

It is important to note that the generic nomenclature covering the statistical treatment of business structures is not necessarily the same as commonly held terminology in the agricultural sector. For instance, usage of the term "enterprise" in a business register is different from the agricultural industry's usage of the term, where it denotes the activity that can occur within a farm, as in the case of a crop or livestock enterprise within a farm.

In this section, descriptions of practices by the Census Bureau, the Bureau of Labor Statistics, and international bodies follow their naming conventions. Recommendations for USDA will use slightly different terminology in order to reduce confusion with agricultural industry norms. Specifically, a *farm establishment* is a business establishment that is engaged in farming. A *farm business* is a collection of business establishments linked by ownership or control that has at least one farm—and corresponds to an "enterprise" in statistical usage.

> **Recommendation 5.1:** USDA should consider adopting definitions of (1) *farm establishment* as a business establishment engaged in farming and (2) *farm business* as a collection of business establishments with at least one farm establishment linked by common ownership or control.

The *Guidelines* includes recommendations specifically directed at the agricultural sector. One recommendation is that agricultural production be included in statistical business registers.

There is a recognition that some countries may organize separate farm registers, as is current practice in the United States. However, the *Guidelines* raises a concern about maintaining a farm register distinct from a statistical register of nonfarm entities, because doing so can make it difficult to maintain consistency of coverage across different economic surveys (p. 29).

The *Guidelines* also recognizes that a farm does not always correspond to an enterprise. A farm may occur within a complex legal entity that could include other activities. Similarly, a farm may involve more than one legal entity. Finally, it is recommended that agricultural household (i.e., unincorporated) enterprises be included in statistical business registers—if an administrative source can be found to identify them.

UN FAO World Program of the Census of Agriculture 2020

The Food and Agriculture Organization of the United Nations (FAO) provides guidance and support to countries undertaking agricultural censuses in the *FAO World Program of the Census of Agriculture 2020* (hereafter referred to as *WPCA 2020).* Here, the recommended statistical unit is the *agricultural holding*, defined as a unit engaged in agricultural production under single management.[6] This is very similar to the way USDA defines a *farm* (see Chapter 2).

This standard splits agricultural holdings into those that are in the household sector (owned by household members) and non-household holdings (owned by corporations or government institutions).[7] In the case of non-household agricultural holdings, the establishment, as defined earlier, is recommended as the unit of measure. In the case of households, the household itself is treated as an enterprise with only one agricultural production establishment.

Given that agricultural activity may be a secondary activity for some establishments, the units included in a census of agriculture should extend beyond establishments whose industrial classification is primary agriculture, meaning agriculture is the activity with the greatest value-added. In other words, to have a complete Census of Agriculture, even establishments for which farming is a secondary activity must be included. When creating a farm register, *WPCA 2020* suggests, ideally one should ensure that it (i) contains information about the unit (land, types of livestock, crops, etc.); (ii) avoids duplications and omissions; and (iii) is regularly updated.

Treatment of Business Lists in the U.S. Federal Statistical System

Two federal statistical agencies, the Census Bureau within the Department of Commerce and the Bureau of Labor Statistics within the Department of Labor, maintain distinct business registers using different administrative data sources. Each is described in detail below.

[6]More specifically, the *WPCA 2020* defines an agricultural holding this way: "An economic unit of agricultural production under single management comprising of all livestock kept and all land used wholly or partly for agricultural production purposes, without regard to title, legal form or size. Single management may be exercised by individual or household, jointly by two or more individuals or households, by a clan or tribe, or by a juridical person such as a corporation, cooperative or government agency. The holding's land may consist of one or more parcels, be located in one or more separate areas or in one or more territorial or administrative divisions, providing the parcels share the same production means, such as labor, farm buildings, machinery or draught animals" (p. 43).

[7]A household is an individual person, or a group of individuals that live together for the common provision of food or other essentials.

The Census Bureau

The Census Bureau maintains the Business Register, a relational database that links data from administrative sources and survey products.[8] As with NASS farm list frames, the Business Register combines data from multiple sources "with the goal of providing comprehensive, accurate, and timely coverage of business units" (NASEM, 2018, p. 34). The statistical structure of the Business Register follows the enterprise-establishment model, which is similar to UNECE's *Guidelines*. The *establishment* is the economic unit that is usually a physical location where activity occurs. It is seen as the smallest, most discrete business unit. A *firm* is a collection of establishments linked by ownership or control; it consists of a top parent company and all of its constituent establishments. As in the case of a simple farm, for a single-establishment business the firm (or enterprise) and the establishment units are the same. Larger firms, on the other hand, may consist of hundreds or thousands of establishments (NASEM, 2018).

As the smallest business unit, each establishment in the Business Register requires a unique *survey unit identifier*. The primary data sources for constructing survey unit identifiers are income and payroll tax filings that the Internal Revenue Service (IRS) shares with the Census Bureau in accordance with Titles 13 and 26 of the US Code. Thus, the Census Bureau faces the following challenge: It must use the tax-reporting behavior of businesses to identify economic activity and then organize that activity within the establishment-enterprise hierarchy.

Within the Business Register, there are four exhaustive and mutually exclusive categories of establishments: (i) nonemployer sole proprietorships, (ii) nonemployer establishments not organized as sole proprietorships, (iii) employer establishments in single-establishment (single unit) firms (SU), and (iv) employer establishments in multi-establishment (multi-unit) firms (MU). For the first three categories, the administrative data provided by the IRS is sufficient for constructing a unique survey unit identifier. For the establishments in multi-unit firms, however, the task is more complicated. Indeed, to maintain the establishment-enterprise hierarchy, the Census Bureau must impose a reporting structure on respondents. For this reason, a somewhat detailed description of the process is instructive.

For nonemployer establishments organized as sole proprietorships, the survey unit identifier can be constructed as a unique transformation of the Social Security Number (SSN) of the proprietor, as reported on *IRS Schedule C: Profit or Loss from Business*.[9] While one individual may operate multiple sole proprietorships, if the business activity taking place in one proprietorship is unrelated to the business activity at another proprietorship the individual must file a separate Schedule C for each. Thus, the construction of unique identifiers for each establishment is quite straightforward, even when multiple establishments share the same SSN.

For all other establishments, the Employer Identification Number (EIN) serves as the tax identification number. Thus, partnerships, corporations, and cooperatives that do not hire employees can each be uniquely identified by the EIN used on their income tax form.[10] In other words, for nonemployers that are not sole proprietorships, the establishment and the (income)

[8]Detailed information on the Business Register can be found at: https://www.census.gov/econ/overview/mu0600.html, and https://unstats.un.org/unsd/trade/events/2015/aguascalientes/10.-Panel%20III%20-%20Presentation%202%20-%20US%20Census%20Bureau.pdf.

[9]Although sole proprietorships that pay excise taxes or operate a Keogh retirement plan must apply for an EIN from the IRS, it is not reported on IRS Form Schedule C.

[10]Partnerships report income on IRS Form 1065; C-Corporations report income on IRS Form 1120; S-Corporations report income on IRS Form 1120-S; Cooperatives and Associations report income on IRS Form 1120-C.

tax-reporting entity are the same. As a result, any one-to-one mapping from EIN to survey unit identifier will suffice.

A similar situation presents itself for SU employers, although the source of federal tax data is different. Regardless of the legal form of organization, employers must report income, social security, and Medicare tax withholding for all eligible workers[11] using their EIN as a unique tax identification number. Again, the establishment and the (payroll) tax-reporting entity are the same, so the survey unit identifier is easily constructed as a transformation of the EIN.

The SSN and the EIN were created to serve administrative purposes, namely the collection of taxes and the crediting of wage earnings. For nonemployer establishments and employer establishments in single-establishment firms, there is an almost one-to-one correspondence between the tax entity and the business entity.

The one-to-one correspondence can break down, however, when a firm reports payroll information covering multiple establishments on one IRS payroll tax form, so that one EIN is associated with economic activity at more than one establishment. Administrative records by themselves are insufficient for maintaining the establishment-enterprise hierarchy in its Business Register.

For this reason, the Census Bureau supplements the administrative information provided by the IRS with responses to the Company Organization Survey. The purpose of the Company Organization Survey is to link business entities within organizational hierarchies through two simple questions: (i) Is the responding company owned or controlled by another company? and (ii) How many establishments were operated by the responding company? If the answer to the first question is affirmative, the company is asked to provide the EIN of the controlling entity. If the answer to the second question is more than one, the company is asked to provide the location, employment, and payroll information of each establishment.

Through the Company Organization Survey, the Census Bureau is able to maintain a register that applies a consistent establishment-enterprise hierarchy with sufficient information to link establishments to parent firms. The benefit of linkages between establishments and firms is two-fold. First, it allows a consistent definition of the survey unit within the Business Register. Second, it allows reporting and analysis at different organizational levels. For some purposes, establishment-level information may be the most pertinent, but for other purposes firm-level aggregates may be most useful (NRC, 2007).

Having reviewed how the Business Register is constructed from a combination of administrative and survey data, it becomes easier to recognize how its design enables the Census Bureau to maintain the structure of complex enterprises and provide links between a range of entities, including:

- Survey units,
- Employer units,
- Address units,
- SSN units, and
- EIN units.

The benefit of having consistent institutional definitions and linkages from establishments to parent firms comes with costs. Because the tax reporting behavior of

[11]This information is reported for nonagricultural work on either IRS Form 941 (quarterly) or IRS Form 944 (annually). Information for agricultural workers is reported on IRS Form 943 (annually).

businesses does not map cleanly onto the establishment-enterprise hierarchy that the Business Register seeks to represent, the Census Bureau must conduct follow-up surveys—and surveys are costly to administer. In addition, the survey instruments force respondents to report information about business activity according to the hierarchy chosen by the Census Bureau, not necessarily the form easiest for the business to provide. Recognizing these costs, it is worth noting two important aspects of the Company Organization Survey: it is short and its directions are clear.

In order to support statistical activities such as surveys and administrative data linkage, the Business Register should contain core information on the following:

- Unique identifier for each unit,
- Business name,
- Address,
- Operational structure (providing linkages between farm establishments,
- Size and activity measures (important for stratification, exclusion of very small units, etc.):
  - Payroll and employment
  - Revenue
  - Assets and liabilities, and
- NAICS classification (important for tracking nonfarm establishments that are part of a farming enterprise).

Beyond these items, which are essential for maintaining a farm register used as the frame for survey sampling and for the Census of Agriculture mailing list, the following additional items should be added if feasible:

- Legal structure, and
- Tax status.

The inclusion of this information allows the Census Bureau to tailor receipt of survey instruments in ways that minimize total respondent burden. For example, the Economic Census is not a true census because smaller establishments, defined as those below certain size thresholds, are only sampled. Moreover, the value of these thresholds varies by industry.

Finally, the Census Bureau regularly updates the information contained in its register. For single-establishment firms and EINs, this is a continuous process as new information is shared from the IRS or data is processed from census survey products. For multi-establishment firms, updating occurs on an annual basis as the Company Organization Survey and Annual Industry Surveys are processed.

A recent CNSTAT report on reengineering the Census Bureau's annual economic surveys (NASEM, 2018) raises some issues concerning the Business Register that are also relevant for the NASS list frame or for a potential Farm Register (described later in this chapter). The report argues that the Census Bureau's Business Register would be more useful as a sampling frame, for the annual economic surveys and other purposes, if it "included information about special reporting units that are used for one or another of the surveys" (p. 36). As with farming, large business enterprises in most sectors account for a large percentage of total employment and output, so accurately measuring their activity is especially important in the

production of reliable statistics. With this concern in mind, the authors of the CNSTAT report (NASEM, 2018) offered the following recommendation:

> The Census Bureau should establish a centralized and coordinated Account Manager Program [in which analysts are assigned responsibility for a specific set of enterprises] that serves as a single point of contact for the largest enterprises with respect to all Census Bureau economic and business data collections. Account managers should have as their primary responsibilities not only the population of the Business Register with up-to-date information about these companies, but also the coordination and facilitation of company responses to the Census Bureau's economic surveys over the course of the year (p. 41).

This guidance would be equally applicable to the maintenance of a Farm Register, as described later in this chapter.

In addition, the Census Bureau, through collaboration between external researchers and Census staff, has linked Business Registers across time into a Longitudinal Business Database and associated Business Dynamics Statistics.[12] Refined intertemporal and firm-to-establishment linkages are the primary difference between these data and County Business Patterns.[13] The USDA could benefit from a similar approach.

The Bureau of Labor Statistics

The Bureau of Labor Statistics (BLS) also maintains a continuously updated business register, but does so using an entirely different administrative data source. Rather than federal income and payroll tax information shared from the IRS, the BLS builds its register from information it receives from state unemployment insurance programs. Despite these differences, the actual structure of the register is remarkably similar, contains much of the same information, and faces the same problem of maintaining the establishment-enterprise hierarchy. Just as a firm may report payroll tax information to the IRS for multiple establishments under one EIN, firms may report unemployment insurance information to the relevant state agency for multiple establishments under one account number. The BLS has its own survey instrument to separately allocate firm-level aggregates to individual establishments, namely the Multiple Worksite Report.

One key difference between Census Bureau and BLS taxonomy is that the latter typically associates an establishment with one activity, generally in a single location. If there is more than one activity at the same location, the BLS creates two different establishments as long as payroll records can make that feasible.

Because the registers for the Census Bureau and BLS are based on different administrative programs, their coverage differs in sometimes important ways. Most obviously, establishments without employees do not participate in unemployment insurance programs. As a result, the BLS register does not include nonemployer establishments, while the Census Bureau register does. In addition, because the unemployment insurance participation threshold for employers of agricultural labor is higher than for employers of nonagricultural labor, the BLS register only captures about 50 percent of agricultural employment.

[12]See https://www.census.gov/ces/dataproducts/bds/.
[13]See https://www.census.gov/programs-surveys/cbp.html.

Options for the Farm Register

The existence of multifarm, multibusiness operations, along with the complexity of the management and decision making structure of these businesses, as described in Chapter 3, require modifications to the current list-frame approach.

> **Recommendation 5.2**: NASS should expand on its list frame to create a Farm Register that provides an ongoing enumeration of all farm establishments in the United States.

This Farm Register would be similar to the current NASS list frame, but it would focus on the enumeration of farms as *establishments* and their characteristics while maintaining links to the farm business (the statistical enterprise) that the establishments are part of. The register would be an "evergreen" product and would be regularly updated as new information becomes available. Survey-specific list-frames would be drawn from this Farm Register at a single point in time to support individual statistical programs, including the Census of Agriculture and ARMS. As additional information on farms is collected through ongoing activities, the Farm Register would be updated. However, a list-frame drawn from the farm register would be a stand-alone product over the course of the collection activity and would not be modified.

Given the existence of a one-to-many or many-to-many relationship between individuals or households and farms, the Farm Register could be structured as a set of relational databases that track the relationship between statistical structure, legal structure, and households.

> **Recommendation 5.3:** The Farm Register should be consistent with other business registers in the federal statistical system. This register should maintain a linkage between statistical units, administrative units, reporting units, and household units.

