AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.

# A SURVEY OF AGRICULTURAL ECONOMICS LITERATURE

VOLUME 2

## *Quantitative Methods in Agricultural Economics, 1940s to 1970s*

*George G. Judge, Richard H. Day, S. R. Johnson, Gordon C. Rausser, and Lee R. Martin, editors*

Published by the University of Minnesota Press, Minneapolis,
for the American Agricultural Economics Association

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.

# Assessment of the Current Agricultural Data Base: An Information System Approach

**James T. Bonnen**
*Professor of Agricultural Economics*
*Michigan State University*

*It is a capital mistake to theorize before one has data.*

SIR ARTHUR CONAN DOYLE

*The discovery of facts . . . depends at least in part on concepts, assumptions, and inferences which can only be defended with reference to normative presumptions.*

MARC J. ROBERTS

*If there is no "given" in experience, then there is no difference between deduction and induction.*

C. W. CHURCHMAN

In 1969 the American Agricultural Economics Association established the Committee on Economic Statistics to evaluate questions that were being raised about the quality and reliability of certain types of agricultural data. In cooperation with the Statistical Reporting Service (SRS) and the Economic Research Service (ERS) of the Department of Agriculture different elements of the agricultural data base were examined (AAEA [1972], Hildreth [1975], Bonnen [1975], Brandow [1976]). The committee worked with many government, university, business, professional, and foundation groups in contributing to a widening sense of the current deficiencies and future needs of the agricultural data base. Instability and uncertainty in the world food situation since 1973 have made the adequacy of information on food and agriculture an urgent worldwide concern.

*Note:* This paper has been adapted from two articles by the author, who wishes to express his gratitude to the *American Journal of Agricultural Economics* (Bonnen [1974]) and the Agricultural Development Council (Bonnen [1976]) for permission to reprint substantial portions of those articles. The critical assistance of a number of reviewers is acknowledged in the original articles. The brief discussion of the Census of Agriculture, which has not been published before, was reviewed by the Bureau of Census and by Karl Wright and Eldon Weeks. The author is responsible for any errors or misinterpretations that may remain. The research for this paper was financed by Michigan Agricultural Experiment Station Project 991.

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



## *The Current State of Our Information Systems*

The AAEA Committee on Economic Statistics concluded that in those instances in which early agricultural data series were not performing as well as they had in earlier years, the problem most frequently was a growing obsolescence in the concepts which the data system attempted to measure (AAEA [1972] ). Some of these concepts, such as the idea of a farm, are so old and so much a part of our historical tradition that we hardly think of them as concepts at all. But 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 section. The concept guided the early development of agricultural data systems especially during their greatest period of growth in the 1920s and 1930s. The structure of the food and fiber industry today only vaguely resembles the structure that prevailed at that time. The world has changed and the concept has not.[1]

### Conceptual Obsolescence in Agricultural Data

The data systems which constitute the agricultural data base of the United States are among the oldest in the federal statistical establishment. The first Census of Agriculture was taken in 1840. Only the censuses of population and manufactures are older. The Department of Agriculture was established in 1862, and by the last quarter of the nineteenth century it was collecting data for many purposes. The United States agricultural data system was given its modern form during the period of the Great Depression, which in agriculture extended from the 1920s until World War II. It was during this period that the system extended its capability beyond that of simply counting things such as agricultural output and acreage by crop, farm numbers, and farm population. Out of this period came the basic concepts underlying the modern usage in farm income and prices and the social and economic accounting capability of the present agricultural information system. The chief focus at that time was the income or welfare of farmers, and a one-to-one relationship was generally assumed to exist between a farm and a farm family. The modern institutional and organizational form of the federal agencies that operate these agricultural information systems also was created during the 1920s and 1930s. There have been modifications, but the general purposes and administrative structure of the system remain fundamentally the same.

Much of our agricultural data base is far more accurate today than the same type of data in the past. Most of these improved data are based on concepts that are biological or physical and have not changed or have changed lit-

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



tle. Examples would be the number of cattle and pigs and the acreage and yields of potatoes or cotton produced. The great improvement in accounting, measurement, and data-processing capability over the last thirty years has combined with conceptual stability to increase the quality of some data. Thus, despite the criticism they receive, modern crop and livestock production estimates, with their biological and physical concept base, tend to be far better statistics than they were fifty or even ten years ago.

Certain statistics based on social science concepts have also retained most of their reliability and in some cases have actually been improved. This tends to be the case for food and fiber statistics in areas where technological and organizational changes have not been rapid. For example, measures of farm production of wheat and most cereals appear to have lost little in conceptual reliability while gaining much in reliability of measurement. Grain prices are another matter. At the other end of the spectrum, where change in the food and fiber sector has been most extreme, statistics for broiler production on farms are weak and broiler prices at the farm level have become nearly impossible to collect or interpret. In poultry and eggs and in many fruit and vegetable products, contracting and vertical integration of both inputs and outputs have undermined, if not destroyed, the traditional concept of the farm which underlies production and marketing statistics. The discovery of beef prices has also grown more difficult and the data ambiguous. Price-spread data present even greater difficulties (Brandow [1976]). Data on other livestock and on cotton, tobacco, peanuts, and other commodities fall between these two extremes.

