Ag. Econ. Report No. 87-54

1987

The Status of Agriculture Data Systems As the Basis For Policy

by

James T. Bonnen
Michigan State University

Agriculture -- Statistics

Presented to the Symposium on Relevance of Agricultural Economics Obsolete Data Concepts Revisited at the annual meetings of the American Agricultural Economics Association in East Lansing, Michigan 3-5 August 1987.

THE STATUS OF AGRICULTURAL DATA SYSTEMS AS THE BASIS FOR POLICY

by

James T. Bonnen
Michigan State University

My assignment is to provide an assessment of where we stand today, 15 years after the first Report of the Economics Statistics Committee in 1972 (AAEA). A real assessment would require a full scale research project which, of course, I have not done. Consequently, what follows is impressionistic and judgmental limited by my own experience and understanding of the changes that have occurred.

Let me remind you of a few major points from the 1972 Report and related work by myself or sponsored by the Committee (Mayer and Ahalt; Carlin and Handy; Weeks, Schutter and Southard). That 1972 Report said that the data base for agricultural economics was suffering from serious obsolescence, primarily conceptual obsolescence. Obsolescence can occur in any of three critical steps in the process that produces information (Bonnen 1975). The first is the development of a concept or set of concepts descriptive of and relevant to the real world phenomena or problem under consideration. Second, since concepts are complete abstractions and lie beyond the veil of the material world, one must select a set of variables or proxies from the real world that represent those conceptual ideas, if anything is to be measured. The decision about what dimension of the real world to measure is critical. The information system can fail entirely, if the representation of the concept is totally inadequate. Finally, there is the act of measurement itself. This involves statistical design and data collection. Thus, the concepts selected, the dimension of the real world chosen to represent those concepts, and the act of measurement, any one or all may be the origin of data system failures.

The Economic Statistics Committee was established in 1970 to examine the criticisms by some analysts that data were performing less well. It was alleged that statisticians were messing-up in their measurement activities. What the Committee

found was that everything statisticians were doing could only have improved the data. The declining performance of some major data sets was caused primarily by economists failure to update the concepts underlying the data, which had been made obsolescent by the great changes in agriculture since the concepts had been developed.

The forces that create obsolescence in an agricultural data system can come not only from changes in the organization and technologies of the food and fiber industry but also from changes in the agenda of food and fiber policy (public and private). Both can change the questions the information system is expected to answer.

In 1978-80, while Director of the White House Reorganization Project for the Federal Statistical System, I began to look at decision failures to discover the origins of these failures in the various components of the information system. For every case of a data or measurement failure, I found (very roughly) ten failures of analysis. In other words, for every decision failure that could be attributed to statisticians, there were about ten that could be attributed to economists or other analysts. In turn, for every example of an analytical failure, I found something like ten or more failures arising out of the policy makers own behavior. Decision makers often allow their ideology and prior policy commitments to be the father of action, interpreting the output of the information system in ways that are selfserving and distorting. This is easy to do since there usually is a considerable amount of noise and ambiguity in the signals from the information system. I have since discovered that studies of intelligence systems demonstrate the same pattern of failure (Betts, 1978).

Let me turn to the question of whether we have improved the data base since 1972 or allowed it to continue deteriorating. I find this task is nearly impossible. In assessing conceptual failure, you are examining something analogous to cultural lag. That is to say, there is always some lag in adjustment that is unavoidable or inherent in any social system, so there is some (unknown and varying) minimal conceptual lag which one will always have to endure. In addition, the agricultural data system is composed of hundreds

of subsystems that are organized and managed separately, are affected differently by change, and receive different levels of budgetary and professional attention. Thus, talking about the data system as a whole makes limited sense. The question, as a practical matter, becomes one of specific data bases and whether the lags are excessive. Even this is very difficult to measure or judge. As an objective matter, aggregating all this for the sector is probably impossible.

Let me suggest two examples where I do, and do not, see improvement. In 1972 the agricultural income statistics were substantially different from and incompatible with the rest of the national income accounts and not too defensible in some agricultural policy uses (Carlin and Handy). Ironically this arises out of the fact that agricultural income data were developed to a higher level of sophistication earlier (in the 1920s and 1930s) and reflect different purposes and conceptual investments. The national income accounts were developed later and have been steadily improved over the years. By the time of the Economic Statistics Committee Report in 1972 the conceptual mismatch of agricultural income measures with the national income accounts was blatant. Since that time substantial improvement has been made in the agricultural income accounts. The USDA has cleaned up many of the accounts that, for example, confused stocks with flows. The agricultural accounts are now much more consistent with the value-added concept that underlies the national income accounts as well as more consistent conceptually with their multiple uses in agriculture. While there are still some problems remaining, these are not as extensive or debilitating as they were 15 years ago.