The design of this Farm Register should be informed by the planned redesign of the Census Bureau's Business Register.

Institutional Units

The Farm Register should follow an institutional unit (enterprise-establishment) structure similar to that of the Census Bureau's Business Register, and unique IDs should be assigned to each of these units.

A *farm establishment* would be the smallest unit that could report agricultural production, including revenue, expenses, and employment. It would have an industrial classification, corresponding to its primary activity; however, secondary activities would also need to be identified. The Farm Register would need to pay closer attention to the household sector (as non-employers) than the Census Bureau register does, given its importance to the agricultural sector.

A *farm business*, which would correspond to a statistical enterprise, would be a collection of farm establishments linked by ownership and control.

Table 5-2 summarizes the distinctions between administrative units, reporting units, and household units, which are described next.

Administrative Units

The administrative unit corresponds to EIN account(s) or the SSN account (in the case of a proprietorship) associated with the farm business. Acknowledging the current legislative and regulatory constraints that prevent the sharing of administrative data across agencies, if it became possible to share these data the Farm Register could then be used to facilitate data linkage. In this respect, it would be beneficial for the Farm Register to be designed to allow for the use of EIN and SSN units in a manner similar to that this is done at the Census Bureau.

Reporting Units

The reporting unit is the entity or element about which information is to be obtained (in contrast, the sampling unit is the person responding to a survey).

> **Recommendation 5.4:** NASS should be more prescriptive in the designation of statistical units but maintain the flexibility to collect from a reporting unit that best suits the respondent. It would be the responsibility of NASS to allocate the reported data back to the statistical unit.

Using the farm establishment or farm business, as defined above, may not be optimal (from the respondent's perspective) for reporting survey or census data. Statistical units may be agglomerated or split into different reporting units to ease respondent burden, but the collected information will be allocated or agglomerated to correspond to the farm establishment or farm business, as more strictly defined.

Household Units

Given the requirement to produce statistics on the financial well-being of farm households, it is important to draw a link between the farm business structures and their operators. The Farm Register would contain information on the linkage between the statistical units (farms or individuals) and households (as described in Section 5.3).

Table 5-2 Relationship between statistical, administrative, reporting, and household units

	Description	Purpose	Linkage with other units
Statistical units	A prescribed consistent representation of the statistical structure of the farm (i.e., the enterprise-establishment structure used in the Census Bureau)	Provide a consistent and unduplicated representation of the structure of American farms. These are the sampling units for USDA surveys	*See below*
Administrative units	Representation of the structure of the farm from the viewpoint of associated EIN (or SSN) accounts	Provide the capacity eventually to link administrative data to survey data, or to provide micro level coherence	The top level of a farm business will have a corresponding set of EINs or SSNs (in the case of proprietorships)

	Description	Purpose	Linkage with other units
		with the Census Bureau or BLS	
Reporting units	A group of statistical units that are combined or split to aid the respondent in reporting information	To aid respondents who may have difficulties, or face significant burden, in reporting for each statistical unit	Reporting units are linked to their constituent statistical units, so that reported data can be allocated to each
Household units	Households associated with farm producers	Provide a link between the household and the farm business	Producers should be responsible for some aspect of decision making on the farm, so there should be an ownership or employment link with the statistical units

Because the current sampling frame creates considerable complexity when dealing with farm operations that are more complex than single-unit farm establishments, the proposed approach would simplify sampling by maintaining separate lists of farms, farm operators, and land holdings, such that a sample unit may be selected that is optimal for the measurement of that characteristic. For instance, information on off-farm income is best obtained from a household-type survey rather than from a survey that targets farms.

> **Recommendation 5.5:** All farm establishments in the Farm Register should be linked to a farm business. In most cases, farm businesses will include only one farm establishment, but they may include more than one.

The following information should be maintained in the Farm Register for each farm establishment:

- Primary NAICS of the farm establishment,
- Commodity output flags (NAPCS),
- Name and address of farm,
- Other geolocation indicator,
- Size indicators (sales, number of employees), and
- Linkage variables (e.g., EIN).

The purpose of the agriculture statistics programs in NASS and ERS is to cover all farm activity, regardless of the industry of the statistical unit. The farm register may therefore contain enterprises and establishments that do not have agriculture as a primary activity. If an enterprise is primarily engaged in processing farm products but also operates its own farms, although most of the value-added could be associated with the processing and the enterprise is thus classified as manufacturing, the farming activity still needs to be captured. In order to maximize coherence across the federal statistical system, it would be ideal if a joint register could be developed with the Census Bureau.

Table 5.3 summarizes the differences in approach between the current statistical sampling frame for surveying farm businesses and the proposed sampling frame.

Table 5-3 Differences between current and proposed approach for a statistical sampling frame

Current approach	Proposed approach
Operation-dominant approach • Statistical units are operations • Selected samples only report for the selected operation • Structure reflects how the operation self-identifies	Institutional unit (enterprise/establishment) • Statistical units are farm establishments • Farm establishments that are related by ownership or control are mapped onto farm businesses corresponding to statistical enterprises • Statistical structure follows an established set of guidelines (e.g., that an establishment must be a profit center)
Operator-dominant approach • Statistical units are identified by the "target name" for the farm or ranch – usually a person's name • The "target name" individual reports for all the operation arrangements he or she is involved in • May include individuals linked to the operation by ownership or employment	Household unit • Includes a roster of the individuals (households) responsible for making decisions for the farm • Will have an ownership or employment relationship with the farm • Characteristics of individuals (members of a household) can be obtained in a second-stage survey
Area frame • Statistical units are segments, tracts, and fields	Land/location • Information on segments, tracts, and fields continues to be collected, but it is important that this information is fed back into the Farm Register and with appropriate linkages to institutional and individual units

Sampling from the Farm Register

The proposed Farm Register differs from the current list-frame approach in that it provides a more prescribed structure for the collection of statistical units, while allowing for flexibility in combining units for reporting (collection entities) and eventual linkage to administrative data. Figure 5-1 demonstrates, in its simplest form, how these linkages would occur in the case of a farm business that controls a single farm establishment. This statistical structure is owned by a single legal entity and has an associated SSN (if sole proprietorship) or EIN (if incorporated or an employer).

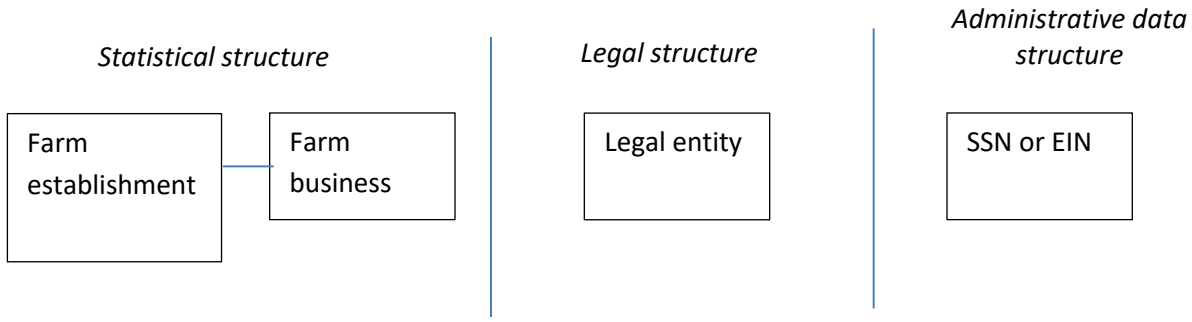


Figure 5-1 Linkages in the proposed sampling frame for a single-farm establishment

The statistical structure can be made more complex with the addition of multiple farm establishments to the farm business. This is depicted in Figure 5.2.

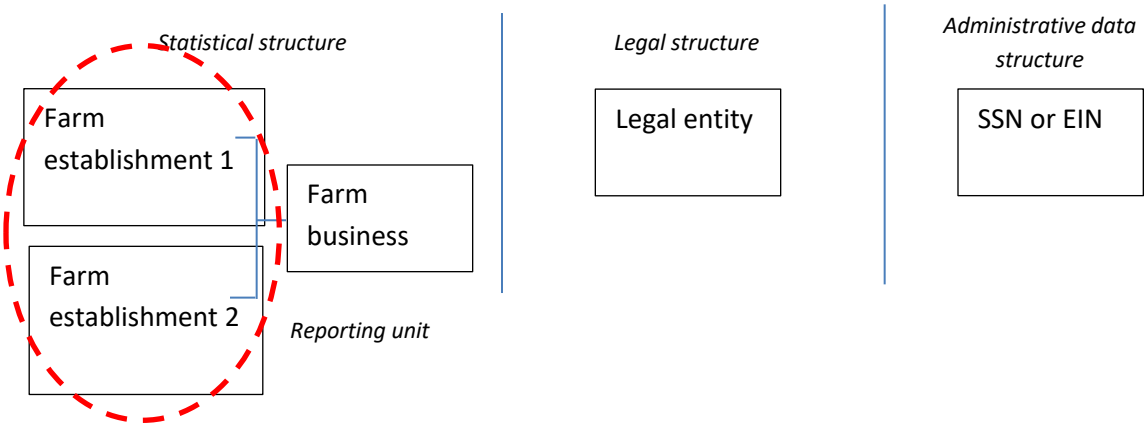


Figure 5-2 Linkages in the proposed sampling frame for a multiple-farm establishments

In the more complex structure illustrated in Figure 5-2, the farm business organizes the farms as two separate profit centers, and each entity that is capable of reporting an economic surplus is identified as a separate farm establishment. If it eases reporting for respondents, these farm establishments could be combined, only for the purposes of reporting, into different reporting units or collection entities; however, reported data would still be allocated to each individual farm establishment.

There may be added complexity if a farm business is engaged in both farm and nonfarm activity. If the nonfarm activity, such as value-added activities like trucking, keeps a separate set of books, then a separate nonfarm establishment should be created in the Farm Register. The farm business may also create a separate legal entity. The legal entities may or may not correspond to individual establishments or farms, but the consolidation of the legal entities should be the same as for the farm business (see Figure 5-3).

In this last instance, all of the activities of the Farm Business should be captured in the Farm Register. This is important if there is to be eventual use of additional administrative records, which may be collected on a legal entity basis but would then need to be allocated within a complex statistical structure.

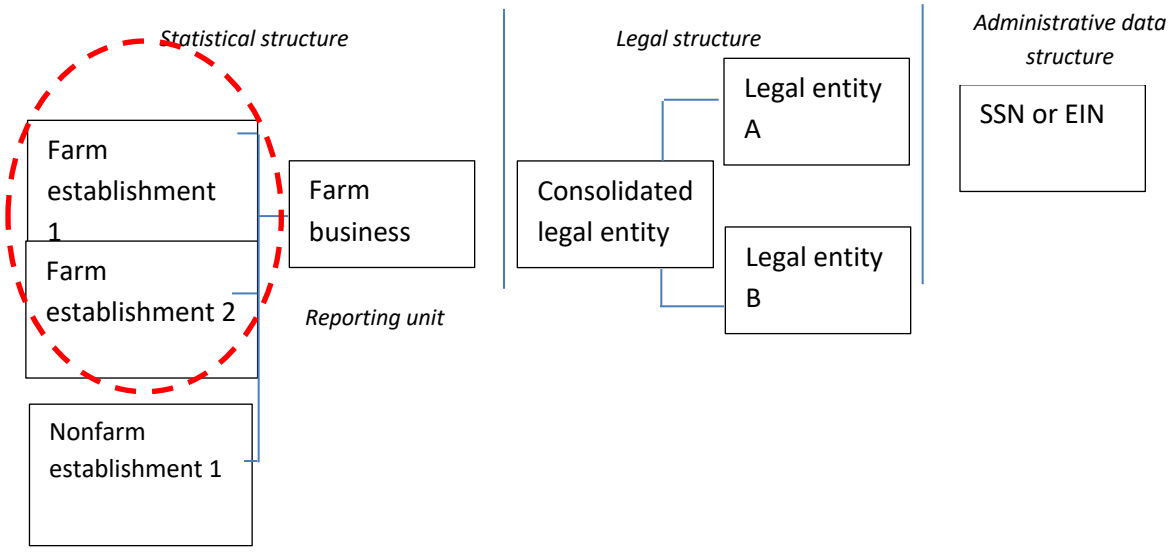


Figure 5-3 Linkages in the proposed sampling frame for a business that combines farm and nonfarm establishments

Treatment of Households and Individuals

The USDA already tracks *some* people associated with farm operations, such as those involved with decision making, employment, ownership, and contacts for surveys. However, this is not done systematically, so there is a lack of coherence about who can be linked. The aim in the past has been to identify a principal operator among all persons involved with the farm. As noted in Section 5.2, this creates an implied population of primary operators who are sampled by NASS but are likely to differ from the true target population of farm households.

Ideally, a household/individual frame would include all operators' households. The operators enumerated are not necessarily the same as the people who should be listed as contacts for survey purposes. Therefore, the enumeration of persons should include their function and an identifier, if possible.

Also, it has become increasingly important to maintain information on all operators in the list frame, including nonprincipal operators. One reason for this is that data appearing to support claims about an aging farm population may, in part, instead reflect a distortion created by sampling frames that fail to represent a younger population and women due to cultural norms regarding who is thought of as the principal operator (Ridolfo et al., 2016). While more careful creation and maintenance of household links within the farm register would in principle be sufficient to sample farm households, using indirect sampling methods, there are important benefits to creating a separate list frame of farm households. In particular, a farm household list frame would make it possible to include additional households or individual-level auxiliary information, and it would allow for a more efficient, more direct sampling of households.

> **Recommendation 5.6:** NASS should create a separate list frame of farm households within the overall Farm Register that would lead to a more efficient sampling of farm households and/or persons involved in farm activities, since the household list itself can be stratified or augmented with auxiliary data.

In so doing, future linkages with tax records or welfare programs could also be facilitated, where appropriate.

Building on the existing operator list frame maintained by NASS, the Farm Register should be a set of relational databases that include information on places and people, identifying households and businesses with suitable links. The Farm Register could also contain household identifiers that could be linked to this frame.

There are benefits with such an approach. For instance, currently ARMS cannot sample households, something ERS should do given its mission to track farm household income and well-being. With this proposed database, ERS could do this. In the case of complex farm operations, it would also simplify the understanding of coverage when the relationship between households and farms is many-to-one or many-to-many. Further, this approach would improve the continuity of operator and household records, and it would also address current issues when the primary operator changed, especially in cases of spouses, two generations of operators (co-principal operators), or partnerships.

5.4. PROPOSED MODIFICATIONS TO THE CENSUS OF AGRICULTURE AND ARMS TO BETTER ACCOUNT FOR COMPLEX FARMS

The panel found that the diversity of topics and conceptual units involved in the Census of Agriculture and the ARMS—concerning businesses, individuals, and households—confuses and burdens respondents, particularly for large complex farms. Adopting a Farm Register, as described in the previous section, would allow ERS and NASS to take a more streamlined approach to both of these surveys, reducing respondent burden and improving the quality of the data collected while still fulfilling all mandates and essential needs.

