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.*

Agriculture-
Statistics
C

1981

UNIVERSITY OF CALIFORNIA
DAVIS
APR 20 1982
Agricultural Economics Library

ECONOMIC STATISTICS FOR AGRICULTURE: CURRENT DIRECTIONS, CHANGES AND CONCERNS

Proceedings of a symposium at the 1981 annual meetings of the
American Agricultural Economics Association,
Clemson, South Carolina
July 28, 1981

CONTENTS

FOREWORD

The Economic Statistics Committee of the American Agricultural Economics Association acts on behalf of the Association to improve the quantity and quality of economic statistics for agriculture. Since its first meeting in July 1970, it has sponsored several Task Forces and seminars to assess the possibilities for improving data collection in various areas. The task forces were those on: Food and Fiber Statistics (R. J. Hildreth, Chm., 1972-75); Social and Economic Statistics for Rural Areas (George Tolley and Keith Bryant, successive chairmen, 1972-76); Price-Spread Data for Foods (George Brandow, Chm., 1975-76); Measuring Agricultural Productivity (Bruce Gardner, Chm., 1977-80); and The Economic Value of Statistical Information (Bruce Bullock, Chm. 1977-80).

In its 1980 meeting, the Economic Statistics Committee felt that it was time to review progress made in agricultural statistics during the decade, and prospects for changes in the future. On behalf of the committee, Edward Reinsel and Richard Perrin organized a symposium for the 1981 annual meetings, "Economic Statistics for Agriculture: Current Directions, Changes and Concerns", at which the papers reproduced here were presented.

Richard Perrin
Edward Reinsel

FEDERAL AGRICULTURAL STATISTICS--
STATUS, IMPROVEMENTS AND THOSE ON THE HORIZON

Introduction

The Statistical Reporting Service (SRS) and the Bureau of the Census have many parallel requirements such as sampling frames, data needs, survey capabilities, statistical techniques and staff resources for conducting their respective agricultural statistics programs. Many are often cited incorrectly as total duplications of effort. The SRS program centers around many sample surveys for currently reporting the potential or realized production, supply, price and other characteristics of the agricultural economy. It issues about 475 reports annually giving State and national estimates of crop acreages, yields and production, livestock inventories and production, stocks and prices of agricultural commodities, value and utilization of farm products and prices paid for production inputs. The entire statistical program centers around small scientifically designed sample surveys that require precise and timely data collection for rigid reference dates that typically precede the announced release data by about two weeks.

The Bureau of the Census conducts the Census of Agriculture, typically at 5-year intervals. This census provides the only periodic comprehensive data about the nation's agriculture at the county level, that covers the entire country. The central focus of the Census of Agriculture is to provide a good historical picture of the changes in American agriculture each five years. These data provide information on farms that can be used for classification of operations based on major types of activity, size, legal form of organization, age of operator, etc. The Census attempts to achieve as complete coverage as possible of all operating units. Most of the data are collected by mail to reduce data collection costs; however, this extends the survey over a six-month period.

The principal overlap in activities is (1) the identical requirement for a general purpose list of farmers' names and addresses that is current and reasonably complete, and (2) the acquiring of data on crop acreage and production, livestock inventories and data on land in farms every fifth year. Both agencies have sought in recent years to minimize overlap to the extent possible within existing legislation and support legislative efforts currently underway that would enhance their ability to expand cooperation.

THE STATISTICAL REPORTING SERVICE SYSTEM

By William E. Kibler 1/

Basic Sampling Methodologies

During the past two decades the methodology for developing current agricultural statistics on crops, livestock and prices has shifted almost entirely from non-probability mail surveys to probability surveys using area and list sampling frames. With probability surveys, estimates for characteristics of interest can be generated without dependence on prior survey relationships or benchmark data such as the 5-year Census of Agriculture. Such surveys also provide the data necessary to derive sampling errors for evaluating the reliability of estimates generated and for optimizing sample designs and allocation of sampling units.