Another major improvement is represented by the Farm Cost and Returns Survey (FCRS). This survey merges two older surveys and integrates farm input production data with financial account data. The FCRS is a national probability sample of all farms providing microdata sets that may be aggregated into almost any format that one chooses, whereas the prior statistical base was composed of several separate data bases, contained large assumption-dependent, synthetic elements, and was very limited and

inflexible. The classical problem of aggregation still has to be handled carefully. Access to these probability-based microdata sets is limited: one must work with them in the USDA under conditions that protect confidentiality of respondents. The sample size still has to be expanded to allow all state aggregates and many important national financial subaggregates to be measured at a reasonable level of statistical reliability.

The FCRS is a good example of how you have to run fast just to stand still. The structure and organization of agriculture has probably changed more over the last 15 years than it has in any comparable period. Agriculture is now much more heterogeneous and production more concentrated. The significance of this lies in the declining capability of any statistical aggregate to represent the reality of a sector when that sector is characterized by increasingly greater heterogeneity within itself. One needs subaggregates and microdata sets to understand what is happening. When some class of activity exhibits heterogeneity that exceeds that which differentiates it from other categories of activity, one must face the question as to whether that category is still valid. The statistical or information imperative that we face is one of describing the sector in more detail at lower levels of aggregation than before. Thus, we now face a growing demand for subsector aggregates and microdata. It would be a fair judgment that while important improvements have been made, we still lag well behind the curve in contending with the problems of obsolescence, conceptual and otherwise.

With the rapid concentration of farm production on a few farms, another major problem looms in the near future. This is the matter of falling response rates to farm surveys as respondent fatigue rises with repeated surveys of a smaller and smaller universe.

Let me focus on the question of what it is that one should measure as a real world representative or proxy for any given concept. It has puzzled me why so few examples of data system problems have arisen that are attributed to this part of the information system. I have recently come to the conclusion that we have all been confusing

conceptual issues with the decision on what real world proxies should be measured. Many problems said to be conceptual are, when examined, discovered to be questions about what is or can be measured to represent a concept. It is often difficult to separate the selection of the proxy from the adjustments that we make in the concept and in the proxy in order to get a better match between the two. In fact some economists do not seem to see this as two distinct decisions. Thus, some of the problems that get identified as conceptual obsolescence may more accurately be seen as having arisen out of the decisions that we made in selecting real world proxies for measurement. In part at least, this confusion arises out of what I have described elsewhere as the decline of the commitment of the profession to the empiric (Bonnen, forthcoming). I have argued that we over-emphasize statistical technique and conceptualization or theoretical modeling while depending on less-than-ideal secondary data. In effect, we are skipping the step of selecting for measurement the most appropriate real world proxy for the theoretical constructs being used. I believe this confusion has played a central role in the problems that we face in maintaining an effective agricultural data system.

I remain of the opinion that we have done better in the actual measurement aspect than in any other. This achievement is limited primarily to the USDA. The universities as a group have withdrawn from data collection compared to the commitment they had, say in the 1940s or 1950s (see Bonnen, forthcoming for a discussion of reasons). I have already mentioned the improvements in the national income accounts and the Farm Costs and Returns Survey, where the USDA has made major innovations that include both conceptual as well as measurement improvements. ERS is responsible for constructing the USDA's economic indicators while the National Agricultural Statistics Service (NASS, formerly SRS) has the major responsibility for gathering economic statistics in the USDA. NASS now has almost all of their national collections on a probability sampling basis, where 15 years ago a substantial portion of these collections were nonprobability surveys. In addition they have begun to integrate their many surveys into a few common

collection vehicles increasing comparability and the capacity to combine data subsequent to collection. This involves not just an improvement in measurement, but improved comparability of the concepts involved. They have also substantially increased quality control for certain surveys. There is underway a review of the Agricultural Reporting Board (formerly the Crop Reporting Board) procedures from which one can anticipate improvements in quality and efficiency of estimates as well as the ease of interpretation of some of the statistical estimates.

On the other hand, there have also been losses in the measurement area. The frequency of many collections has been reduced. Labor as a farm input has practically disappeared since its measurement is now limited to an annual survey in July, which cannot possibly be representative of the labor input for the year. The sample size of some surveys has been reduced and therefore the reliability of the data, especially for subnational aggregates. Some series have been dropped entirely. These typically are "minor crops" (primarily fruits and vegetables) that individually may be relatively unimportant in gross sales at the national level, but together account for 14% of gross farm income and which in any number of states in the United States are major farm income sources. This latter raises the point, that I have made before, of the growing demand for small area agricultural data as specialization of production proceeds. The substantial increase in state and local government development effort has also contributed to the demand for local area data. Budget constraints on statistical agencies have led to substantial losses in the last ten years. Many data are simply not available today that were available a decade ago.