The panel acknowledges that any major changes to the Census of Agriculture or the ARMS occur only after careful reflection on what is lost and gained by the changes. In the following, we describe one series of promising changes.

Census of Agriculture

The Census of Agriculture should be recast as a source of basic structural characteristics that creates a sampling frame for more focused surveys and that creates reliable small-area estimates for such characteristics. The new Census of Agriculture would be a survey at the farm-business (statistical enterprise) level, enumerating all farm and nonfarm activity that occurs within the farm business. It would identify all farm establishments within each farm business, the producers associated with each farm establishment, and the households containing the producers. The information collected in the Census of Agriculture would update information collected on the Farm Register and could thus be used to draw more targeted sample surveys.

NASS and ERS would have to decide what characteristics are needed for sampling or for small-area estimates. The essential characteristics could include:

1. *Farm characteristics*: Rent land? Irrigate? Use hired labor? Participate in government programs? Family or nonfamily farm?
2. *Producer characteristics*: Demographic information, primary occupation, education, households the producers belong to.

3. *Production characteristics*: Total sales or value of production, production quantities for particular crop and livestock categories.

Items that could be left for more specialized surveys include these:

1. Values for items that are not essential characteristics of the farm (e.g., revenues from horse breeding and stud fees),
2. Practice-related questions, unless linked to the focus of a follow-up survey (section 24 in the 2017 Census of Agriculture),
3. Machinery and equipment details (section 28),
4. Production expenses (section 30), and
5. Income from farm-related sources (section 32).

The Agricultural Resource Management Survey

The Agricultural Resource Management Survey (ARMS) and its various phases and versions should be reformulated into an annual component along with specialized, periodic components. The annual component should be a farm establishment survey, which would collect the information needed for measurement of the costs of production and the financial health of farms, including the information needed by the BEA for national economic statistics.

Periodic, specialized surveys could be used for any questions not needed for these purposes or to meet mandates that explicitly require annual collection. Linkages and comparisons between Census of Agriculture records at the farm business level and ARMS records at the farm establishment level should remain possible, using information from the Farm Register. This is particularly important in calculating the number of farms, which would now be measured at the farm establishment level.

Household information should be collected in a periodic survey of producer households. Conducting a household survey every three years, for example, would allow ERS to fulfill its responsibility to report on the well-being of farm households. And by maintaining a link between households and farm establishments in the Farm Register, it should be possible to link the operational characteristics of the farm establishment with producers and the associated households. To reduce respondent burden, ERS and NASS should reconsider which households it collects financial information from. This includes reevaluating the merits of collecting household information from households with little involvement in production agriculture or from households involved in very large operations with complex and diffuse ownership.

> **Recommendation 5.7:** The USDA should consider alternative strategies for collecting information to meet its mandates. The Census of Agriculture could be revised as a farm business survey, and the ARMS could be revised as a farm establishment survey with linkages between farm businesses and farm establishments.

Other information currently collected in ARMS but not needed annually could be collected in specialized, periodic surveys, which could target farm establishments, farm businesses, or fields. These could include focuses on pressing topics, such as antibiotic use or seed technology, or on more general farm topics, such as production practices, labor arrangements, sources and uses of debt, or participation in government programs.

6
A Broader Data Infrastructure: Administrative and Other Nonsurvey Data Sources

In the previous chapter, a set of recommendations was forwarded to provide a data management framework that could improve the measurement of complex farm operations by addressing the challenges they create for reporting on the farm economy. Two themes run throughout these recommendations. First, USDA must strive to consistently apply an organizational hierarchy through well-defined structures that allow for linkages both within farm businesses and from farm businesses to land and households. This capacity could be enabled through a Farm Register. Second, respondent burden must be minimized by surveying fewer individuals, asking them to answer fewer questions, and ensuring that questions are more carefully defined.

These themes reappear in this chapter, where we discuss how well-designed registers that generate reliable sample frames for survey products can be paired with other sources of data to improve data quality and utility while reducing respondent burden. There is significant untapped potential for the use of nonsurvey data—from administrative, commercial, and nonstructured data sources—in the production of agricultural statistics for the United States. This is an approach that statistical agencies across the globe are increasingly undertaking and one that is already being used for some purposes within USDA.

While this chapter emphasizes administrative data, it is important to acknowledge that the recommendations in Chapters 4 and 5 also increase the possibility of incorporating commercial and other types of data into USDA's data infrastructure. This is a rapidly developing area, with new data products being introduced continually; indeed, the absence of a detailed discussion reflects the panel's concern that specific recommendations may be particularly vulnerable to obsolescence. Nonetheless, satellite data and commercial data, such as Monsanto's Fieldscript, and web-based platforms, such as Farmobile,[1] have already become valuable sources of information on land, land use, yields, and production methods. Such data may also offer unique opportunities for collaboration between USDA, university researchers, and industry. While our

[1]See https://www.farmobile.com/.

discussion does not address commercial issues explicitly, we believe our discussion of the *framework* necessary to better incorporate existing administrative data can inform how USDA must develop a collection infrastructure to incorporate commercially sourced data as well.

While methods for exploring the use of alternative "big data"[2] sources are being pursued by statistical agencies, the statistical validity of analyses based on them are not yet widely established. In contrast, survey methodology provides a known inferential framework for dealing with questions of data accuracy, representativeness, and confidentiality. For this reason, the primary method for compiling information about the nation's farm businesses and households will, for the foreseeable future, continue to be survey-based, albeit with a growing need for linkages to administrative data. The emergence of information that can be captured from digital sources without intrusion on individuals' time and resources does point to the likelihood that less obtrusive methods of collecting data will continue to grow.

According to a survey of 93 national statistical offices conducted by the U.N. Statistical Commission, respondents were most interested in using big data for "faster, more timely statistics" (88 percent), "reducing response burden" (75 percent), and creating "new products and services" (72 percent). Other reasons included "modernization of the statistical production process" (69 percent) and cost reduction (63 percent) (U.N. Economic and Social Council, 2016). A key finding was that, although many countries are pursing options for exploiting large digital (public and commercial) data sets, "very few have yet been able to actually produce official statistics based on these sources" (NASEM, 2017a).

Nevertheless, there are examples of success in the use of nonsurvey, digital data in producing policy-relevant statistics both within and outside of government. Among these are MIT's Billion Prices Project, which produces price indexes or measures of inflation using online posted prices for goods and services;[3] a U.S. Bureau of Justice Statistics project to use web-scraping to improve statistics on arrest-related deaths; and a Statistics Netherlands project that uses data from road sensors to generate transportation and traffic statistics (Puts et al., 2016). Future improvements to federal statistics will largely depend on the capacity of the agencies to leverage multiple data sources. Private-sector data that are continuously generated have emerged as an information source capable of improving the timeliness and detail of national statistics. In the agriculture sector, demands for small-area estimates of economic activities, for which sample sizes are often inadequate to provide precise estimates directly, will continue to increase—as will the need for model based estimates, and to incorporate massive digital datasets for the purpose.

Because modern farms maintain data about their businesses on own their computer systems, some farmers may prefer to share that data in digital form instead of filling out paper forms. In the years to come, agriculture will continue to transition to digital infrastructure, replacing paper invoices and accounts. NASS and ERS, together with the land grant universities, could take an active role in promoting such a digital infrastructure.

One option, as a small step in this direction, might be to work with agricultural accounting software companies and provide them with the algorithms to reformat their data (and add an extra screen for data entry for some non-accounting data) and make it digitally available

[2]The term "big data" is becoming outdated. As noted by Oremus (2017), we are now taking for granted that "data sets can contain billions or even trillions of observations and that sophisticated software can detect trends in them." http://www.slate.com/articles/technology/technology/2017/10/what_happened_to_big_data.html

[3]It should be noted that the Billion Prices index relies on official (BLS) price indexes for benchmarking purposes.

to NASS and ERS.[4] Doing so, even if only on a small scale at first, could help the agencies learn about potential measurement errors that result from mismatches between software systems and survey questions. Farmers might also find that burden can be reduced with this kind of data collection approach. Even if farmers were unwilling to share their data electronically, the computerized forms could ease the completion of surveys such as ARMS. Further, if these algorithms generate cost and price calculations for products, or average income and balance sheet estimates (as published by ERS), the approach could even be of value to farms not selected for ARMS as it could allow them to benchmark their own results with those against sector averages. Such a benchmarking service could be at the heart of a digital research infrastructure, in which the request or obligation to farmers to fill out surveys develops into a more collaborative relationship.

Recognizing that the farmer is the owner of all of his (or her) data, he (or she) could be asked to sign a consent authorization to transfer a copy of his digital data to ERS. This could also include contributing to data sets that are already available at universities or to commercial benchmark software, such as Farm Business Network or Farmobile. Cooperating farmers could be rewarded for such cooperation by receiving additional benchmark reports, if desired, or by being invited to engage in calculations of the effect of potential future policies ("citizen science"). This approach could be initiated with farms that are selected for ARMS; once the marginal costs of taking in the digital data of an extra farm declines considerably, more farms could be welcomed into the process, since that would improve the reliability of the estimates and support small-area estimates.

Such a digital research infrastructure could also advance standardization in data exchange. AgGateway[5] is an American industry organization that promotes such standardization. Another initiative that could possibly share in such a collaboration might be the Sustainability Consortium[6], which runs a project on Data Landscape Mapping in Agricultural Supply Chains. USDA's Natural Resources Conservation Service is also a potential partner; it maintains a considerable amount of environmental and geographical data at the farm level, including subsidy information that could be forwarded to ERS with the consent of the farmer. In Europe, several Farm Accountancy Data Networks (equivalent to ARMS) have experience with this type of data collection, which involves collaboration with accounting offices, agricultural accounting software, extension services, banks, and supply chain partners. Countries where such collaborations are operative today include Denmark, France, Germany, the Netherlands, and Norway. Their experiences might be inspirational for setting up a project in the United States that helps complex holdings share their data digitally with NASS and ERS instead of having to cope with piles of paper.

[4]The Panel chose to present its ideas here somewhat informally since technology recommendations often become obsolete almost immediately after being issued. The use of electronic records by farms is currently undergoing rapid change, so it would be difficult to pinpoint the future use of particular forms of electronic accounting data at this time. It is likely, however, that NASS would benefit from monitoring developments and experimenting with different, evolving forms of accounting and tax data that could potentially provide information that is similar to what is currently being collected using survey questionnaires.

[5]See http://www.aggateway.org/.

[6]See https://www.sustainabilityconsortium.org/.

6.1. MOTIVATIONS FOR PURSUING ALTERNATIVE DATA SOURCES

Federal statistical agencies face increasing demands to improve the accuracy, granularity, and timeliness of their statistical products while simultaneously reducing programmatic expenditures. Accomplishing these goals requires optimizing the use of data already collected across the federal statistical system, as well as incorporating new data products as they become available from nongovernment sources.

Administrative Data

International and national statistical agencies are increasingly using administrative data to support and supplement existing survey programs. Although definitions of administrative data vary across organizational bodies, all reference the same basic attributes. The Evidence-Based Policymaking Commission Act of 2016 defines administrative data as data that are "(1) held by an agency or contractor or grantee of an agency (including a State or unit of local government); and (2) collected for other than statistical purposes" (CEBP, 2017, p. 9). Unlike survey data collected specifically for statistical purposes, administrative data are typically collected in support of an agency's or other organization's routine program operations.[7] Examples of administrative data include federal tax information, vital records, criminal justice records, and information on participants in a wide range of programs, such as unemployment insurance, Medicaid, Medicare, the Children's Health Insurance Program, the Supplemental Nutrition Assistance Program, and federal student aid.

Administrative data are central to the ability of many government departments to fulfill their statutory responsibilities of program operation, management, evaluation, and oversight. In addition, statistical agencies often draw from administrative data to more efficiently meet their statutory obligations. For example, the Economic Census conducted by the U.S. Bureau of the Census has used a mail-out/mail-back design since 1905. It does not print and mail the entire survey instrument to every address in the United States, as that would be an incredibly wasteful exercise. First, most addresses are not associated with business operations, and thus far too many surveys would be produced. Second, industry-specific questions would apply to only a small share of the businesses that received the survey. Therefore, instead of doing a universal mail-out, the Bureau of the Census (Department of Commerce) uses information collected by the Internal Revenue Service (Department of the Treasury) through the latter's administration of the income and payroll tax system to identify businesses, their addresses, and their primary industry.[8] With this information, a sample frame can be constructed that permits individual businesses to receive an Economic Census survey form that is specific to their own economic activity. This process reduces both the administrative cost of conducting the Economic Census and the businesses' respondent burden.

In addition to supporting the development of sampling frames prior to survey data collection, administrative data can also support survey programs during data collection, for

[7]This definition of administrative data is similar in spirit to that provided by the United Nations: "information collected by public sector offices to meet demands of government regulations" (UNECE, 2011).

[8]The Census of Agriculture adopted a mail-out/mail-back design beginning in 1959 and, prior to its transfer from the Census Bureau to NASS, the construction of an initial sample frame from IRS administrative records, including IRS Schedule F and IRS Form 943, was a key component of the planning process. Today, federal tax data from the IRS are still used to assist in the construction of the NASS list frame, but they are not directly incorporated.

example to flag suspicious or missing records for follow-up investigation. They can also be used after data release for evaluating survey performance.[9] Increasingly, however, statistical agencies are looking to expand the role that administrative data play in their survey programs, moving from support to supplementing, or in some cases replacing, their own survey-sourced data. This expansion is being driven in part by the urging of the Office of Management and Budget (OMB), with broad support from the scientific community.[10] Reflecting these trends, the next section of this chapter (Section 6.2) addresses the specific benefits and challenges of using administrative data to improve the accuracy and efficiency of agricultural statistics in the United States.

Other Nonsurvey Data Sources

Other nonsurvey sources of data, such as commercially produced data, are also becoming increasingly relevant in generating statistics on the economy. For example, experimentation is underway to tap into information collected by credit reporting agencies during the credit card application process concerning individuals' debt and repayments. "Unstructured" data, such as that continuously generated by social media or as a byproduct of internet-based commercial activities, have been used to track prices,[11] employment,[12] subjective well-being,[13] and for a wide range of other research purposes.

NASS is looking to expand its use of nonsurvey data in a number of areas. For example, it is exploring the use of web-based information, from sources such as state and local permits, Facebook and Twitter feeds, and interest groups, in building lists to detect nontraditional entities such as urban farms.[14] And, for the agency's 2015 Local Food Marketing Practices Survey—which was designed to produce statistics on the number of farms that market food directly through farmers' markets or roadside stands—a second list frame of potential local food operations was derived from web-based information in an effort to increase coverage of the population. Although widely available, data generated for nonresearch purposes may suffer from incomplete coverage of the population or may be biased toward particular subgroups in the population. For example, they may be more readily available for less mobile or higher-income persons. On the positive side, such data often provide many more observations, greater detail, or greater timeliness than data from survey or administrative records.

---

[9]A review of administrative data, and the scope, purpose, and principles and guidelines for their use, has been developed by Statistics Canada (http://www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm).