Conceptual obsolescence in data is of two types. It can occur not only because of changes in the organization and nature of the food and fiber industry, as just described, but also because the agenda of food and fiber policy (public and private) shifts drastically, as it has recently, changing the questions which the information system is expected to answer. When the questions change, it is almost always found that the conceptual base of some data, especially secondary data, is not a fully appropriate representation and also that some data critical to the new questions are not even being collected. When normative or positive change occurs either in the object being represented by data or in the environment of the object, some degree of conceptual obsolescence is almost certain to follow.

Recent major examples of conceptual obsolescence of data arising from changes in the environment of agriculture can be seen in the entirely new questions which agricultural economists are asked to answer today, as a consequence of new values held and new positive knowledge about the environment, the energy economy, and the world food situation. The overall agenda of urgent agricultural policy issues has changed a great deal since the Great

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



Depression when the better part of our present data system was designed and built. Some older data have been conceptually redesigned to respond to new questions, but by and large we have "made do," fiddling with different definitions of the same concept. Thus, for example, the farm has been periodically redefined in recent agricultural censuses, but the concept itself has slowly become obsolete in so many uses that no matter how sensible the new definition we still measure something that in some major degree no longer exists.

Farm income is an example of both types of conceptual obsolescence. While improvements have been made in the concept and in the accounting rules which make operational the farm income concept, we still fail to net out certain expenses and assets and we miss some income flows entirely (Weeks [1971a], Carlin and Smith [1973], Simunek [1976], Hildreth [1975]). Changing the design and implementing the farm income concept are difficult and are often distorted by the congressional political imperatives of the day. Farm income data are still inconsistent with the current conceptual design of national income accounting (AAEA [1972]).

Farm input and output measures have long exhibited many conceptual deficiencies, even though some improvements have been made periodically. As the American farm became industrialized, specialization spun off many production, processing, and marketing functions from the farm to agricultural business firms. As a consequence, agriculture long ago ceased to be just farms. We still lack an adequate paradigm with which to describe and categorize a modern food and fiber industry and its subsectors and to provide a general conceptual basis for sector statistics. We do not, for example, have an integrated and consistent set of descriptors of the size and productivity of the food sector or its social performance.

## The Census of Agriculture

The Census of Agriculture exhibits problems which are a function of its unique role in the United States agricultural data base. The Census has long served as the five-year baseline to which less accurate data are periodically adjusted. The Census collects data not otherwise available: detailed descriptive data on the structure of agriculture by county, state, and national levels; certain farm characteristics by type of farming, economic class, and various other farm classifications, enterprise data on size, physical inputs, outputs, costs, and cash receipts for major crops and livestock, much of it by national, state, and county levels of aggregation; and data on the farm family. Perhaps the most valuable aspect of the Census is that it has been the only reliable, nationally uniform source of time series data at the county level.

The Census shares with other agricultural data systems the serious problems of conceptual obsolescence, especially in its basic unit of observation,

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



the farm, and in the lack of an adequate paradigm to describe the economic activities of the total food sector. There are many problems involved, but the "establishment" concept, which is the basic unit of observation in the economic censuses, is considered by some analysts to be the appropriate starting point in developing a new basic unit of observation for the collection of data from farmers. In any event, the 1974 Census provided, for the first time, some data by the Standard Industrial Classification (SIC) code using a modified establishment concept.

The Census remains essentially a census of farms, although farming, despite its absolute economic growth, has declined to less than 15 percent of the economic activity of the total food sector. Both the Census of Agriculture and the USDA collect data on some activities of nonfarm agricultural firms, but the data, while quite valuable, are not comprehensive and there is no food sector paradigm that allows the data to be meaningfully aggregated into useful sector statistics. The current plan to move the agricultural census into the same time frame as other economic censuses will allow little integration of sector data until a common conceptual base, including an adequate paradigm for the food sector built around a common basic unit of observation, has been developed.

Historically, any census was a complete enumeration. In the agricultural censuses after World War II basic data continued to be obtained through complete enumeration, but sampling was introduced to reduce the cost of obtaining certain additional kinds of data. In the 1969 Census all data for farms with $2,500 or more in sales were collected on a 100 percent basis, but all data for farms with less than $2,500 in sales were collected on a 50 percent sample basis. For the 1974 Census all data were collected on a 100 percent basis. In addition, the 1969 Census and the 1974 Census represented a shift from an area frame to a list frame universe. The latter universe was developed from 1964 Census records augmented primarily by farm addresses maintained for administrative purposes by the Agricultural Stabilization and Conservation Service, the Social Security Administration, and the Internal Revenue Service. This consolidated list of addresses provides a less expensive way of specifying the universe of farms, but it contains many nonfarm names and is currently incomplete and thus introduces error. The USDA has also increased its use of list frames for sampling. At best a list frame is never really complete or up to date, this raises a question about whether a list frame alone can ever provide an adequate base for the collection of complex and detailed data.

After some experimentation in earlier years the Census was shifted almost entirely in 1969 and 1974 from the traditional method of having enumerators interview farmers and fill out the questionnaires to having mail questionnaires filled in by the farmers and sent back to a central location. This change

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



helped somewhat to control the continual rise in costs, but it substituted respondent variability for the previous enumerator bias and introduced problems of error as a result of incomplete list frame coverage and, especially in 1974, a higher nonresponse rate. Since 1969 the Census of Agriculture has been a mailed-in report from a list frame universe. This complete transformation in methodology slowed the escalation of costs of the Census but at the expense of a significant reduction in the timeliness and quality of the data. This has been especially true of its most valuable feature, county data, some of which now appear to be quite unreliable and no longer very useful in some of the most important traditional applications.