A basic requirement for any probability survey is a complete sampling frame which is an aggregation of the elements from which a sample can be selected. An area frame is the principal frame used for estimating major crop acreages, yields and production. This frame is made up of small geographic units of land called "segments" which may be sampled. It is constructed using the most current aerial photography available to classify (stratify) all land according to its current use./1/ The stratification is based on extent and type of farming and can be described in four broad categories: (1) intensively cultivated areas where a significant portion of the land is under cultivation, (2) extensive agricultural areas used primarily for grazing and producing livestock, (3) highly developed land found in city residential, shipping and industrial areas, and (4) non-agricultural land such as parks, military reservations and other recreational areas./2/

As frames for individual States are periodically updated, by using additional materials such as satellite imagery, more sophisticated stratification procedures have been used. Examples include the addition of an agri-urban stratum which is used as a transition zone between the city and agricultural strata. Within the intensive agricultural stratum refinements have been made by including additional information such as soil type and topography to develop crop-specific strata. For example, a fruit/vegetable stratum in California, a dry land wheat stratum in Oregon and Washington, and rice, peanuts, wheat/sorghum, and cotton strata in Texas. Geographic stratification is sometimes used, in addition to the land use stratification, to separate differing agricultural areas. This is accomplished by grouping counties into type-of-farming districts.

About two decades ago research showed that an optimum size segment should include about two farms and be about one square mile in intensively cultivated areas, several square miles in extensively farmed areas and one-tenth square mile in industrial or urban areas. As additional refinements have been made in both the area sampling frame and sampling methodology over the past decade, segment size has generally been reduced to an average of about .7 square mile for the intensively cultivated strata. In many states .5 square mile segments are used. This, combined with increasingly sophisticated sample designs, has permitted

1/ Acting Administrator and Chairman of the Crop Reporting Board, Statistical Reporting Service, USDA, Washington, D.C.

significant reductions in data collection costs (up to 30 percent in some States) without adverse impact on the error level of the estimates from the June Enumerative Survey. This has been of primary benefit in helping to cover inflation costs in the absence of increased appropriations.

Two sampling methods are followed in selecting sample segments. A systematic-sample approach is used in some States with the frame units arrayed geographically to ensure proper dispersion over the area of interest. Since 1974 all new samples have used interpenetrating designs to provide flexibility in computing sampling variation and segment rotation.

Crop Sampling Methodology

A sample of 15,700 segments (about 350 per State) representing about 0.4 percent of the universe is selected and enumerated about June 1 to collect data on crops planted and livestock numbers as well as characteristics of farms. Both direct expansion and ratio estimators are used for aggregating sample data to State, regional and national totals. Survey data from each segment are expanded by the reciprocal of the probability of selection (typically a factor of about 250) to obtain the direct expansion estimate. A ratio estimate is also computed using current and previous years' data since about 80 percent of the segments are enumerated in successive years. This estimate is particularly useful in evaluating changes from year to year for identical segments. Sampling errors for acreage planted to major crops are about 2 percent at the national level, 3 to 4 percent at the regional level and 4 to 6 percent at the State level.

Sampling Errors from the 1981 June Enumerative Survey
for Planted Acreages at State, Regional and National Levels

Commodity	National	East North Central Region	Illinois	Mississippi
	- - - P e r c e n t - - -			
Corn	1.2	1.2	2.5	18.7
Soybeans	1.3	1.6	3.0	6.0
W. Wheat	1.6	3.2	6.6	15.6
Cotton	3.3	-	-	10.8
Sorghum	4.3	-	32.9	24.4

Followup objective yield surveys are made for wheat, corn, soybeans and cotton to update planted acres for fields actually planted after June 1, to collect information for forecasting yields during the growing season, and to estimate actual yields at harvest. These surveys provide information based directly on counts, measurements and weights of the crop made from small randomly selected plots in sample fields. Samples are designed to produce estimates of at harvest yield with sampling errors of 1 to 2 percent. Large nonprobability mail surveys are conducted to gather data for strengthening State and sub-State estimates for crops important to the State's agricultural sector and to support cooperative State-Federal programs. Samples for such surveys vary in size from 150,000 to 200,000 and operate

fairly effectively for disaggregating accurate annual benchmarks based on probability surveys.

Livestock Multiple Frame Sampling

Multiple frame sampling utilized more than one sampling frame to cover the universe of interest. The theory for multiple frame sampling was first developed in the 1960's. /3/ Its use has rapidly grown because of its distinct advantages in efficiency of costs in data collection and its ease of adaptation for specialized characteristics associated with a small portion of firms in a universe. The theoretical concepts of multiple frame sampling are basically the same as those for probability sampling concerning known probabilities and randomness of selection. In addition, two other characteristics must hold: (1) every element of the population must belong to at least one of the frames, and (2) it must be possible to specifically identify the frame(s) to which, if any, each selected sample unit belongs other than the one from which it was drawn. The use of the area frame as described earlier satisfies the first characteristic. The second characteristic requires the proper classification of each farm operator as to whether his name is included on the list(s) frame(s). Multiple frame sampling technology is used for rice, potatoes, quarterly hog surveys in 14 States, and semi-annual cattle surveys in 28 States.