[10]From the 2017 CNSTAT report, *Innovation in Federal Statistics: Combining Data Sources while Protecting Privacy*: "OMB and the federal statistical agencies have engaged in a number of efforts in recent years to facilitate greater use of administrative records for statistical purposes, with the goal of improving federal statistics and facilitating program evaluation. ... Statistical agencies have worked together to identify and document important case studies that demonstrate the utility of administrative data for statistical purposes and have documented difficulties in being able to access and use administrative data (see Prell et al., 2009). To address those difficulties, OMB issued a memo to all federal agencies that specifically encouraged the use of administrative data for statistical purposes and discussed the legal, policy, and operational issues with using administrative data (U.S. Office of Management and Budget, 2014a)" (NASEM, 2017b).

[11]The most prominent effort is MIT's Billion Prices Project, which uses web-scraped data from online retailers to track inflation for 60 countries (http://www.thebillionpricesproject.com/).

[12]See https://lsa.umich.edu/lsa/news-events/all-news/search-news/twitter--big-data-and-job-numbers.html.

[13]For example, a "happiness index" has been constructed by analyzing word usage in social medial (https://www.facebook.com/notes/facebook/how-happy-are-we/150162112130/).

[14]See https://www.istat.it/storage/icas2016/f29-young.pdf.

The remainder of this chapter discusses the potential and limitations of administrative data, the current uses to which NASS and ERS put administrative data, and the importance of data linking for the use of these data.

6.2. BENEFITS AND CHALLENGES OF USING ADMINISTRATIVE DATA IN STATISTICAL PROGRAMS

The survey-centric statistical system created during the 20th century is at a crossroads. Despite technological advances in data collection and processing, producing accurate statistics from survey-based instruments has become increasingly cumbersome and costly. Of particular concern to the quality of social statistics is the well-documented decline in survey response rates which, as discussed in Chapter 2, increases the cost of data collection and can decrease the quality of the statistics produced from those data.[15] For example, decreasing response rates have forced NASS to report crop acreage, crop yield, and cash rents for a declining number of counties over time.[16] New approaches are needed, not only to reverse the decline in response rates to the extent possible but also to collect data more efficiently from those farms that do participate in surveys.

A second issue complicating survey-based data collection, discussed at length throughout this report, is the rapidly changing structure of economic activity in key sectors of the U.S. economy. Production activities that were once almost exclusively undertaken by a central business entity are now commonly performed by multiple entities, including specialist providers. For example, instead of hiring and managing its own workers, a farm may outsource field preparation and planting to one operation; insecticide, herbicide, and fertilizer application to another operation; and crop harvesting to a third. The characteristics of outputs produced by farm businesses are also diverse. Farms may process their crop into higher-value products that are sold directly to consumers (manufacturing and retail), and they may operate restaurants, catering facilities, or bed-and-breakfasts (food and lodging services). The boundary of the farm-firm is changing at the same time that the output portfolio of many operations is expanding. Identifying and correctly attributing the many inputs and outputs associated with production on a farm has become increasingly challenging for survey respondents.

A third issue affecting the quality of survey data is the limited budgetary support that federal statistical agencies receive to fulfill their statutory and regulatory obligations. Collecting and processing high-quality survey data to produce accurate and timely estimates of economic activity for use by policy makers, businesses, and citizens is a high-value but nonetheless expensive enterprise.

Administrative data, combined with other sources, offer partial solutions to these problems. Among the advantages offered by administrative sources is that, since their data are already collected as part of program operations, using them creates no additional costs of collection or added burden to the public; using tax data to maintain business frames is one

[15]For a discussion of declining response to social science surveys, see National Research Council (2013).

[16]To address this problem, NASS engaged the Committee on National Statistics of the National Academy of Sciences, Engineering and Medicine to assess county-level crop and cash rent estimates and make recommendations on methods for integrating multiple data sources–including NASS surveys, data from other agencies, and automated field-level information collected by farm equipment dealers—to provide more precise county-level estimates of acreage and yield for major crops and cash rents by land use. See NASEM (2017c).

example. At the same time, accessing administrative data may be associated with costs, with meeting confidentiality and privacy requirements, and with making them fit for purpose.

Achieving efficiencies in the production of one statistic, through the use of administrative data or other nonsurvey sources, frees up resources for the production of other statistics. Administrative data may also be used to improve upon the quality of survey data by reducing variance of and bias due to nonsampling errors, increasing the timeliness of data releases, and facilitating small-area estimates. In some cases—such as when program participation is involved, where respondent recall is a problem, or where quantitative estimates are difficult to calculate—administrative records may be more accurate than survey responses. Even in the absence of falling response rates and budgetary pressures, using existing data more efficiently is still good practice.

> **Conclusion 6.1:** As has been documented in numerous reports—most recently and prominently that of the Commission on Evidence-Based Policymaking (CEBP, 2017)—the use of administrative data can improve the overall efficiency of data programs by reducing agency expenditures, lowering respondent burden, encouraging the sharing of information across agencies, and potentially increasing the accuracy of the information collected. In some cases, administrative data may be used to replace survey data.

Among the disadvantages of using administrative data are the lack of researcher control over content and the need to overcome accessibility constraints.[17]

Using Administrative Data for Official Statistical Reporting

National statistical offices in several countries have been on the leading edge of administrative data use. At Statistics Canada, both survey enhancement and survey replacement are being actively explored. For example, the agency uses tax data to replace data on farm expenditures collected through the Canadian Agricultural Census (Smith et al., 2013). The use of tax data to replace questions on detailed expenses in Statistics Canada's 2016 Census of Agriculture was found to reduce respondent burden by allowing the length of the questionnaire to be reduced by approximately 7 percent. Furthermore, a Tax Replacement Study of Canada found that tax data were also reliable for estimating detailed farm expenses and acknowledged their potential to improve accuracy: "information prepared for tax data might be more thorough and complete" than information reported on the Census form (Smith et al., 2013).[18]

In the United States, both the Census Bureau and Bureau of Labor Statistics (BLS) use administrative data directly in the production of population estimates and in support of programmatic needs. As discussed in Chapter 5, the Census Bureau maintains the Business Register, which serves as the sample frame for its various survey instruments and is constructed from administrative tax records provided to it by the Internal Revenue Service (IRS). The data contained in the Business Register also serve as the primary inputs for the County Business Patterns (CBP) program. The CBP program publishes national-, state-, county-, and zip-code-level estimates of the number of nonfarm establishments with nonagricultural employees, the number of such employees employed on March 12, first-quarter payroll, and annual payroll.

[17]UK Administrative Data Liaison Service, www.adls.ac.uk/adls-resources/guidance/introduction/

[18]There are also instances where administrative data can be checked for accuracy using survey data. See Berent, Krosnick, and A. Lupia (2016) as well as Kreiner, Lassen, and Leth-Petersen (2013).

The BLS operates an analogous sample frame and employment statistics program but relies on a different source of administrative records. The BLS Business List originates from unemployment insurance records provided by state and federal agencies. These are then linked to other administrative records to assign industry classifications. In addition to using these data as the sample frame for BLS survey programs, the BLS tabulates the underlying establishment, employment, and payroll data and publishes them through the Quarterly Census of Employment and Wages (QCEW) to provide industry-level estimates at various reporting geographies (national, state, county, and metropolitan statistical area, among them).[19]

It is worth noting that, while the CBP and QCEW rely primarily on administrative data to construct estimates of the number of operating establishments and employment and payroll numbers, survey data are critically important to both programs. First, the reporting unit for both the CBP and the QCEW is the establishment. To reduce reporting burden, firms with multiple establishments are not required to report payroll withholding and unemployment insurance contributions at the establishment level. Such firms are then requested by the appropriate agency to complete a separate survey that provides employment and payroll information for each establishment separately.[20]

Second, updated information about industry classification is largely collected through survey responses. For example, when a business begins operation, it must register with the Social Security Administration and designate its primary economic activity. This is the source of information when firms enter the Business Register. The Census Bureau updates this information, however, if an establishment reports on a subsequent Census Bureau survey that its primary economic activity has changed. The BLS sends the Annual Refiling Survey to one-third of businesses on its register to provide updated industry classification information.

These two examples demonstrate the value that administrative records offer in producing statistical products for businesses. Because employment and payroll must be reported to various government agencies, processing these administrative records to construct estimates rather than conducting a survey saves valuable resources without increasing respondent burden. Nonetheless, these examples also demonstrate that administrative data often require complementary surveys to ensure that data quality and comparability are maintained.

Administrative data may also reduce respondent burden by decreasing the amount of information that a survey questionnaire must elicit. The Census Bureau is exploring the use of tax records to replace questions about income sources and amounts in household surveys such as the American Community Survey and the Current Population Survey. The results are promising. They suggest that administrative data can in some cases also yield significant improvements in accuracy over survey responses (NASEM, 2017a, p. 25).

Challenges to Using Administrative Data for Official Statistical Reporting

There are significant startup costs for the use by statistical agencies of administrative data. It may take years to test, validate, and automate these sources to ensure their successful incorporation into statistical programs. The statistical agencies of USDA are no exception, as departmental policy dictates that significant vetting is required to understand administrative data

[19]A major reason U.S. statistical agencies must produce two different business registers is that access to IRS data (in this case for BLS) is extremely limited due to laws about confidentiality.

[20]The Census Bureau requests this information through the Company Organization Survey, and BLS requests this information through the Multiple Worksite Report.

sufficiently before integrating them into statistical programs.[21] Since administrative data arrive at NASS with errors in coding, processing, and logic, "quality control efforts, such as time-series and cross-sectional analyses, two or more independent computational cross-checks, and record-level data evaluations, are undertaken before administrative data are used by NASS, particularly where principal economic indicator surveys or market sensitive releases employ administrative data."[22]

There are other challenges as well to the use of alternative data sources. Realistically, these challenges ensure that at least for the very near term survey data will remain at the core of the statistical system. Among the challenges are the following:

- *Administrative objectives often differ from statistical objectives*. Program agencies collect data for the primary purpose of administering their programs; as a result, these sources of information are typically not research-oriented, and data from survey and administrative origins may not be comparable. Stewards of administrative data make decisions that are optimal for their mission and rightfully seek to avoid negative side effects. For example, knowledge among program participants that their records will be used for research purposes may change either their willingness to report or the quality of their reports.
- *Inconsistent concepts and definitions*. Administrative reporting/statistical units, as well as definitions of variables and populations, often do not match those used in surveys, especially for complex farm operations. Surveys often focus on the decision-making entities, while administrative units are typically concerned with smaller parts (such as the field level) of multi-entity operations. Farm Service Administration (FSA) data, for example, are a potentially useful source of data that are redundant with information in the ARMS, but they are reported using the FSA definition of a farm, which differs from the definition used by NASS and ERS. Ideally, to increase the versatility of these data, statistical and program agencies would collaborate on data collection—for example, to harmonize definitions across the agencies, or to add information in surveys, such as the last four digits of the respondent's social security number, to facilitate linkages.
- *Administrative data may reflect biased reporting*. For example, there are incentives to underreport taxable income to reduce tax burden. In support programs, operators may seek to maximize the benefits of participation, which then influences the information they report on administrative forms.
- *Administrative data sets are often characterized by incomplete coverage of a population.* For example, there may be selective participation in administrative programs.
- *Acquiring and documenting administrative data are often problematic*. This is often due to legal constraints concerning confidentiality and privacy.

Due to these challenges, administrative data, while likely to prove increasingly integral to the statistical system, are not a panacea, and at least for the foreseeable future they cannot replace the need for surveys. The influential report by the Commission on Evidence-Based Policymaking (2017) recommends the creation of a federal agency responsible for overseeing the

[21]USDA, Policy and Standards Memorandum, No. PSM-ASMS-15, April 2012. http://nassportal/NASSdocs/Documents/PSM-ASMS-15.pdf.

[22]Edwin Anderson and Daniel Beckler, *Administrative Data Used by NASS,* paper prepared for and presented to the panel (meeting #2).

use of administrative records across federal agencies.[23] This recommendation has not yet been implemented, and if it is it still may not have an impact on practice for several years to come.

6.3. CURRENT, AND POTENTIAL FUTURE, USE OF ADMINISTRATIVE DATA BY USDA

Administrative data are used by USDA for a range of purposes, including survey planning and design, frame construction and stratification, and assessing selection probabilities.[24] This section reviews current sources of administrative data and suggests how usage could be expanded in the future.

Farm Services Administration (FSA)

One particularly rich source of information currently accessible across statistical programs at USDA is the administrative data set maintained by the FSA to implement farm conservation and regulatory laws. FSA data are drawn to estimate production and price statistics for various crop, dairy, poultry, and livestock programs. Of particular importance are the FSA 578 data, which NASS uses to estimate minimum planted acreage indications, calculated by summing acreage for planted, failed, and other status codes, such as double cropping.

FSA data on acreage mixes self-reported data with "determined" data, that is, values taken from satellite images or other sources; and FSA and NASS use different coding schemes for crops. Due to these kinds of differences, NASS reports that "approximately 70 percent of FSA names map to NASS records with few complications . . . the remaining 30 percent require probabilistic record linkage techniques to associate possible matches" (Andersen, 2017, p. 6). The final matches are then manually completed.

Improved linkage accuracy has the potential to greatly improve the utility of these data. In particular, using administrative data to link farm owners to FSA farm numbers could allow for a reasonable method to generate aggregate estimates of key farm economy indicators. This would be useful in its own right, but it would also allow for benchmarking of the Census of Agriculture.

Hurdles exist to the creation of these linkages. ERS research indicates that some data on farm operations are clearly missing and that incorporating additional data (such as crop insurance policies) causes sample selection bias because the added data depend on voluntary participation. In addition, networks of operations or owners captured in administrative data are not stable over time, in part because land moves between them.[25] Nonetheless, these challenges may be no more daunting or deleterious than nonresponse to the Census of Agriculture.

[23]Congress passed the Evidence-Based Policymaking Commission Act of 2016, which created an expert panel—the Commission on Evidence-Based Policymaking—to conduct a comprehensive study recommending strategies for making administrative and other nonsurvey data available for research and policy purposes, while ensuring individual privacy and confidentiality.

[24]In discussing administrative data, it is important to distinguish between administrative data housed at USDA (e.g., subsidy and insurance programs; conservation programs), administrative data at other agencies (most notably, IRS and BLS), and survey data from other agencies (e.g., demographic data at Census, and employment/wage data at BLS). As discussed below, using any of these requires "linkage" with NASS farm and person identifiers.

[25]Presentation to the panel, meeting no. 2, by Steve Wallender.

Federal Tax Information

The use of federal tax information to conduct surveys of the farm economy has changed markedly over time. The Census of Agriculture adopted a mail-out/mail-back design beginning in 1959 and, prior to its transfer from the Census Bureau to NASS, the construction of the mailing list began with extensive use of IRS administrative records. Figure 6-1 below, taken from the 1992 Census of Agriculture documentation, summarizes the initial sources of potential survey recipients before the linkage and validation process.