The decline in the quality of Census data has been compounded by longer and longer publication delays. Many data have been published so late as to be nearly useless except for historical purposes. This has contributed to the erosion of the Census as a baseline for "truing up" other independently collected data. Indeed, following the 1973 flap over the Secretary of Commerce's intention not to take the 1974 Census, bills were introduced in the Congress and hearings were held on a proposed transfer of the agriculture census functions to the Department of Agriculture. These bills did not emerge from committee, but such notions will persist if the quality of agricultural census data does not improve and if the commitment of the Commerce Department to the Census of Agriculture continues to wane.

Many of the problems of the Census of Agriculture arise out of inadequate resources and a long-term lack of administrative support in facing the extremely complex problems of a census in a rapidly changing economic sector of more than two and a half million farmer entrepreneurs. The top political leadership of the Department of Commerce has too often viewed the Census of Agriculture as a service activity which is marginal to the mission of the department and thus a natural candidate for budget cuts and administrative neglect.

One study of the agricultural data system suggested the uncoupling of agricultural census data collection from its five-year cycle, redesigning it as a series of annual sample surveys in which sample size can be more closely related to the desired level of statistical reliability (American Agribusiness Associates [1973]). This would reduce the statistical design and organizational inefficiencies created now by the peak-load problems of the five-year census cycle. This approach might be implemented for the data now collected in the agricultural census whether the responsibility is assigned to the Bureau of the Census or the USDA. Like any other statistical design, this approach would also have some inherent limitations. If an adequate investment is made by the Bureau of the Census in developing a satisfactory list frame and in reconceptualizing the food sector paradigm, then there would be a genuine organiza-

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



tional advantage in keeping these data functions in the Bureau of the Census where they could be conceptually integrated with the other economic census statistics, especially those for nonfarm agricultural firms.

One factor often missed in attempting to understand the data quality problems of the agricultural census is the great increase in the number and complexity of the questions asked in the agricultural census during the period since World War II. This is a reflection of the growing complexity of the food and fiber sector itself under the fragmenting impact of progressively greater industrial specialization. Specialization of production processes invariably leads to greater informational requirements for the coordination of fragmented production and marketing processes. In short, the explosive growth in the need for food and fiber sector information for both private and public decision making appears to have grown beyond the capacity of the Census of Agriculture to sustain as a statistical vehicle. Rising complaints of respondent burden combined with an increase in distrust of government has resulted in congressional pressure on the Bureau of the Census to reduce the respondent burden substantially in planning for the 1978 Census. This leaves the Bureau of the Census with a Hobson's choice of either drastically reducing the number of questions and, therefore, the volume of useful information or greatly reducing the size of the sample, perhaps completely eliminating the capability for producing reliable county-level agricultural data—or some combination of the two alternatives. If the Census ceases to produce county data, its primary rationale for existence as a unique statistical vehicle will have ended.

## Statistics for Rural Society

In the case of social and economic statistics for rural society, the overpowering problem, as the AAEA Committee on Economic Statistics [1972] pointed out, is the lack of data. This often is because there has been no demand for the financing of data collection. But even in areas of increasing public concern, as in rural development and natural resource management, and in the various dimensions of human welfare, little coherent data and few well-developed information systems exist. The primary reason is the absence of a satisfactory conceptual or theoretical base for either data collection or analysis. Economists cannot even define adequately what is meant by economic or rural development.

## Institutional Obsolescence

Rapid or steady long-term technological, organizational, and associated value changes not only create obsolescence and mismatching in the conceptual base but also in the institutional structure of statistical systems. This is often compounded by the reorganization or development of new administra-

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



tive structures without adequate care for the integrity or capability of involved data systems. Changes in basic statistical measurement techniques (for example, shifting the agricultural census from complete enumeration of an area frame to list frame surveys) which are unmatched by an implementing organizational adjustment can create another form of institutional obsolescence and inefficiency (American Agribusiness Associates [1973]). As a result of institutional obsolescence or reorganization, current administrative structures often do not bring the necessary information together at the time and places in the structure where it is most needed by decision makers.

## Vested Interests in Data

Changing the design of data or its collection and the related information system always involves property rights, some publicly held and others privately held. The redesigning of information output and availability always redistributes those property rights and thus is a difficult feat to achieve in the face of the vested interests in data and information.

Bureaucracies, those who staff them, and various user groups and clientele develop substantial vested interests, not only in specific data output but also in existing concepts and measurement procedures. Thus, they behave as if they had a property right in certain data or analysis systems and often are able to enforce their interest politically. Any change in the design of the system must face this problem as a cost of eliminating or replacing an old statistic with newly designed data. The same observations can be made about any attempt to modify an established analytical process. Arrow [1974] rightly characterizes this as one of human capital made obsolete by change. Information systems are the human capital side of the development process. When human capital becomes obsolete, its replacement or redesign is far more difficult than the replacement of physical capital.

Some data problems arise because many of the property rights vested in information are privately held by firms and interests with considerable economic and thus political influence. As we attempt to redesign or create new data to respond to the public interest in problems of international trade with the Soviet Union or China or in new public policy issues involving the behavior and performance of the food and fiber sector, we find that essential information is often held by a few firms whose immediate interests would not be served by the release of that information. As industrial concentration continues to grow in food and fiber markets, the issue of private ownership of information versus the public's right to know will become more and more critical and heated. Giant firms acquire with their great size not only an impact on markets but also a major responsibility for public information. Where the data on a market are collected from and distributed to firms by a trade asso-

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



ciation, the tendency to withhold data from the public is even greater (Stigler [1961]).