The Census of Agriculture is another data base which was seen as deficient in 1972 (AAEA, Bonnen 1977). Since then the Census of Agriculture has been moved to the same five year cycle as the major economic censuses that report on the rest of the food and fiber sector. While this has improved temporal comparability, there still are major conceptual incompatabilities. Eliminating these was supposed to have been the goal in

shifting the Agriculture Census from the demographic to the economic fields of the Census Bureau. No action has been taken to develop a food and fiber sector paradigm as was suggested in the early 1970s, so that the farm input and output marketing activities could be integrated into a complete description of the food and fiber sector. Perhaps this conceptual challenge is primarily a USDA and agricultural economics profession responsibility, but the fact remains that we continue to lack performance indicators for the food and fiber sector. The Census of Agriculture thus remains a census of farms -- a concept that is empirically more heterogeneous today than in 1972.

Perhaps the greatest direct loss of data in the Census has been the elimination of the follow-on surveys to the Census to capture the growing specialized services and inputs for agriculture. The Census is managed far better today and is available on time, although the continued changes in design and collection methods from one census to the next still leaves those who use Census data as a time series with a substantial lack of comparability between each Census of Agriculture.

The changing agenda of issues since World War II has raised questions that the agricultural data system had not faced in 1972 and still has not fully faced. Despite some efforts, we have failed to face the so-called "structure" issue -- the progressive concentration of farm production and marketing firms in agriculture and its social implications. Many of our most frequently use agricultural statistics assume one large U.S. farm and present measures of inputs, production and prices at the farm gate. But what does that "farm gate" or even the "farm" mean any longer (Carlin and Hardy)? Vertical integration, contracting, artificial subdivision of farms to avoid payment limitations and many other practices have undermined the clarity of these long-used concepts. I do not perceive any action to change the conceptual base or the way the data are presented that represent progress on this front. We have as well few measures of farm size, its distribution or other structural descriptors that are comparable over time as well as useful.

The increasing interdependence of the food and fiber sector with the rest of the

national and international economies means that linkages to other sectors and functions of the economy have become more important in agricultural decision making, public and private. This makes data system demands more interactive and less predictable. We need a more flexible, real time data base and analytic capability to contend with todays problems. Great change has occurred since 1972 in the importance of international trade in U.S. agricultural products. We still have considerable difficulty in achieving a conceptual mesh with what can be measured well in trade statistics for agriculture. Many data that would be routinely useful are available only in one-time research efforts.

One of the obstacles to a more fully developed data system is the economist's failure as yet to develop a clear statement of the economics of information as a good to be produced, bought and sold. Public good aspects are especially poorly understood. One could go on to many other problem areas such as rural development and nonfarm data needs, but I do not want to depress anyone unduely. I will not even discuss the eternal, unanswered question of what the concepts of rural and urban should mean.

Where does this leave us? About where we were 15 years ago. Better off in some cases, but worse in others. We have made progress in many areas, but the changing nature of agriculture and its policy environment have flowed on, offsetting some of our good works and in any case continuing the erosion of the capability of our data system. As the Red Queen said to Alice "Now, here, you see, it takes all the running you can do, to keep in place" (Carroll).

REFERENCES

American Agricultural Economics Association, Committee on Economic Statistics. "Our Obsolete Data Systems: New Directions and Opportunities." American Journal of Agricultural Economics. 54(1972):867-75.

Baum, Kenneth and James D. Johnson. "Microeconomic Indicators of the Farm Sector and Policy Implications." American Journal of Agricultural Economics. 68(1986) 1121-29.

Betts, Richard K. "Analysis, War, and Decision: Why Intelligence Failures are Inevitable." World Politics. 31(1978)61-84.

Bonnen, James T. "Assessment of the Current Agricultural Data Base: An Information System Approach." A Survey of Agricultural Economics Literature. Vol. 2. (eds. Lee R. Martin and George G. Judge et. al.) Minneapolis, Minnesota: University of Minnesota Press, 1977, pp. 386-407.

Bonnen, James T. "Improving the Data Base" Agricultural and Rural Areas Approaching the 21st Century: Challenges for Agricultural Economics. (eds., James Hildreth, Katherine Lipton, Ken Clayton and Carl O'Connor). Ames: Iowa State University Press (Forthcoming).

Bonnen, James T. "Improving Information on Agriculture and Rural Life." American Journal of Agricultural Economics. 57(1975):753-63.

Carroll, Lewis. The Annotated Alice. New York: Clarkson N. Potter. 1960, pp. 210.

Carlin, Thomas A. and Charles R. Handy. "Concepts of the Agricultural Economy and Economic Accounting." American Journal of Agricultural Economics. 56(1974)964-75.

Mayer, Leo V., and J. Dawson Ahalt. "Public Policy Demands and Statistical Measures of Agriculture." American Journal of Agricultural Economics. 56(1974):984-88.

Weeks, Eldon E., Gerald E. Schulter, and Leland W. Southard. "Monitoring the Agricultural Economy: Strains on the Data System" American Journal of Agricultural Economics. 56(1974):976-83.