Preliminary list. The Phase 1 (spring 1992) linkage operation involved approximately 9.1 million records drawn from the following sources:

Source Records Total 9,158,514
National Agricultural Statistics Service (NASS):
- Farms 1,594,125
- Nonfarms 631,274
- Special list 69,627
- Special list (other) 107,603

1987 Census of Agriculture:
- In scope. 1,826,042
- Out of scope 1,534,398
- Nonresponse. 585,810

1990 Internal Revenue Service (IRS) files:
- 1040F (Schedule for Farm Income and Expenses (attached to form 1040 Individual Tax Returns)): 2,242,356
- 1120 (Corporation Income Tax Return (equivalent to standard industrial classification (SIC) codes 01 and 02): 21,152
- 1065 (Partnership Return of Income (equivalent to SIC codes 01 and 02): 67,710
- 941/943 (Employers' Quarterly Tax Returns): 406,772
- Business Master File (BMF—IRS 1120/1065 and 941/943 combinations): 71,645

Figure 6-1 Sources used in the 1992 Census of Agriculture, Phase 1
SOURCE: U.S. Bureau of the Census (1996).

Today, while federal tax data from the IRS is still used to assist in the construction of the NASS list frame, it is not directly incorporated into the frame. Instead, NASS uses it to contact tax filers with agriculture activities not already present in the NASS list frame. As detailed in Anderson and Beckler (2017), usage of these data is governed by three principles: (i) the limited number of NASS staff allowed direct access to actual Federal Tax Information (FTI) data; (ii) a promise of restricting FTI data within a limited-access secure area; and (iii) a promise that FTI data will never be shared with anyone.

FTI is obviously highly sensitive, and the public desires assurances that such data will remain confidential after it is collected. As a result, there are numerous federal laws governing the appropriate use of such data, even for statistical purposes, and the penalties associated with its misuse are severe. Nevertheless, both USDA and the Department of Commerce have a statutory right to request FTI from the Department of Treasury for the purposes of conducting the Census of Agriculture and the Economic Census.[26]

[26]The statutes are as follows:
26 USC 6103(j)(1): Upon request in writing by the Secretary of Commerce, the Secretary shall furnish—

Use of FTI for statistical purposes is not available to all federal statistical agencies. Thus, while the Census Bureau can produce estimates of payroll and employment based on payroll tax filings through the County Business Patterns program, an analogous program undertaken by USDA would not be feasible under current law.

Making better use of FTI to improve list-frame construction and reduce the reporting burden should be a priority within USDA, but it would require modification of federal statutes to extend beyond currently allowed use, which is limited to that required in "conducting the Census of Agriculture." In the statute's language, it is a matter of interpretation whether FTI can be used as an alternative source of household and farm income information within the Census of Agriculture. Nonetheless, ongoing dialogue has the potential to achieve greater collaboration on statistical reporting programs that leverage the statutory authority available to the Census Bureau.

USDA Conservation Programs

Conservation programs also generate administrative data that are useful for statistical purposes. For example, record-level information is available, disaggregated at a scale near the farm field level, since conservation contracts generally apply to individual fields or small collections of fields. Many commodity payments and risk management programs also apply at close to the field level.[27] However, the use of administrative data with common land unit information is highly restricted due to privacy concerns and language in the Farm Act. Currently, linking by researchers must be done on-site at ERS by the agency's Geospatial Information System team.

Summary

USDA currently uses administrative data for statistical and other purposes. However, there is even greater scope for their use within USDA's statistical reporting programs—to facilitate the construction of sample frames, validate data collected from survey instruments, augment existing collection efforts to handle nonresponse or missing information, and contribute to data processing through model-assisted calibration, model-based estimation, and imputation of survey responses. Since administrative data are maintained to support many USDA programs, the scope of these potential applications is vast. A previous report commissioned by USDA and conducted by CNSTAT (NRC, 2008) recognized this potential:

> NASS and ERS should explore the collection of auxiliary information on a formal basis, as well as feasibility of enriching the ARMS data files with information from administrative data sources, geospatial data, and the like (p. 162).

(A) such returns, or return information reflected thereon, to officers and employees of the Bureau of the Census, and (B) such return information reflected on returns of corporations to officers and employees of the Bureau of Economic Analysis, as the Secretary may prescribe by regulation for the purpose of, but only to the extent necessary in, the structuring of censuses and national economic accounts and conducting related statistical activities authorized by law.

26 USC 6103(j)(5): Upon request in writing by the Secretary of Agriculture, the Secretary shall furnish such returns, or return information reflected thereon, as the Secretary may prescribe by regulation to officers and employees of the Department of Agriculture whose official duties require access to such returns or information for the purpose of, but only to the extent necessary in, structuring, preparing, and conducting the Census of Agriculture pursuant to the Census of Agriculture Act of 1997 (Public Law 105–113).

[27]Presentation to the panel, meeting no. 2, by Cynthia Nickerson and Steve Wallender, both of ERS.

NASS and ERS have demonstrated a willingness and capability to successfully expand their usage of administrative data. Following the CNSTAT report quoted above, NASS and ERS responded by participating in an OMB-led initiative to incorporate selected administrative data into surveys. Part of this involved an effort by USDA to synchronize the reporting of administrative (program) data for FSA, the Risk Management Agency, and the Natural Resources Conservation Service in a way that promotes the use of common definitions and reporting. The purpose of this effort was to allow for more direct linking to ARMS and other survey records with the goal of developing agricultural production and conservation.

These efforts have led ERS to conclude that it is possible to use administrative data to support research into complex farm operations. Two key challenges to achieving this goal remain, however. First, as noted earlier, it requires harmonizing the definition of a farm operation sufficiently across the NASS surveys and these sources of administrative data, so that the quality of data linkages can be established. Second, complex farm structures may provide information at a level of aggregation that does not match the desired level of reporting. We have addressed aspects of these challenges in previous chapters of this report. In the next section, we directly address the challenges of data linkage with administrative data sources.

6.4. THE ESSENTIAL ROLE OF DATA LINKING

When adopted for research or statistical purposes, the value added from administrative and other nonsurvey data is often realized when they can be combined with survey data. Record-level survey data may be augmented with information—on income, demographics, geolocation of residence or business, program participation, employment, and potentially many other variables—from administrative records or other sources. Administrative sources often contain data useful for creating descriptive estimates, such as on levels of program support. Other sources may provide supplemental, contextual information about counties and states for subnational-level analyses. For example, ERS has done some exploratory work using county land-value records, by parcel. Many administrative data sources serve needs at the state and local levels of government, which often are the geographical units to which agriculture policies and programs are most relevant.

In applications where federal surveys generate insufficient sample sizes to support local-level estimates, linking to additional data sources may reduce the variance of estimates and feed into small-area modeling.[28] Although nonsurvey data are rarely sufficient on their own to support analyses of farm policy or to evaluate program impacts, they are becoming increasingly essential for filling in key pieces of information.

[28]Small area estimation methods include generalized linear mixed models (e.g., Fay and Herriot, 1979) and hierarchical models (e.g., Lindley and Smith, 1972). The National Academies of Sciences, Engineering, and Medicine report on the use in federal statistics of multiple data sources (NASEM, 2017a) provides a full discussion of linking methods and of the potential benefits of using administrative data for statistical purposes: as a complete frame or supplement to an existing frame for individuals, households, or businesses; to replace surveys when the administrative data contains all needed information; for editing survey responses or making imputations for missing responses; as a source of auxiliary information that can be used to improve survey-based estimates; and for survey evaluation (e.g., to compare the number of program beneficiaries in program records with estimates based on a survey).

Conclusion 6.2: The effectiveness of the federal statistical system to meet future data demands will largely depend on the extent to which data sources—survey and nonsurvey, national and local, public and private—can be combined in synergistic ways. Any redesigns of the Census of Agriculture and ARMS should be done with the presumption that these instruments will need to be linked to other data sources maintained by USDA, other statistical agencies, and even nongovernment organizations.[29]

The key element in the data system for promoting data linkages—for example, between household records and farm business records—is created during questionnaire design. If units of measurement are consistent, then in principle a crosswalk between the FSA data and NASS survey IDs can be maintained. Other options include asking respondents for a limited amount of personal identifying information, such as the last four digits of the social security number, FSA field IDs associated with the operation, or date and place of birth.[30] However, adding questions like these in order to improve linkage rates would require extensive testing; for example, simultaneous impacts on nonresponse would need to be carefully evaluated.

New Opportunities for Data Linking within USDA

Based on experimental research to determine whether discrepancies across sources are meaningful, linkages across USDA data are already reported to be working well, particularly across the Census of Agriculture, ARMS, and the June Area Survey (Young, Lamas, and Abreu, 2017). Data linking efforts are progressing along a number of fronts. Some efforts take advantage of geo-referenced common land units which, by serving as a basic unit for geographically based list frames, allow linkage to FSA and Risk Management Agency administrative records and also to GIS data generated by remote sensing and precision agriculture. A recent expert panel—on Methods for Integrating Multiple Data Sources to Improve Crop Estimates—recommended in its report that NASS adopt FSA's "common land unit" as its basic spatial unit (NASEM, 2017c, Recommendation 2-8).

An ERS project on the northern plains is exploring the capacity for linking geospatial data on soils and cropping history, data that are increasingly available at the field level. The purpose of the project is to study what happens to conservation tillage on fields following participation in a conservation contract or on neighboring fields that are not in such a contract. The geospatial data add key variables to field-level administrative data that can be exploited by researchers. ERS is currently working with the Peterson Institute and USDA's Agricultural

[29]NASEM (2017b), a report initiated to examine the potential of combining data sources for research and policy purposes, discusses the methodological considerations for designing surveys with administrative data linking in mind.

[30]One hurdle to attaching social security or Employer Identification number to a farm establishment or business is that respondents may be reluctant to provide it. Getting the same information from the IRS is fraught with difficulties. NASS sought to avoid this situation by not using IRS information directly in its farm list. NASS does receive some tax information from farm filers and links it to commodity and other lists that the agency receives. For matching records, NASS uses the information from these other lists, so that it is not incorporating any IRS data directly into the NASS farm list. This caution notwithstanding, it should be noted that other National Academies panels (e.g., National Research Council, 2007; NASEM, 2017b) have recognized the advantages—in terms of both response burden and improved quality of statistics—of improving the ability of federal agencies to share data, including tax data. There are certainly practical issues with using these data, but sharing across agencies is a method that has the potential to reduce survey burden, even if the ability to use these data relies on overcoming some legal hurdles and on the uptake of recommendations aimed at all the statistical agencies.

Research Service to develop satellite-based, field-level estimates of conservation tillage and then link these data to administrative and survey data. More generally, small-area estimates, such as for yields or acreage devoted to a particular crop, could be made more accurate and comprehensive by combining survey information based only on a subset of farms with area-level satellite imagery, which may be available for all areas (Cruze, 2015; NASEM, 2017a).

Linking data sources often proves to be challenging in practice. One reason is that, as described above, even among USDA agencies the differences in the definition of the farm and other measurement units create inconsistencies in the way data are reported to NASS, FSA, and the Risk Management Agency. The Panel on Integrating Multiple Data Sources to Improve Crop Estimates reported that "a project in support of the 2012 Census of Agriculture to link the FSA payments database and the NASS list frame resulted in only a 63 percent match rate with 6 percent possible matches" (NASEM, 2017c, p. 38). A NASS study of Nebraska showed that there were 2.4 FSA-identified farms for every NASS-identified farm in the state, and that difficulties in aligning the NASS farms and FSA farms were most acute in the case of complex operations, where there is ambiguity in defining the reporting unit. The available geospatial information on farm operations can help analysts understand differences in the list frames by "making it possible to track down the NASS farms to identify matches." Expensive manual efforts to match farms, according to the same panel report, would best be directed toward achieving matches for the largest farms (NASEM, 2017c).

New Opportunities for Data Linking: Survey and Administrative Data from Other Federal Agencies

Taking full advantage of multiple data sources to transform statistical programs at ERS and NASS will require coordination that extends beyond USDA to include data sources housed at other statistical agencies and beyond. The decentralized U.S. statistical system creates complications that international counterparts can largely avoid. An example is Statistics Canada which, as a centralized statistical agency, has broad powers to exploit administrative data. Across the U.S. system, each statistical agency has its own set of approval, confidentiality, and clearance procedures.[31] No doubt related to these systemic contrasts, along with the differing laws that govern interaction and collaboration between agencies, national statistical offices in Europe report that of all the information they collect, the proportion that originates with administrative data is roughly 80 percent, as compared with 20 percent originating from surveys. Meanwhile, the ratio for the United States is just the reverse (20 versus 80 percent) (Prewitt, 2010). This relatively modest baseline means that the potential for exploiting administrative data for U.S. agencies is relatively large. Indeed, U.S. agencies are now moving more quickly to create opportunities for these kinds of coordinated efforts.

The U.S. Census Bureau, in particular, has already cultivated a significant capacity to link data from a range of sources, a capacity that is being advanced by researchers using data

[31]Prior to 2002, the legislative authority for maintaining the confidentiality of identifiable information collected for statistical purposes was not uniform across statistical agencies. In 2002, the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) was enacted to (i) provide a uniform standard of privacy and confidentiality for statistical agencies to ensure that information supplied by individuals or organizations to an agency under a pledge of confidentiality is used exclusively for statistical purposes, and will not have that information disclosed in identifiable form to anyone not authorized in the legislation; and (ii) promote statistical efficiency through limited sharing of business data among three designated statistical agencies: the Census Bureau, the Bureau of Economic Analysis, and the BLS.

available on a restricted-access basis through its Federal Statistical Research Data Centers (FSRDCs). Titles 13 and 26 of the U.S. Code, which govern much of what is possible at the FSRDCs, provide guidelines for protecting and accessing high-value information about the nation's population and economy. These laws convey the rules for accessing and utilizing records to the greatest extent possible for statistical uses; for supporting reimbursable studies and joint statistical projects; and for protecting confidential individual and establishment data.[32] Section 23(c) of Title 13 allows researchers to be sworn in to access federal data, including data housed and secured by the IRS, the Social Security Administration, and the departments of Housing and Urban Development and Veterans Affairs. Some state data and third-party data are also accessible from within these centers.

Research based on confidential data, whether using administrative or survey data, is only approvable if it supports the mission of the Census Bureau by contributing to improved data quality or the estimation of population characteristics.[33] But the interpretation of mission, and in turn the approvable scope of research, have each been broadened under the FSRDC program, which has established partnerships between federal statistical agencies and leading research institutions; it now serves as a data host in enclaves for other agencies as well.

Through its Center for Administrative Records Research and Applications (CARRA), the Census Bureau has developed infrastructure to help other statistical agencies move forward on record linkage so a broader range of data sources can used for statistical reporting. This has included developing expertise in combining data and in meeting legal requirements and hiring the personnel to write data-use agreements and the staff needed to curate the data. CARRA identifies, acquires, processes, links, curates, and analyzes administrative data, and it creates products that demonstrate the value of data linkage and linked data. It now has many years of experience in identifying administrative sources and figuring out, based on precedent, how to tackle the governance and legal issues that throw up hurdles to their use. It also strives to promote a sustainable and scalable model for accessing a range of high-value, sensitive, and confidential information.