## Empiric Failure in Design and Collection of Data

Another problem is the increasing tendency of economists to propagate endless theories, concepts, and models of unknown value because they fail to design and collect data for an adequate empirical test. In his 1970 presidential address to the American Economic Association Wassily Leontief indicted economists for this failing. Leontief faulted economists for being satisfied with secondary data which do not match and thus cannot adequately test their theoretical concepts. His point was that theory will never be improved without empirical test, and in its absence, economists are playing sterile games.

Variations on Leontief's criticism have been voiced in many such addresses by economists (Bergmann [1974], Blackman [1971], Hahn [1970], Phelps Brown [1972], Maisel [1974], Worswick [1972]). Bergmann argued that the situation is worse than Leontief imagined: ". . . these days the best economists don't even look at secondhand data; they get them on magnetic tape and let the computer look at them. Economists have voluntarily set for themselves the limits on data collection faced by students of ancient history." Rivlin [1975] has lamented that "disdain for data collection is built into the value and reward structure of our discipline. Ingenious efforts to tease bits of information from unsuitable data are much applauded; designing instruments for collecting more appropriate information is generally considered hack work."

Leontief pays a high compliment to agricultural economists by explicitly exempting agricultural economics from his indictment. He describes the discipline as "an exceptional example of a healthy balance between theoretical and empirical analysis and of the readiness of professional economists to cooperate with experts in the neighboring disciplines . . ." However, the AAEA Committee on Economic Statistics argued in 1972 that the honor Leontief accords us "properly belongs to an earlier generation" and that agricultural economists are now falling into the same errors which Leontief ascribes to other branches of the economics profession.

The capacity and reputation of agricultural economics were built around a balanced investment in theoretical and empirical analysis. We have now lost much of our early interest in the design and collection of data and often fail to collect needed data or to respect those who do. There is evidence that we are failing also to update our conceptual base at a pace sufficient to keep up with major changes in agriculture. Conceptual failure directly undermines the deductive process of knowing, and empirical failure directly undermines the

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



inductive process of knowing. Thus, we must contend with two kinds of failure, either of which could prove disastrous if not corrected.

## *Data, Analysis, and Information: A Paradigm*

One of the first problems encountered by the AAEA Committee on Economic Statistics [1972] was a confused but common vocabulary which erroneously equates data with information and fails to differentiate the distinctive steps in the process by which data and information are produced. There also seems to be no clear understanding of how the analytical process or system of inquiry over which the agricultural economist presides relates to data collection and to the information system.

### The Nature of Data and Data Systems

Every data system involves an attempt to represent reality by describing empirical phenomena in some system of categories, usually in quantified form. Data are the result of measurement or counting, but when one sets out to quantify anything, the first question that must be answered is, "What is to be counted or measured?"[2] If the configuration of data produced is to be internally consistent and to have some correspondence with reality, the quantified ideas must bear a meaningful relationship to each other and to the reality of the world being described. In other words, there must be some concept of the reality of the world that is to be measured. Reality is nearly infinite in its variation and configuration and must be simplified or categorized if the human mind is to handle it in a systematic way. Thus, in producing accurate data, one either implicitly or explicitly develops a set of concepts which in some significant degree is capable of portraying and reducing the nearly infinite complexity of the real world in a manner that can be grasped by the human mind. Data are a symbolic representation of those concepts. If the concepts are not reasonably accurate reflections of that real world, then no amount of sophisticated statistical technique or dollars invested in data will produce useful numbers. (See figure 1.)

Although data presuppose a concept concepts cannot be measured directly (or, in a strictly logical sense, cannot be measured at all). Rather, we make the concepts operational by establishing (defining) categories of empirical phenomena (variables) which are as highly correlated as possible with the reality of the object of our inquiry.

Thus, there are three distinct steps which must be taken before one can produce data which purport to represent any reality. These are (1) conceptualization, (2) operationalization of concept (definition of empirical variables, and (3) measurement. The failure or deficiency of any one of these

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



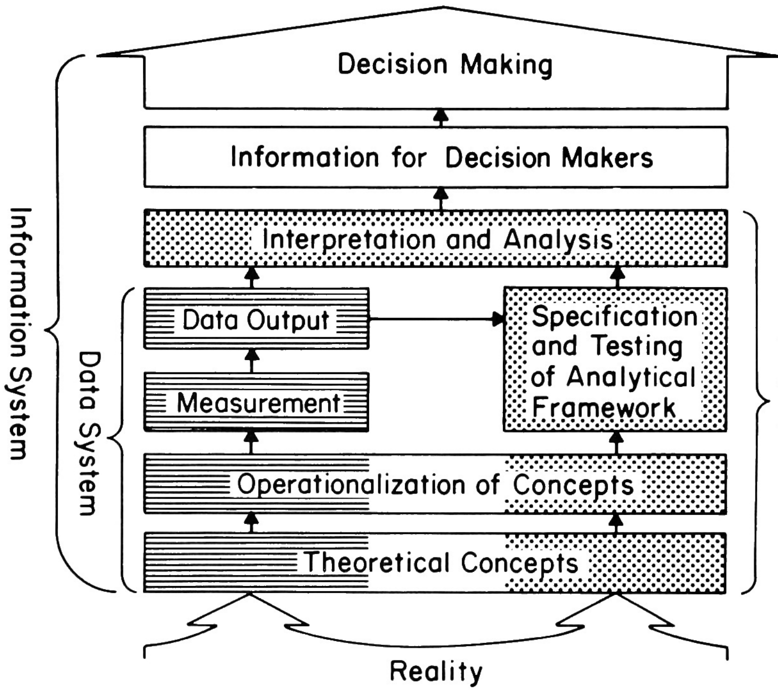


Figure 1. An agricultural information system.