Providing further proof of the multiple-data-source concept, the Census Bureau is currently engaged in a number of joint projects geared toward maximizing the value of existing surveys.[34] Additionally, CARRA is engaged in a longitudinal linkage project with 10 institutions in seven FSRDCs (the Census Longitudinal Infrastructure Project, or CLIP), and 12 pilot projects at Chapin Hall, University of Chicago. The pilot projects are on topics that range widely, from labor market outcomes for public school students in Chicago, to causes of poverty in Cook County, to service utilization by families and children experiencing homelessness, and they are

[32]Title 13 governs the Census Bureau directly, but Title 26 (governing IRS) includes provisions for collaborating with the Census Bureau on statistical reporting. Specifics about these laws can be found at: https://www.census.gov/history/www/reference/privacy_confidentiality/title_13_us_code.html, and https://www.census.gov/history/www/reference/privacy_confidentiality/title_26_us_code_1.html.

[33]The role of the "Predominant Purpose Statement" is described at https://www.census.gov/ces/pdf/Research_Proposal_Guidelines.pdf.

[34]These joint projects are collaborations with the Bureau of Justice Statistics, Bureau of Prisons, Centers for Medicare and Medicaid Services, IRS, Social Security Administration, and departments of Veterans Affairs and Housing and Urban Development, as well as USDA's ERS. Among the surveys identified by Amy O'Hara (in a presentation to the panel on February 10, 2017) that link to administrative or other survey data sources are the Rental Housing Finance Survey, American Housing Survey, the Consumer Expenditures Survey, the National Survey of College Graduates, the American Community Survey, and the Survey of Business Owners.

oriented toward using linked data to enhance evidence-based policy at the local, state, and federal levels.[35]

Growth in the FSRDC system has been enormous during the past five years. In 2010, there were 12 research data center locations, and there are now 28 locations with more on the way.[36] If USDA were fully partnered into the FSRDC program, its research and statistical capacity could be greatly enhanced.[37] For example, person-level data from ARMS and the Census of Agriculture could be linked to BLS employment data, and possibly to tax data, to generate fuller profiles of farm operation entities. The free labor of academic researchers granted access to FSRDC centers could help further develop analytic tools to answer questions about the income of farm households, about sub-businesses of farm households, about on-farm and off-farm value added activities, and about many other dimensions of complex operations. Jumping into the FSRDC "sandbox" would allow researchers to learn more about ARMS respondents who also have records of their farm operations in various business datasets.

> **Recommendation 6.1:** USDA should explore opportunities for record linkage at the person level to obtain information on key demographic and off-farm employment variables, and perhaps with IRS on farm income and expense information. These opportunities can be explored through participation in the Federal Statistical Research Data Centers program, a partnership between federal statistical agencies and leading research institutions that provides secure access to restricted-use microdata for statistical purposes.

NASS and ERS have already developed a data access mechanism in which ARMS data are accessible for statistical purposes through a cooperative agreement with the University of Chicago's National Opinion Research Center. This agreement suits the needs of the agencies and their researchers and is governed by rules established by the Confidential Information Protection and Statistical Efficiency Act. The arrangement works well for those who want to work with ARMS data alone, but NASS does not provide certification of analyses through the National Opinion Research Center and the center does not provide opportunities for linking with data from other agencies. Any proposed expansion in the use of tax data should be accompanied by research to assess producer sentiment toward the idea as well as a campaign to educate producers on why this would be beneficial to them (for example, in reducing burden).

> **Recommendation 6.2:** NASS should pilot efforts to participate in the Federal Statistical Research Data Centers program and identify one or more high-value projects through which USDA researchers could engage with academic researchers and Census staff.

Reflecting the growing importance of drawing from multiple sources of data for policy-relevant program evaluation and research, the report of the Commission on Evidence-Based Policymaking recommended creating a secure digital portal for researchers to use to study the

[35]The University of Chicago, Chapin Hall webpage documents these programs: http://www.chapinhall.org/pages/RFP-Linked-Data-Evidence-Based-Policymaking.

[36]https://www.census.gov/about/adrm/fsrdc/locations.html

[37]It should be noted that use of research data centers for NASS/ERS is not a new idea. A 2008 CNSTAT report (NRC, 2008) recommends that "USDA should consider extending the availability of ARMS microdata through the Census Bureau research data centers to increase access opportunities for using additional data sets and enabling researchers to match ARMS files with other data sets" (p. 157).

impact of U.S. government spending on health care, education, housing, labor markets, and other sectors of the economy. If this recommendation were legislatively enacted, the portal, which the commission refers to as the National Secure Data Service, would mark the next step in the evolution of the Federal Statistical Research Data Centers. The report recommends housing this National Secure Data Service within the Department of Commerce, where the Census Bureau (and, crucially, CARRA) is housed, but with assistance from the 12 other key statistical agencies scattered across the government. The National Secure Data Service would be organized to "temporarily link existing data and provide secure access to those data for exclusively statistical purposes in connection with approved projects" (CEBP, 2017, p. 1). It would build on the infrastructure and expertise already developed at the Census Bureau's CARRA to ensure that data linkages and access to confidential data for statistical purposes are conducted in the most secure manner possible.[38] Additional state-collected data about federal programs would also be made available for statistical purposes: "Where appropriate, states that administer programs with substantial Federal investment should in return provide the data necessary for evidence building" (CEBP, 2017, p. 2).

> **Conclusion 6.3:** Given the work of the Commission on Evidence-Based Policymaking to improve the climate for legislative changes that would make data linking more routine across the statistical agencies, now is the time for NASS and ERS to begin mapping out a strategy to coordinate their survey and administrative data programs within USDA and across other key agencies such as the Census Bureau and the BLS.

One example of how combining data collected across agencies creates new opportunities is in the reporting of off-farm food and agricultural activities, a key data need discussed at several points in our report. To produce statistics on agriculture and the food chain more broadly, as opposed to just on-farm economic activities, the role of data from nonagricultural agencies is crucial. Coordinating output and employment data on farming with data on manufacturing and services could help fill the gaps in our understanding of businesses that operate in close proximity to farm businesses but are not picked up in the Census of Agriculture.

The Bureau of Economic Analysis (BEA) currently reports value-added, employment, GDP contributions, and other statistics for six supplemental (or "satellite") accounts. These accounts combine economic activity across NAICS categories to provide aggregate reporting on sectors defined by criteria different from those used for NAICS.[39] Although it does not have a supplemental account for food and agricultural industries, and it has problems accurately reflecting agricultural sectors with NAICS data, BEA reports on the categories listed in Table 6.1 in its input-output tables which underlie the national income product accounts. The first column contains relevant BEA codes from the IO tables representing the national economy with 15 aggregate industries, while the second column contains relevant codes from the 71-industry IO tables. The final column contains the number of relevant 6-digit NAICS codes from the IO tables used to represent the economy in its most disaggregated form (for NIPA reporting), which includes 389 industries. ERS currently uses data from BEA IO tables for its "food dollar"

[38]The commission's report details promising approaches to quantify the additional risks to privacy associated with record linkages and then set boundaries on acceptable levels of privacy loss. Several methods, such as differential privacy and the use of synthetic data that substitute for values in the original data, are already in use at the Census Bureau as well as at private sector companies like Google and Uber.

[39]See https://www.bea.gov/index.htm.

series.[40] Code 311FT captures many of the value-added activities that occur downstream from farms in food preparation, but there are significant components of the nonfarm agricultural sector not included in current BEA industry aggregations.

Table 6-1 BEA categories and codes for food and agriculture, with corresponding NAICS codes

BEA code	Title	6-digit NAICS codes
11	Agriculture, forestry, fishing, and hunting	
111CA	Farms	11
113FF	Forestry, fishing, and related activities	3
31G	Manufacturing	
333	Machinery	1
325	Chemical products	2
311FT	Food/beverage/tobacco products	28
44RT	Retail	
445	Food and beverage stores	1
7	Service	
722	Food services and drinking places	3

SOURCE: https://www.bea.gov/industry/io_annual.htm

Table 6-2 contains a subset of potential additional NAICS categories that could be part of a supplemental account for more comprehensive reporting on food and agricultural industries.

Table 6-2 Potential (proposed) additional NAICS categories for a supplemental account

NAICS	Description
237990	Farm drainage tile installation
423320	Lime (except agricultural) merchant wholesalers
423820	Farm machinery and equipment merchant wholesalers
424590	Raw farm products (except field beans, grains) merchant wholesalers
424910	Chemicals, agricultural, merchant wholesalers
424910	Farm supplies merchant wholesalers
424910	Lime, agricultural, merchant wholesalers
424910	Pesticides, agricultural, merchant wholesalers
444220	Farm supply stores
484220	Farm products hauling, local
484230	Farm products trucking, long-distance
493120	Farm product warehousing and storage, refrigerated
493130	Bonded warehousing, farm products (except refrigerated)
493130	Farm product warehousing and storage (except refrigerated)
493190	Warehousing (except farm products, general merchandise, refrigerated)
522292	Farm mortgage lending
522294	Federal agricultural mortgage corporation

[40]https://www.ers.usda.gov/data-products/food-dollar-series.aspx

532490	Farm equipment rental or leasing
532490	Farm tractor rental or leasing
541711	Biotechnology research and development laboratories or services in agriculture
561710	Pest control (except agricultural, forestry) services
811310	Farm machinery and equipment repair and maintenance services
811310	Tractor, farm or construction equipment repair and maintenance services

Summary

In summary, administrative data have the potential to improve the efficiency of survey programs and the accuracy of statistical estimates derived from them. Challenges in the use of administrative data arise for several reasons. First, data collected for programmatic purposes and data collected from surveys are pursued with different objectives, so they are not always optimal for linking for the purpose of improving research and evidence-based policy. Second, the decentralized nature of the U.S. statistical system creates legal and administrative barriers to efficient cross-agency collaboration. However, recent work at the Census Bureau by CARRA, coupled with developments such as the recommendations of the Commission on Evidence-Based Policymaking, have greatly increased the chance of overcoming these hurdles. The latter recommendations, especially, hold the promise of motivating legislation to push mechanisms forward for broader data sharing and linkage across the nation's statistical agencies.

References

Adjemian, M. K. 2012. "Quantifying the WASDE Announcement Effect." *American Journal of Agricultural Economics* 94(2012): 238–56.

Administrative Data Liaison Service (ADLS). 2015. "Administrative Data Introduction." Retrieved 2015-10-21. http://www.adls.ac.uk/adls-resources/guidance/introduction/

Ahearn, Mary. 2013. "Challenges in Collecting Data from Complex Farm Operations: Review of Perspectives from an International Conference." In *Complex Farms and Sustainability in Farm Level Data Collection*, edited by H. C. J. Vrolijk (LEI Proceedings 13-054, The Hague).

Ahearn, M., D. Banker, D. M. Clay, and D. Milkove. 2011. "Comparative Survey Imputation Methods for Farm Household Income." *American Journal of Agricultural Economics 93*(2): 613–18.

Allan and Lueck, 2002. The Nature of the Farm: Contracts, Risk, and Organization in Agriculture, Cambridge, MIT Press.

Anderson, Edwin. 2017. *Administrative Data Used by NASS: A Discussion of Administrative Data Used by NASS Headquarters*. Staff Report, National Agricultural Statistics Service (NASS). Washington, DC: US Department of Agriculture.

Anderson, Edwin, and Daniel Beckler. 2017. *Administrative Data Used by NASS*. Paper prepared for the Panel on Improving Data Collection and Reporting about Agriculture with Increasingly Complex Farm Structures on February 10, 2017 in Washington, DC.

Bako, Dramane. 2018. *Guidelines on Improving and Using Administrative Data in Agricultural Statistics,* Global Strategy to Improve Agriculture and Rural Statistics. https://www.researchgate.net/publication/323280869_Guidelines_on_improving_and_using_administrative_data_in_agricultural_statistics [accessed May 31 2018].

Beckett, M. K., M. N. Elliott, S. Gaillot, A. Haas, et al. 2016. "Establishing Limits for Supplemental Items on a Standardized National Survey." *Public Opinion Quarterly* 80(4): 964–76.

Beckler, D. G. 2013. "Administrative Data Used by the National Agricultural Statistics Service." In *Proceedings of the Sixth International Conference on Agricultural Statistics*, Rio de Janeiro, Brazil, October 23–25, ISI Committee on Agricultural Statistics.

Berent, M. K., J. A. Krosnick, and A. Lupia. 2016. "Measuring Voter Registration and Turnout in Surveys: Do Official Government Records Yield More Accurate Assessments?" *Public Opinion Quarterly* 80(3): 597–621.

Berg, Emily. 2015. *Improving the Methodology for using Administrative Data More Broadly Across the Agricultural Statistics System: Administrative Data and the Statistical Programmes of Developed Countries.* Working Paper no. 3, Global Strategy to Improve Agricultural and Rural Statistics (GSARS), FAO. Available: http://gsars.org/wp-content/uploads/2015/10/WP2-Improving-the-Methodology-for-Using-Administrative-Data-in-an-Agricultural-Statistics-System_14102015.pdf

Bonnen J. T. 1977. "Assessment of the Current Agricultural Data Base: An Information System Approach." In *A Survey of Agricultural Economics Literature,* edited by L. R. Martin and G. G. Judge. St Paul: University of Minnesota Press.

_____. 1989. "On The Role of Data and Measurement in Agricultural Economics Research." *Journal of Agricultural Economics Research* 4.

Bonnen, J. T., Richard J. Hildreth, George G. Judge, George Tolley, and Harry Trelogan. 1972. "Our Obsolete Data Systems: New Directions and Opportunities." *American Journal of Agricultural Economics*.

Bradburn, N. M. 1978. "Respondent Burden." In *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 35–40). Available online: https://ww2.amstat.org/sections/srms/Proceedings/papers/1978_007.pdf

Calus, M. 2009. "Factors Explaining Farm Succession and Transfer in Flanders." PhD thesis, Ghent University, Ghent, Belgium.

Commission on Evidence-Based Policymaking (CEBP). 2017. *The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking.* Washington, DC.

Cruze, N. B. 2015. "Integrating Survey Data with Auxiliary Sources of Information to Estimate Crop Yields." In *Proceedings of the Survey Research Methods Section* (Washington, DC: American Statistical Association).

Davis, J. H. and R. A. Goldberg. 1957. "A Concept of Agribusiness." *American Journal of Agricultural Economics* 39(4): 1042–45.

Duncan, G. 2011. "Commentary: Future U.S. National Statistics Use of Administrative Data." *Journal of Privacy and Confidentiality* 3(2). https://doi.org/https://doi.org/10.29012/jpc.v3i2.603

Dunn, Richard A., and Brent Hueth. 2017. "Food and Agricultural Industries: Opportunities for Improving Measurement and Reporting." *American Journal of Agricultural Economics* 99(2): 510–23.