data system components constrains the quality and characteristics of the data produced. An inadequacy at any stage can be offset only to a very limited extent by improvements or manipulations at the other stages. Thus, the great improvements in statistical methodology and data-processing techniques over the last generation cannot offset failures at the conceptual level, for no matter how well one measures and manipulates the numbers one will still be measuring the wrong thing. For example, the parity price concept, no matter how well measured, is a poor representation today of farmer welfare. The "cost of production" concept central to the operation of the Agriculture Act of 1973 is so inadequate as a representation of the complexities of farm cost structures that no amount of genius in making it operational or measuring it can redeem its inadequacy as a concept.

It is worth noting that the term "reliability of data" has three possible meanings in this paradigm: reliability of measurement, which is the meaning the statistician usually attributes to the term; reliability of operationalization; and conceptual reliability.[3]

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



## The Nature of Information

Data are not information (Eisgruber [1967], Dunn [1974]). An information system includes not only the production of data but also analysis and interpretation of these data in some purposeful policy decision or problem solution context. The demand for data is generated by the need to make decisions on problems, but decision makers rarely use raw data. Rather, there are intervening acts of interpretation ranging from statistical and economic analysis through less complex program and political evaluations, and the like which transform data into information by placing them in a specific problem context to give the data meaning and form for a particular decision or decision maker. Data are symbolic artifacts which acquire most of their meaning and value from the context and design of the information system in which they appear. Information, then, is a process which imposes form and gives meaning. Thus, an information system includes a data system, the analytical and other capabilities necessary to interpret data, and, finally, the decision maker.

## Analysis as a Function of Information

What does the agricultural economist do when he plays the role of analyst? In our training we all acquired much the same epistemological sense of how to analyze and solve problems; that is, there is a base of theoretical concepts, a body of theory purporting to represent reality which we make operational through the definition of variables, often specified formally in a model which must be matched with data or measured representations of the variables. The model or analytical framework is then tested against the data and conclusions are drawn. Thus, in these three steps in analysis, we find two of the same components observed in a data system—that is, theoretical concepts and operationalization of those concepts.

Thus, in data systems (left side, figure 1) and in analytical systems of inquiry (right side, figure 1), we operate from the same set of theoretical concepts and, ideally, the same set of definitions which make those concepts operational. Unless economic theory and economic statistics meet on common conceptual ground, there can be no mesh between empirical analysis and theory—between the inductive and deductive processes.

Agricultural economists—especially the academic economists—are clearly responsible not only for the design and maintenance of the profession's analytical framework but also for the design of the conceptual base of the data systems which provide the empirical content for analysis. The notion held by some economists that statisticians alone are responsible for the design and production of data represents a gravely distorted view of our professional re-

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



sponsibilities. It reflects an epistemological weakness and also a lack of understanding of the historical development of data systems. From earliest times data systems have been created to solve specific problems, and specialists in the appropriate areas of knowledge have always been involved in the design of the data systems.

An information system is the total process by which knowledge is generated and brought to bear on social decisions—public and private. As social scientists and statisticians, we are concerned with social information processing. The design of the information system establishes the nature of the relationship between the decision maker, the information on which decisions are based, the analytical process which transforms data into information, and the design and collection of data.

## The Imperative of Information System Design

It is the conscious design of the information system as a system which must be fully respected if data are to be accurate and if the information upon which decisions are based is to be relevant and reliable—or even available. One of the most fundamental errors which we now make as agricultural professionals is the failure to perceive and design activities as subsets of the information systems which give them their meaning and significance. This is a design failure which takes different forms.

There are many different "ways of knowing," or epistemological positions. The validity of any one depends on its teleological context. Thus, the purpose of the system of inquiry controls or limits the appropriate epistemological basis of any specific information system (Churchman [1971]).

All information systems are problem solving or purposive because they are subsets or components of social systems which are designed for some problem-solving purpose. Thus, data collection and analysis always have a purpose and can only be understood fully in a social system or decision context.

Data collected for societal decision making must have a social theory base. No matter how ad hoc the collection of data may seem, every measurement act is guided consciously or unconsciously by conceptual and value structures which exist prior to the act of measurement. Data and information are never value-free or theory-free. Conversely, all concepts or theories have a prior empiric basis that is explicit or experiential. Theory and data are, thus, epistemologically interdependent.

For this reason any quantitative or statistically based information system exhibits both inductive and deductive epistemological bases. Consequently, in an information system we do not *know* anything until, as a necessary condition, a deductive analytic mode of inquiry is tested against and combined

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



with an inductive empiric model of inquiry. What is known from such a process grows in extent and reliability by a repetition of interaction between the deductive and inductive modes, in which both the analytic content and the empiric content of the process are reformulated and improved on the basis of what is learned from each prior iteration.

An analytical hypothesis or model and the data for its empirical test must have the same conceptual and definitional base. This is perhaps too logical and obvious to mention, yet a failure to appreciate this fact lies at the heart of our apparent inability to understand and deal with the problem of the accuracy of information provided in agricultural economics. It also is related to the progressive deterioration in the sense of professional responsibility on the part of economists (including agricultural economists) for the design of the data they use.

The points made in the last three paragraphs are implicit in Leontief's insistence on the necessity for empirical testing of all theoretical formulations with data which are designed around the proper concepts. They are also implicit in the statement by the AAEA Committee on Economic Statistics [1972] that accurate and useful data can be collected only in a conceptual frame which is an accurate representation of the reality the data are supposed to describe. Data are symbolic of some phenomena which they are designed to represent. The quality of that representation is only as good as the adequacy of the conceptual base or its operationalization or its measurement.

When the phenomenon that is being represented changes rapidly, as it has in the food and fiber industry, the conceptual base of the information system must be redesigned frequently to keep up with the changes in reality and the problems being studied. If the rate of change is high enough, the need for conceptual redesign becomes nearly continuous. This is the fundamental problem in the design of information for agriculture. Failure to keep up with the changes in the policy agenda and in the reality of agriculture leads to significant conceptual obsolescence, and the system begins to lose its capacity as an accurate guide for problem identification and solution or management. This paradigm of the constituent processes of an information system provides a conceptual template with institutional analogues for the design of data and information systems.

## *The Design and Management of Information Systems*

Let us turn now to some of the general information system design and management questions raised by the development or industrialization of agriculture, by economic organization differences, and by the behavior of various information system participants.

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



## Development and Information

Information systems are an essential part of any decision process, public or private. Thus, the social returns on any investment in information are derived from the benefits generated by that information in the decision structure of a social system. Information systems appear to have a unique role in an industrializing or developing society. Without an adequate information system the potential gains in productivity from specialization and new technologies are lost in inadequate coordination and management of the developing industry or economy. As one moves into a developmental mode and begins to set conscious national goals for economic and social development, the need for data and the social value of that data increase greatly. Agricultural information systems have played a strategic role in the growing productivity of this country over the past century. Although it is generally understood that industrialization and development increase the demand for information and the social returns, it is not often recognized that development also brings about a change in the kind of information demanded.

The earliest systems for data collection in any society usually arise out of administrative and management needs. The data required can be described as primarily static and descriptive and as involving clear, relatively fixed goals and simple or low levels of information processing. As a society's economic structure grows more complex and specialized, the demands are not just for more data and greater accuracy in the articulation of detail. Increasingly the demand is for data and information in a "learning or developmental mode" (Dunn [1974] ), in which the goals of decision making are not completely specified; and one purpose of the information system is to assist the decision maker in specifying the goals in a progressively more complete form (that is, in redesigning the information system). In a developmental mode goals and problems may continue to change as learning takes place and thus may never be completely or finally specified. It is obvious that one is not well served in this situation by data and information which are basically static.

Note too that in the learning or developmental mode the information system which perceives and acts on data is itself changing in structure and behavior in response to the information it processes. Thus, the information system must be capable of perceiving changes not only in the environment but also in itself, even under conditions in which such changes themselves become goals (Dunn [1974] ).

As if this were not demanding enough, in the most industrially and agriculturally advanced countries where the reality of the world (especially in agriculture) continues to change rapidly, the need to redesign the system eventually becomes continuous. It follows that if agricultural information is

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



to be accurate and reliable the capacity for redesign must be a normal internal function of the information system. If the designer does not become a part of the system in this situation, the capacity of the system to produce useful information will deteriorate.

Another significant observation can be made about the design of information systems. Any system designed to solve problems will inevitably combine and use different fields of knowledge. Therefore, the concepts underlying the information system will be derived from different disciplines. Agricultural information systems are an excellent example. If such a system is to produce useful data and, in the process, manage its own continuing redesign, a general "theory of social information processing" or theory of theories, a "meta-theory," is needed. In other words, we must have a means of synthesizing concepts from different bodies of knowledge into a meaningful relationship to each other (Dunn [1974] ).

A meta-theory for information system design may well be an impossible practical goal. But the logic of its necessity is valid and has the virtue of keeping in front of the designers of information systems the true complexity of the task. The design of data and information systems is not a job we can assign to any but the best minds. The theory of information from a single discipline such as economics or statistics is not adequate to the task.

## Economic Structure and Information

In the design of information systems the configuration of major economic sectors in the society and the degree of concentration of these economic structures make a great deal of difference in public and private sector information needs. As one looks across the entire economy, from the agrarian sector to the most industrialized manufacturing sector, it is evident that differences in concentration or industrial structure have great impact on public and private sector data needs.

In some economic sectors an industry or firm can recapture the gains from data collection and analysis financed by that industrial association or firm. The government has little business collecting this kind of data unless, as is sometimes the case, there is an urgent or overriding public interest in such data. There are other economic sectors where the benefits of private investment in data can never be captured by an individual firm or even an entire industry. The difference is found primarily in the nature of the industrial structure itself.