Earp, Morgan S., Jaki S. McCarthy, Nick D. Schauer, and Phil S. Kott 2008a. *Assessing the Effect of Calibration on Nonresponse Bias in the 2005 ARMS Phase III Sample Using 2002 Census of Agriculture Data*. Washington, DC: United States Department of Agriculture, National Agricultural Statistics Service, Research and Development Division.

_____. 2008b. *Assessing the Effect of Calibration on Nonresponse Bias in the 2006 ARMS Phase III Sample Using Census 2002 Data*. Washington, DC: United States Department of Agriculture, National Agricultural Statistics Service, Research and Development Division.

Endres, A. B., and L. Schlessinger. 2015. "FDA Revises Definition of "Farm" *farmdoc daily* (5):185. Published by the Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. http://farmdocdaily.illinois.edu/2015/10/fda-revises-definition-of-farm.html.

Eurostat. 2017. "Agricultural Holding." In *Glossary* (online). http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Agricultural_holding

Fay, R. E., and R. A. Herriot. 1979. "Estimation of Income from Small Places: An Application of James-Stein Procedures to Census Data." *Journal of the American Statistical Association* 74: 269–77.

Featherstone, Allen M., Mark A. Wood, Kevin L. Herbel, and Michael R. Langemeier. 2012. "Multiple Entity Farms: A Growing and Challenging Phenomenon." *Agricultural Finance Review* 72(2): 210–21. https://doi.org/10.1108/00021461211250447.

Federal Statistical Office (FSO). 2008. *NOGA 2008 General Classification of Economic Activities*. Neuchâtel: Federal Statistical Office.

Food and Agriculture Organization of the United Nations (FAO). 2017 *FAO World Program of the Census of Agriculture 2020,* Volume 1. Rome.

Galesic, M., and M. Bosnjak. 2009. "Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey." *Public Opinion Quarterly* 73(2): 349–60.

Galt, Ryan E. 2013. "The Moral Economy Is a Double-edged Sword: Explaining Farmers' Earnings and Self-exploitation in Community-Supported Agriculture." *Economic Geography* 89(4): 341–65.

Gardner, B. L. 1983. "Fact and Fiction in the Public Data Budget Crunch." *American Journal of Agricultural Economics* 65(5): 882–88. https://doi.org/10.2307/1240384

_____. 2002. *American Agriculture in the Twentieth Century: How It Flourished and What It Cost.* Cambridge, Mass.: Harvard University Press.

Gasson, R., and A. Errington. 1993. *The Farm Family Business.* Wallingford, UK: CAB International.

Gerling, M., H. Tran, and M. Earp. 2008. *Categorizing Nonresponse Occurring in the 2007 June Area Survey (California, Kansas, New York, Virginia and Washington*). Research and Development Division Report RDD-08-09. Washington, DC: United States Department of Agriculture, National Agricultural Statistics Service.

Global Strategy to Improve Agricultural and Rural Statistics. 2017. *Improving the Methodology for Using Administrative Data in an Agricultural Statistics System: Final Report*. Technical Report no. 24. Rome: Global Strategy.

Goodhue, Rachael E., Dale M. Heien, Hyunok Lee, and Daniel A. Sumner. 2003. "Contracts and Quality in the California Winegrape Industry." *Review of Industrial Organization* 23: 267- 282.

Groves, R. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70(5): 646–75.

Hancock, E., and K. Ott. 2017. "Total Respondent Burden for NASS for 2017." Presentation to the Panel on Improving Data Collection and Reporting about Agriculture with Increasingly Complex Farm Structures on February 10, 2017 in Washington, DC.

Hansen, K. M. 2007. "The Effects of Incentives, Interview Length, and Interviewer Characteristics on Response Rates in a CATI-Study." *International Journal of Public Opinion Research* 19(1): 112.

Herzog, A. R., and J. G. Bachman. 1981. "Effects of Questionnaire Length on Response Quality." *Public Opinion Quarterly* 45(4): 549–59.

Hoppe, Robert A. 2014. *Structure and Finances of U.S. Farms: Family Farm Report*. Washington, DC: United States Department of Agriculture, Economic Research Service.

Hoppe, Robert A., and James M. MacDonald. 2013. *Updating the ERS Farm Typology*. Washington, DC: United States Department of Agriculture, Economic Research Service.

Horowitz, Karen J., and Mark A. Planting. 2006. *Concepts and Methods of the U.S. Input-Output Accounts*. BEA Paper 0066. Washington, DC: Bureau of Economic Analysis.

Hotz, V., B. Jospeh, J. D. Balzekas, N. Bradburn, et al. 1998. *Administrative Data for Policy-Relevant Research: Assessment of Current Utility and Recommendations for Development*. Report of the Advisory Panel on Research Uses of Administrative Data of the Northwestern University/University of Chicago Joint Center for Poverty Research. http://public.econ.duke.edu/~vjh3/working_papers/adm_data.pdf

International Financial Reporting Standards Foundation (IFRSF). 2017. "Biological Assets Growing on Bearer Plants (IAS 41)." http://www.ifrs.org/projects/2017/ias-41-biological-assets-growing-on-bearer-plants/

Isengildina-Massa, Olga, Berna Karali, and Xiang Cao. N.D. "The Market Impact of USDA Crop and Livestock Reports in the Big Data Era." Unpublished manuscript (Submitted to the *Journal of Agricultural Economics*).

Judge, George G., Richard H. Day, S. R. Johnson, Gordon C. Rausser, and Lee R. Martin, pp. 386-407. Minneapolis: University of Minnesota Press, 1977.

Just, R. E. 1983. "The Impact of Less Data on the Agricultural Economy and Society." *American Journal of Agricultural Economics* 65(5) (Proceedings Issue): 872–81.

Karali, B. "Do USDA Announcements Affect Co-movements Across Commodity Futures Returns?" Journal of Agricultural and Resource Economics 37(2012):77-97.

Katchova, Ani, and Mary Ahearn. 2015. "Strategies to Link Farm Business and Household Data Over Time: Insights from Research on Farm Transitions." Linking Federal Data on Agriculture Workshop, Economic Research Service, U.S. Department of Agriculture, Washington, DC, October 2, 2015. http://www.copafs.org/UserFiles/file/FarmTransitionsDataLinkagesKatchova2.pdf.

Kreiner, Claus Thustrup, David Dreyer Lassen, and Søren Leth-Petersen. 2013. *Measuring the Accuracy of Survey Responses using Administrative Register Data: Evidence from Denmark*. Working Paper No. 19539. Cambridge, Mass.: National Bureau for Economic Research.

Lavallée, P. 2007. *Indirect Sampling.* New York. Springer.

Lindley, D. V., and A. F. M. Smith. 1972. "Bayes Estimates for the Linear Model." *Journal of the Royal Statistical Society* 34:1–41.

MacDonald, James, Robert Hoppe, and Doris Newton. 2018. Three Decades of Consolidation in US Agriculture. United States Department of Agriculture, Economic Research Service, Economic

Information Bulletin, Number 189, March 2018: Available: https://www.ers.usda.gov/webdocs/publications/88057/eib-189.pdf?v=0

MacDonald, James M., Penni Korb, and Robert A. Hoppe. 2013. *Farm Size and the Organization of U.S. Crop Farming.* ERR-152. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Mattos, F. L., and R. L. F. Silveira. 2016. "Futures Price Response to Crop Reports in Grain Markets." *Journal of Futures Markets* 36(10): 923–42.

McCarthy, J. S., Daniel G. Beckler, and Suzette M. Qualey. 2006. "An Analysis of the Relationship Between Survey Burden and Nonresponse: If We Bother Them More, Are They Less Cooperative?" *Journal of Official Statistics* 22(1): 97–112.

McCarthy, Jaki, James Wagner, and Herschel Lisette Sanders. 2017. "The Impact of Targeted Data Collection on Nonresponse Bias in an Establishment Survey: A Simulation Study of Adaptive Survey Design." *Journal of Official Statistics* 33(3): 857–71.

Miller, D., and T. O'Connor. 2012. *Item Response Rates for the Agricultural Resource Management Survey III in 2006 and 2007*. National Agricultural Statistics Service Research Report, RDD-11-07. Washington, DC: United States Department of Agriculture, National Agricultural Statistics Service.

Moss, Charles B., Allen M. Featherstone, and Christine A. Wilson. 2012. "Review of the financial data provided by the agricultural resource management survey." *Agricultural Finance Review* 72(2). https://doi.org/10.1108/afr.2012.42172baa.001

Myers, R. J., R. J. Sexton, and W. G. Tomek, 2010. "A Century of Research on Agricultural Markets." *American Journal of Agricultural Economics* 92: 376–402.

National Agricultural Statistics Service (NASS). 2014. *2012 Census of Agriculture, Appendix B: General Explanation and Census of Agriculture Report Form.* Washington, DC.

National Agricultural Statistical Services/National Institute of Statistical Sciences Technical Expert Panel (NASS/NISS). 2017a. *Publication of Agriculture Census Data on Farm Operator Demographics.* Washington, DC.

National Academies of Sciences, Engineering, and Medicine (NASEM). 2017a. *Federal Statistics, Multiple Data Sources, and Privacy Protection: Next Steps*. Washington, DC: National Academies Press. doi: https://doi.org/10.17226/24893

National Academies of Sciences, Engineering, and Medicine (NASEM). 2017b. *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy.* Washington, DC: National Academies Press. https://doi.org/10.17226/24652.

National Academies of Sciences, Engineering, and Medicine (NASEM). 2017c. *Improving Crop Estimates by Integrating Multiple Data Sources.* Washington, DC: National Academies Press. https://doi.org/10.17226/24892.

National Academies of Sciences, Engineering, and Medicine (NASEM). 2018. *Reengineering the Census Bureau's Annual Economic Surveys*. Washington, DC: National Academies Press. doi: https://doi.org/10.17226/25098.

National Research Council. 2007. *Understanding Business Dynamics: An Integrated Data System for America's Future* (J. Haltiwanger, L. M. Lynch, and C. Mackie, editors). Panel on Measuring Business Formation, Dynamics, and Performance, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

_____. 2008. *Understanding American Agriculture: Challenges for the Agricultural Resource Management Survey*. Panel to Review USDA's Agricultural Resource Management Survey, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

_____. 2013. *Nonresponse in Social Science Surveys: A Research Agenda* (R. Tourangeau and T. J. Plewes, editors). Panel on a Research Agenda for the Future of Social Science Data Collection, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

O'Connor, T. 1992. *Identifying and Classifying Reasons for Nonresponse on the 1991 Farm Costs and Returns Survey*. SRB Research Report SRB-92-10. Washington, DC: United States Department of Agriculture, National Agricultural Statistics Service.

O'Donoghue, Erik J., Robert A. Hoppe, David E. Banker, and Penni Korbe. 2009. *Exploring Alternative Farm Definitions: Implications for Agricultural Statistics and Program Eligibility*. Economic Information Bulletin no. 49. Washington, DC: United States Department of Agriculture.

O'Donoghue, E.J., R.A. Hoppe, D.E. Banker, R. Ebel, K. Fuglie, P. Korb, M. Livingston, C. Nickerson, and C. Sandretto. 2011. "The Changing Organization of U.S. Farming," EIB-88. USDA Economic Research Service.

O'Hara, Amy. 2016. *The Census Bureau Linkage Infrastructure*. Presentation at the Joint Statistical Meetings on August 1, 2016 in Chicago, IL. Available: http://ww2.amstat.org/misc/OHaraJSM2016.pdf

Oremus, Will. 2017. "How 'Big Data' Went Bust--And what comes next." *Slate* (online), October 16, 2017.

Parsons J. L. 2011. *USDA/NASS Current and Imminent Data Collection Challenges Faced by Statistical Agencies*. Paper presented at the Enhancing Data Collection for Complex Agricultural Establishments Workshop, Niagara on the Lake, Ontario, Canada, June 26-28, 2011.

Peytchev, A., and E. Peytcheva. 2017. "Reduction of Measurement Error Due to Survey Length: Evaluation of the Split Questionnaire Design Approach." *Survey Research Methods* 11(4): 361–68.

Porter, S. R., and M. E. Whitcomb. 2005. "Non-response in Student Surveys: The Role of Demographics, Engagement and Personality." *Research in Higher Education* 46(2): 127–52.

Prell, M., H. Bradsher-Fredrick, C. Comisarow, S. Cornman, et al. 2009. *Profiles in Success of Statistical Uses of Administrative Data*. Available at: http://www.bls.gov/osmr/fcsm.pdf [November 2017].

Prewitt K. 2010. "Science Starts Not After Measurement, But With Measurement." *ANNALS of the American Academy of Political and Social Science* 631(1): 7–16.

Pruitt, J. R., G. T. Tonsor, K. R. Brooks, and R. J. Johnson. 2014. "End User Preferences for USDA Market Information." *Food Policy* 47: 24–33.

Puts, M. J. H., M. Tennekes, P. J. H. Daas, and C. D. Blois. 2016. "Using Huge Amounts of Road Sensor Data for Official Statistics." Proceedings of the European Conference on Quality in Official Statistics (Q2016). Madrid. Available at: http://www.pietdaas.nl/beta/pubs/pubs/q2016Final00177.pdf [November 2017].

Reinhardt, N., and P. Bartlett. 1989. "The Persistence of Family Farms in United States Agriculture." *Sociologia Ruralis* 29: 203–25.

Ridolfo, H., V. Harris, J. McCarthy, D. Miller, N. Sedransk, and L. Young. 2016. "Developing and Testing New Survey Questions: The Example of New Questions on the Role of Women and New/Beginning Farm Operators." In Proceedings of the Fifth International

Conference of Establishment Surveys, June 20-23, 2016, Geneva, Switzerland: American Statistical Association

Robbins, M. W., and T. K. White, T. K. 2011. "Farm Commodity Payments and Imputation in the Agricultural Resource Management Survey." *American Journal of Agricultural Economics* 93(2): 606–12.

Sanginabadi, Bahram. 2018. *USDA Forecasts: A Meta-analysis Study. A*vailable online at: https://arxiv.org/abs/1801.06575

Smith, J., M. Beaulieu, E. Kumar, and L. O'Neill. 2013. "Recent Developments in the Production of Agriculture Statistics." *Proceedings of the 6th International Conference on Agricultural Statistics*. Rio de Janeiro, Brazil.

Sumner, Daniel A. 1982. "The Off-Farm Labor Supply of Farmers." *American Journal of Agricultural Economics* 64(3): 499–509.

_____. 1985. "Farm Programs and Structural Issues." In *U.S. Agricultural Policy: The 1985 Farm Legislation,* edited by Bruce Gardner. Washington, D.C.: American Enterprise Institute.

_____. 1990. "Targeting and the Distribution of Program Benefits." In *Agricultural Policies in a New Decade,* edited by Kristin Allen. Washington D.C.: Resources for the Future.

_____. 1991. "Targeting Farm Programs." *Contemporary Policy Issues* IX: 93–106.

_____. 2007. "Farm Subsidy Tradition and Modern Agricultural Realities." Paper prepared for the American Enterprise Institute project on Agricultural Policy for the 2007 Farm Bill and Beyond. Available online at: http://aic.ucdavis.edu/research/farmbill07/aeibriefs/20070515_sumner Rationalesfinal.pdf

_____. 2008. "Agricultural Subsidy Programs." In *The Concise Encyclopedia of Economics*, 4 Library of Economics and Liberty. http://www.econlib.org/library/Enc/AgriculturalSubsidyPrograms.html, accessed 11/18/13.

_____. 2014. "American Farms Keep Growing: Size, Productivity, and Policy. *The Journal of Economic Perspectives* 28(1): 147–66.

_____. 2017. "Economics of US State and Local Regulation of Farm Practices, with Emphasis on Restrictions of Interstate Trade." 2017. *Annual Review of Resource Economics* 9.

Sumner, Daniel A., and James D. Leiby. 1987. "An Econometric Analysis of the Effects of Human Capital on Size and Growth Among Dairy Farms." *American Journal of Agricultural Economics* 69(2): 465–70.

Sumner, Daniel A., and Christopher A. Wolf. 2002. "Diversification, Vertical Integration and the Regional Pattern of Dairy Farm Size." *Review of Agricultural Economics* 24(2): 442–57.

Thompson, K., and S. Kaputa. 2017. "Investigating Adaptive Nonresponse Follow-up Strategies for Small Businesses through Embedded Experiments." *Journal of Official Statistics* 33(3): 835–56. doi:10.1515/jos-2017-0038.

United Nations Economic Commission for Europe (UNECE). 2011. *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*. Available at: http://www1.unece.org/stat/platform/display/adso/Using+Administrative+and+Secondary+Sources+for+Official+Statistics [November 2017].

_____. 2015. *Guidelines on Statistical Business Registers*. Available from: http://www.unece.org/index.php?id=40574

United States Bureau of the Census. 1996. 1992 census of agriculture: Volume 2 Subject series. Part 4 History. Washington, DC. Available: http://usda.mannlib.cornell.edu/usda/AgCensusImages/1992/02/History.pdf.

United Nations Economic and Social Council. 2016. *Report of the Global Working Group on Big Data for Official Statistics.* Available: http://unstats.un.org/unsd/statcom/47th-session/documents/2016-6-Big-data-for-official-statistics-E.pdf [November 2016].

United States Department of Agriculture (USDA). 2014. *2012 Census Highlights: Farm Economics.* ACH12-2, May 2014. Washington, DC.

United States Department of Agriculture, Economic Research Service (USDA, ERS). 2017. *America's Diverse Family Farms: 2017 Edition*. Washington, DC.

_____. 2018. *Agricultural Productivity in the U.S*. Washington, DC. https://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us.aspx.

United States Department of Agriculture, National Agricultural Statistical Service (USDA, NASS), Advisory Committee on Agriculture Statistics. *Summary and Recommendations.* Annual Meeting, November 8-9, 2016. Available: https://www.nass.usda.gov/About_NASS/Advisory_Committee_on_Agriculture_Statistics/2017/ACAS%20November%202016%20Meeting%20Executive%20Summary.pdf

_____. 2015. "ARMS Progress Report: USDA's National Agricultural Statistics Service and Economic Research Service Respond to Recommendations by the Agricultural Resource Management Survey Review Panel." https://www.nass.usda.gov/Surveys/ARMS_Progress_Report.pdf

United States Office of Management and Budget (OMB). 2006. *Standards and Guidelines for Statistical Surveys*. Washington, D.C. Retrieved September 11, 2017 from www.whitehouse.gov/omb/inforeg/statpolicy/standards.pdf

Vrolijk, H. C. J., and K. J. Poppe. 2016. "Structural Change in Dutch Agriculture; Impact on Farm Level Statistics." In *7th International Conference on Agricultural Statistics* (Rome, October 10–28, 2016). https://www.researchgate.net/publication/309616377_Structural_change_in_Dutch_agriculture_impact_on_farm_level_statistics [accessed Aug 17 2018].

Weber, J. G., and D. M. Clay. 2013. "Who Does Not Respond to the Agricultural Resource Management Survey and Does It Matter?" *American Journal of Agricultural Economics* 95(3): 755–71.

Weber, J. G., N. Key, and E. O'Donoghue. 2016. "Does Federal Crop Insurance Make Environmental Externalities from Agriculture Worse?" *Journal of the Association of Environmental and Resource Economists* 3(3): 707–42.

Wolf, Christopher, and Daniel A. Sumner. 2001. "Are Farm Size Distributions Bimodal? Evidence from Kernel Density Estimates of Dairy Farm Size Distributions." *American Journal of Agricultural Economics,* 83(1): 77–88.

Young, Linda J., Andrea C. Lamas, and Denise A. Abreu. 2017. "The 2012 Census of Agriculture: A Capture–Recapture Analysis." *Journal of Agricultural, Biological and Environmental Statistics* 22 (4): 523–39.

Appendix A
Biographical Sketches of Panel Members

Catherine L. Kling *(Chair)* is a professor in the Dyson School of Applied Economics and Management, as well as faculty director at the Atkinson Center for a Sustainable Future, at Cornell University. She was formerly the Charles F. Curtiss distinguished professor of agriculture and life sciences, a professor of economics, and director of the Center for Agricultural and Rural Development, all at Iowa State University. Kling is undertaking research to examine how agricultural practices affect water quality, wildlife, soil carbon content, and greenhouse gases. She is an elected member of the National Academy of Sciences and an elected fellow of both the Association of Environmental and Resource Economists and the American Agricultural Economics Association. She currently serves as chair of the National Academies of Science, Engineering, and Medicine's (NASEM's) Water Science and Technology Board and has been a member of several previous NASEM committees. Kling holds a B.S. in business and economics from the University of Iowa and a Ph.D. in economics from the University of Maryland.

J. Gordon Arbuckle, Jr., is an associate professor in the Department of Sociology at Iowa State University. His research and extension activities focus on improving the social and environmental performance of agriculture. His primary area of interest is drivers of farmer and agricultural stakeholder decision making and action realted to soil and water quality. He is codirector of the Iowa Farm and Rural Life Poll, a survey research program that collects and disseminates information on issues of importance to agricultural stakeholders across Iowa and the Midwest. Arbuckle holds an M.S. in agricultural economics and a Ph.D. in rural sociology, both from the University of Missouri.

Norman M. Bradburn is the Tiffany and Margaret Blake distinguished service professor emeritus at the University of Chicago, where he also served on the faculties of the Department of Psychology, the Irving B. Harris Graduate School of Public Policy Studies, the Booth School of Business, and the College. He is also a senior fellow at the National Opinion Research Center (NORC) at the University of Chicago. Bradburn previously served as assistant director for social,

behavioral, and economic sciences at the National Science Foundation. His research focuses on psychological well-being and the assessment of quality of life using large-scale sample surveys. He has served on more than two dozen NASEM activities, including a term as chair of the Committee on National Statistics. He is a fellow of the American Statistical Association, a fellow of the American Association for the Advancement of Science, and an elected member of the International Statistical Institute. He holds an M.A. in clinical psychology and a Ph.D. in social psychology, both from Harvard University.

Richard A. Dunn is an associate professor in the Department of Agricultural and Resource Economics at the University of Connecticut. He is also an affiliate of the Connecticut Center for Health Improvement and Policy and the Charles J. Zwick Center for Food and Resource Policy. Dunn's research focuses on the use of federal administrative data to better measure the contribution of food and agriculture industries to the U.S. economy, with emphasis on the effect of different reporting requirements across federal administrative and survey programs. Dunn was previously an assistant professor in the Department of Agricultural Economics at Texas A&M University. He holds a B.A. in economics and mathematics from Williams College, an M.Sc. in econometrics and mathematical economics from the London School of Economics and Political Science, and a Ph.D. in economics from the University of Wisconsin–Madison.

Allen M. Featherstone is the department head and a professor in the Department of Agricultural Economics, as well as director of the masters in agribusiness program, at Kansas State University. As an agriculture finance scholar, he has studied land markets, investment decisions, mergers in the financial services industry, the probability of agricultural loan default and loan loss severity, the influence of taxes on farmland, and alternative federal tax systems. Featherstone worked to create the Comparative Food and Agriculture Systems course to give students first-hand knowledge of agriculture and cultural situations in other parts of the world. He is currently an executive director of the Agricultural and Applied Economics Association. Featherstone holds an M.S. and a Ph.D. in agricultural economics, both from Purdue University.

Joseph W. Glauber is a senior research fellow at the International Food Policy Research Institute (IFPRI) in Washington, DC, where his areas of interest are price volatility, global grain reserves, crop insurance, and trade. Prior to joining IFPRI, Glauber spent more than 30 years at the U.S. Department of Agriculture (USDA), including as chief economist from 2008 to 2014. As chief economist, he was responsible for USDA's agricultural forecasts and projections, oversaw climate, energy, and regulatory issues, and served as chairman of the board of directors of the Federal Crop Insurance Corporation. He is an elected fellow of the Agricultural and Applied Economics Association and holds a Ph.D. in agricultural economics from the University of Wisconsin.

Brent Hueth is an associate professor in the Department of Agricultural and Applied Economics and director of the University of Wisconsin Center for Cooperatives at the University of Wisconsin–Madison. His research and teaching focus on agricultural firms and markets, cooperative enterprise, and economic development. Hueth has published in top economics journals, including the *American Journal of Agricultural Economics* and the *Journal of Regulatory Economics*, and is a research fellow at the Institute for Exceptional Growth Companies. He also serves as executive director of the Census Bureau's Research Data Center at

the University of Wisconsin–Madison. Hueth holds a Ph.D. in agricultural economics from the University of Maryland.

Ani L. Katchova is farm income enhancement chair and associate professor in the Department of Agricultural, Environmental, and Development Economics at The Ohio State University. In her role as chair of the Farm Income Enhancement Program, she manages a research team of post-doctoral students and graduate research assistants to conduct research and outreach on U.S. agricultural economics issues. Katchova's research areas include agricultural finance, cooperatives, agribusiness management and marketing, and applied econometrics. Her research has been published in the *American Journal of Agricultural Economics, Agricultural Finance Review,* and *Agribusiness*. She currently serves as chair of the review panel of the USDA-Economic Research Service's (USDA-ERS's) Farm Income and Wealth Forecast Program, and is an executive board director of the Agricultural and Applied Economics Association. She holds a Ph.D. from The Ohio State University.

Doris Mold is the President of Sunrise Agricultural Associates, LLC, an agricultural consulting firm. She is an Agricultural Consultant, Agricultural Economist, Educator, as well as a farm co-owner/operator. She also teaches Farm and Agri-Business Management at the University of Minnesota for MAST International. In 2015 she served on an expert national Panel on Statistics on Women and Beginning Farmers in the USDA Census of Agriculture. Mold served six years on the Agricultural Statistics Advisory Committee for NASS and chaired the committee for three years. Mold is Past President of American Agri-Women, the nation's largest coalition of women in agriculture. Mold maintains a unique position as a producer who uses NASS data and provides data to NASS (via her active dairy farming operation), and as an economist who utilizes the data in research, teaching, business, and in volunteer advocacy. She holds an M.S. in Agricultural and Applied Economics from the University of Minnesota.

Jean Opsomer is a Vice President at Westat in Rockville, Maryland. He was formerly a professor and department chair in the Department of Statistics at Colorado State University, as well as a faculty member at Iowa State University. His research focuses on shape-constrained and nonparametric methods in survey estimation and on several interdisciplinary projects with survey components on a range of topics. He is a member of Statistics Canada's Advisory Committee on Statistical Methods. He previously served on the Bureau of Labor Statistics' Technical Advisory Committee, the USDA's Advisory Committee on Agricultural Statistics and NASEM's Panel to Review USDA's Agricultural Resource Management Survey. Opsomer is an elected fellow of the Institute of Mathematical Statistics and the American Statistical Association, as well as an elected member of the International Statistical Institute. He holds a Ph.D. in statistics from Cornell University.

Greg Peterson is director general for Agriculture, Energy, Environment and Transportation Statistics in the Economics Statistics Field at Statistics Canada. Since joining Statistics Canada in 1990, Peterson has worked in many areas, covering manufacturing, culture and tourism, retail trade, and recently was director of the Investment, Science and Technology group. In that capacity, he directed the statistical program that measures science, technology and innovation, the digital economy, and capital spending, as well as building permits and property values. Peterson also leads Canada's census of agriculture, a program that sends out and compiles data

from 250,000 questionnaires every five years. He holds a B.A. from Concordia University and an M.A. from Queen's University, both in economics.

Krijn J. Poppe is a senior economist and chief policy advisor at Wageningen University and Research. Poppe also manages various research programs for the European Union covering the food industry, including several studies for the Farm Accountancy Data Network. For several years, he co-led the European Union's Standing Committee on Agricultural Research's Strategic Working Group on Agricultural Knowledge and Innovation Systems. The Dutch government appointed him as a member of the Netherland's Council for the Environment and Infrastructure. He previously served as secretary-general of the European Association of Agricultural Economists, of which he is a fellow. Currently he is involved in managing its journals the European Review of Agricultural Economics and EuroChoices. Poppe has served as chief science officer of the Dutch Ministry of Agriculture. He is a board member of SKAL, the Dutch Inspection Organization for Organic Farming. He holds an M.Sc. in business economics from Erasmus University Rotterdam.

Daniel A. Sumner is the Frank H. Buck, Jr., professor in the Department of Agricultural and Resource Economics at the University of California, Davis and the director of the University of California's Agricultural Issues Center. He participates in research and teaching and directs an outreach program on public issues related to agriculture. Sumner has served on the President's Council of Economic Advisers, is a former assistant secretary for economics at USDA, and is a former chair of the International Agricultural Trade Research Consortium. He is also a consultant for farm organizations, government agencies, and firms and is a frequent speaker at national and international conferences and symposia. Sumner is a fellow of the American Agricultural Economics Association and holds a Ph.D. in economics from the University of Chicago.

James Wagner is a research associate professor at the University of Michigan Survey Research Center as well as associate director of the Michigan Program in Survey Methodology. He also teaches in the Joint Program in Survey Methodology and serves as principal investigator on several large studies; he is currently the chief mathematical statistician for the National Survey of Family Growth. Wagner was recently an invited summer scholar at the U.S. Census Bureau. His research interests include nonresponse error, quality indicators for survey data, and responsive or adaptive design. Wagner serves as associate editor of *Survey Research Methods* and *Journal of Official Statistics.* He holds an M.S. in political science and a Ph.D. from the program in survey methodology, both from the University of Michigan.

Jeremy G. Weber is an associate professor in the Graduate School of Public and International Affairs and the Department of Economics at the University of Pittsburgh. He previously worked at the World Bank and at USDA's Economic Research Service—where he produced numerous reports on farm income characteristics—and was an adjunct faculty member for the master's program in applied economics at Johns Hopkins University. Weber's current research focus is on energy, natural resources, and agricultural economics; and he has published more than a dozen articles in journals such as *Energy Economics*, *Resource and Energy Economics*, *World Development*, *Land Economics*, and the *American Journal of Agricultural Economics*. He holds a Ph.D. in agricultural and applied economics from the University of Wisconsin–Madison.