If an industry is a monopoly involving only one firm, then the benefits of any investment in statistics for that industry will accrue directly to the private gain of the monopoly. Thus it can afford to pay for, and can expect to

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



get, the benefits of any investment in information needed to manage the industry. The more concentrated the industry, the greater the returns to private investments in data. And the more concentrated the industry, the less the justification for public investment in data for private use. Over the same continuum the justification for public investment in data for public use at first declines from significant levels, but as higher levels of industry concentration are approached the need for data and its social rate of return should rise again. This is especially true if the society has policies which constrain or regulate monopolies.

At the other extreme of industrial structure, where an industry may be made up of thousands or even millions of independent firms as in agriculture, the amount of private sector investment in data collection and analysis that can be justified by an individual firm (because it can be recaptured in the firm's balance sheet) would be extremely small, if not zero. In fact, the public returns to private investments in information would probably exceed private returns.

Increasing demands for data are generated by the greater and greater specialization that has resulted from the modernization or industrialization of agriculture. Social returns on the investment in data for improved coordination and management of specialized industrial processes are usually very high. In very competitive unconcentrated industries such as agriculture these gains are only realizable through public investment. The great increases in productivity in United States agriculture over the past century can be traced in substantial part to the contribution of publicly collected data for private management decision making. The improved efficiency in the use of resources has accrued to the society in the form of lower food costs and the release of most of the farm labor force for employment in nonfarm pursuits. Returns to information and coordination in the development process have always been high, particularly in atomistic economic sectors such as agriculture (Hayami and Peterson [1972] ).

It is thus no accident that most national governments collect far more detailed statistics on highly competitive industries such as agriculture than on highly concentrated industries such as steel. This is a logical allocation of public resources which follows from the very nature of the industrial structures themselves.

## The Dysfunctional Behavior of the Actors: A System Management Problem

The difficulty with most of the literature on the problems of agricultural data and information systems is that the authors often overlook the forest and see only the trees. There has been a general failure to see any aspect of

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



these problems in their systems context. We need to think more about the nature of the social systems of which information systems are an inherent part. This does not keep us from discussing data as such, but it is well to stand back occasionally and look at the whole as we try to diagnose problems. We are only beginning to do this and no one can claim more than a partial understanding of the difficulties we face.

The initial designers of an information system will usually perceive the whole of the system. They will at least understand the system well beyond the subset they inhabit. In the early stages of the development of an information system there is limited organizational specialization, and often no professional statisticians are available. In fact, it is a rare administrative data system which has a statistician in attendance at its birth.

As the agricultural sector of the United States has developed, its public and private agricultural information systems have exhibited progressively greater specialization in organization and growth of bureaucracy and professional staff. This is an inevitable consequence of growth and creates a situation in which the information system management problem is largely one of managing bureaucracies and highly specialized professionals.

The skills of different types of professionals are necessary for developing sophistication and capacity in any information system. Nevertheless, the attitudes of professionals and various professional groups increasingly seem to affect the design and functioning of agricultural data systems in a deleterious manner. As science and society have become more and more specialized, and as specialized vocations have been transformed into professions, both specialization and professionalism have begun to have an effect on the perception of the nature and role of data systems and analysis.

Once professionalized, the agencies which produce statistics tend to see data as an end in itself—unless the agency is tied very directly to client or user groups. Many statisticians have only a limited grasp of the analytical methods and the information system needs for which specific data sets are inputs. Statisticians as professionals also tend to view their responsibility in the data system as limited to the application of statistical methods to the production of data (in their view, the final product). Thus, to many statisticians the improvement of data means little more than the improvement of the professional quality of statistical agency performance, which is quite important but is not a very large part of the total design problem in data collection. As organization men they also believe in rapidly growing budgets and freedom from interorganizational commitments and restrictions. Such narrow perceptions of professional responsibility get in the way of adequate information system design, coordination, and management. Nevertheless, the statistician's performance must be judged to be better than that of the economist.

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



Economists have come to be so specialized that their common attitude toward data and data systems is narrowed to a bored yawn—it's someone else's responsibility. Frequently they do not today understand the nature of the relationship between their analysis and the data systems upon which they must draw, particularly as it concerns the quality of data. Few understand the responsibility they have for maintaining the conceptual foundation of that data base. Indeed, judged by how they behave rather than by what they say, economists with depressing frequency fail to perceive that the theoretical structure of their discipline is also simultaneously the conceptual base of the data system which produces the numbers they use. To repeat, economic theory and economic statistics unavoidably meet on a common conceptual ground. If they do not, there can be no mesh between empirical analysis and theory. As this mesh deteriorates, the capacity of the profession for doing accurate analysis also deteriorates in an equal degree (Morgenstern [1973], Nieto-Ostolaza [1973]). Often contempt is openly expressed for economists who spend time on the collection and design of data. There is thus a failure to understand the nature of the data system and the epistemological interdependence of data systems and analysis. This, of course, is an oversimplification—and there are honorable exceptions in all the disciplines—but increasingly the consequence is that few social scientists perceive the information system within which they work and, as a result, almost no one acts to manage, improve, or renew the system.

The attitudes of many political appointees or managers in the federal government, including those who direct the varied elements of national data collection and analysis, also often have a very unfortunate impact on information systems. Information tends to be viewed by politicians as a free good, which because of the scale and complexity of the federal executive and its great capacities, should be forthcoming without cost and on demand, no matter how esoteric or unusual that demand for data may be. In commenting on the quality of the data and information produced in response to the typical political demand for instant data, a colleague once remarked of the political managers of his agency, "When they want it bad, they get it bad!"

There are several other traits that tend to be quite disruptive of a well-designed and well-managed information system. Agencies are repeatedly reorganized with little care for the structure and integrity of the information and data systems involved. After three or four reorganizations one finds that what may once have been a coherent data or information system is now scattered across the agency, and if it is coordinated at all it is by individuals acting informally and not by the structure of the system itself. The older the agency, the greater the validity and frequency with which this observation may be made.

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



The effective time horizon or planning span of political decision makers rarely goes beyond the next budget or the next election. They will assure you (and rightly so, by their own standard) that anything that exists as a potentiality at any greater distance in time is not a real or urgent problem because the probabilities are high that the problem will not materialize in the form envisioned or that the political decision maker will not be there or will no longer be responsible. The political and budgetary costs of changes in statistical systems must be faced immediately. The benefits are rarely perceptible except in periods of time running well beyond the next budget and the next change in administration. This often severely disturbs the management of agricultural data systems, particularly when the systems are simply subordinate parts of administrative structures and are not organized as formal data systems.

Kings, pharaohs, and khans of ancient times used to behead the bearer of bad tidings. Today political decision makers still have a tendency to sack, demote, or punish those who produce data that are embarrassing or that make them uncomfortable. In a modern bureaucracy this is simply transmitted down through the hierarchy so that ultimately some statistician or economist is made the goat. Farm income is a politically sensitive statistic in domestic agricultural policy. Through administration after administration, various political decision makers have blamed statisticians and economists for all the deficiencies of that statistic. Yet when examined closely, one will find that the statisticians and economists—the technicians—have for decades recommended remedial action, generally with limited results because the political decision maker in that generation refused to assume the costs of making the change, thus transmitting to another political decision maker in a subsequent administration (with luck, of another party) the even greater political costs of failing to reform the system. Perhaps the immediate political decision maker's hands are tied, but they should at least stop blaming statisticians and economists for the failures of politicians.

## *Conclusion*

It is worth reflecting on the impact of progressively greater specialization of society and subsequent organizational fragmentation on information system design. The problems of any society dominate its policy agendas and the information systems which are the basis of society's capacity for problem solving. But when specialization begins seriously to fragment the social organization, the scope for externalities in the society grows and with it the problems of the society. However, the capacity for problem solving tends to decline because the same social organizational fragmentation also

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



shatters the information systems, making it more difficult to maintain a coherent, integrated information design for any problem purpose. Perhaps it is this phenomenon which we are seeing in the dysfunctional behavior of information system actors and in the lack of integration today in many agricultural information systems. In any case, it makes evident the serious need for information system design and management.

No information system has unlimited capability. The act of design is one of progressive elimination of some potentialities in order to sharpen other specific capabilities. The purpose of social system decision processes, which information must serve, provides the primary principle of information design.

There are two distinctive but interdependent parts of the design process. One involves the design of the system, the other the design of information proper. Specification of the system involves elimination of the impossibilities and the potential system subsets that are not relevant, given one's purpose and context. Specification of the system also involves the design of the institutions within which data are to be designed, collected, analyzed, and used in the decision process. Many of these decisions are made without reference to data or analysis, since they are primarily political or social decisions.

The design of information involves decisions on what information to collect and analyze and how to do it. The purpose of the system provides some of the decision criteria, but the dominant element within those limits is found in the economics of information. Thus, the assessment of the cost of a bad or erroneous decision versus the cost of information becomes critical.

My objective here precludes an adequate discussion of the complex and important problems of the economics of information. But it is worth noting that in any social system the greater the level of uncertainty (up to a limit), the higher will be the value of information. Appropriately designed information allows one to reduce uncertainty and to manage its undesired consequences. But uncertainty is inherent in the human condition. While "sufficient expenditure" on information will keep the effects of uncertainty "upon people . . . within tolerable or even comfortable bounds . . . it would be wholly uneconomic to eliminate all its effects" (Stigler [1961] ).

World food and feedgrain stocks vanished in 1972-73. The consequent price instability and production fluctuations in subsequent years have dumped every nation's food and agricultural policy into a sea of uncertainty. The value of information has increased many times over, thus exposing more clearly the many weaknesses in our information systems and giving rise to a call for construction of a world food information system (United Nations, World Food Conference [1974] , U.S. Congress, Office of Technology Assessment [1976a] ). During the past several decades of excess stocks and shelter from market uncertainty, we have undervalued our agricultural information sys-

Copyright © 1977 by the American Agricultural Economics Association. All rights reserved.



tems so greatly that we have not invested adequately in some systems and we have allowed others to decay seriously.

Information is an expensive commodity as well as a valuable one. Returns to careful decisions about data and information are high. In the search for an effective information system the economic and statistical models, the estimation and optimization procedures, and the corresponding inferences and choices are interdependent links in the information chain. The opportunity decision cost of considering any one of the above ingredients in isolation is very high.

The cost of poor decisions and subsequent lack of appropriate information is extremely high (Bonnen [1973]). The foundation of effective information management for agricultural decisions is careful design of data and information.

## Notes

1. Conceptual obsolescence is not limited to agricultural statistics. All of our older social and economic statistics share in this problem. It is also obviously a difficulty that will continue to plague all data systems involving social and economic behavior in a society in which change is rapid.

2. Data, strictly speaking, are not limited to quantified forms, but this discussion is confined to statistical data. Implicit in the question of what is to be measured is also the question of why.

3. This observation was contributed by L. V. Manderscheid.