With multiple frame sampling, data can be collected more efficiently by mail or telephone and more efficient sampling can be accomplished by stratification of the list by size of operation. A variety of list sources such as ASCS, State Farm Censuses, brand inspections, etc., is used in assembling list frames. However, due to rapid organizational and operational changes that occur, lists must be updated periodically to retain their advantages in sampling and cost efficiencies. There are also some complex operating problems associated with identifying and measuring overlap between the two frames (area and list) that increase nonsampling errors. Typical sampling errors for these multiple frame surveys for cattle and hogs are shown in the following table:

Sampling Errors for 1980 and 1981 Based on
Multiple Frame Surveys for Hogs and Cattle
at Various Geographic Levels

Survey	23 State Level	14 State Level	28 State Level	Iowa	Georgia
	- - - P e r c e n t - - -				
December 1, 1980 Hogs	2.1	2.3	--	3.5	11.9
June 1, 1981 Hogs	1.8	2.0	--	4.1	9.2
January 1, 1981 Cattle	--	--	1.3	3.6	4.9
July 1, 1980 Cattle	--	--	.8	3.5	6.5

New Probability Surveys for Prices

The area and list sampling frames described earlier are not suited for collecting current information on prices farmers receive for commodities they sell or prices

paid for inputs used in production. An indirect method is used to establish these frames and select appropriate samples. For obtaining prices received for grains, a list of all grain and oilseed elevators (about 14,000) is maintained from administrative records, available as a byproduct of licensing requirements. These are stratified by storage capacity and a probability sample of about 1 in 6 selected for surveying. Similar lists for cotton, peanut, and rice buyers serve as frames for these crops.

For commodities such as cattle, hogs, vegetables and fruits, a periodic point of sale survey is conducted to determine what portion of the total production is sold through each marketing channel such as auctions, dealers, commission firms, processors and packers. The universe list of these firms is then stratified by type of marketing channel and sampled using probabilities proportional to the channel's importance in the marketing of the commodity.

Firms sampled are surveyed monthly on about the 15th to collect actual quantities purchased and dollars paid farmers for each commodity during the previous month. These data are used to derive a self-weighted average published as the revised price received for the entire month. In addition, the exact price being paid to farmers about mid-month is obtained and published as the preliminary price as of the 15th of each month. The data on quantities purchased are used by the Economic Research Service (ERS) in estimating current cash receipts for aggregation and calculation of farm income. The typical entire monthly price received for corn has a sampling error of about 3 cents while the error for the aggregated 5-month average used for determining the level of deficiency payments is less than 1 cent.

In collecting data on prices paid for inputs used in commodity production, a periodic point of purchase survey is conducted to ascertain the portion of the various inputs that are bought through cooperatives, brokers or wholesalers, dealers or manufacturers. Lists of firms are assembled from phone directories, licensing bureaus, and the American Business Lists Inc., and classified by specific inputs sold. The listed firms are geographically grouped by counties to form a frame of primary sampling units. For primary sampling units selected to be surveyed, a second stage of sampling is performed to identify the individual firms to be included in the sample. The clustering by counties makes data collection more efficient by reducing travel. Much of the work requires personal interviews for establishing accurate specifications on inputs priced.

Reliability and Completeness of Principal Statistical Series

Many data users have requested that the Crop Reporting Board provide additional information on the sources of data used in establishing official estimates and measures of their reliability since social or economic costs of errors in forecasts can be significant./4/ Beginning in 1977, most major reports have included a general summary of survey procedures, comments about errors from sampling and non-sampling sources and typical sampling errors for surveys or Root Mean Square Errors for forecasts. The following is typical of the summaries provided for livestock reports:

> RELIABILITY AND ESTIMATING PROCEDURES: Primary data used in setting these hog estimates were obtained from a sample of farmers across the U.S. using probability surveys. Information was collected by mail,

telephone and personal interviews. Since all operations raising hogs were not included in the sample, survey estimates are subject to sampling variability. This variability, as measured by the relative standard error, is about two percent at the U.S. level for hog inventory. This means that chances are approximately 95 out of 100 that survey estimates will be within four percent of the complete coverage value if the same procedures were used to survey all producers. Survey estimates are also subject to non-sampling errors such as omissions, duplications, and mistakes in reporting, recording and processing the data. These errors cannot be measured directly, but they are minimized through rigid quality controls in the data collection process and a careful review of all reported data for consistency and reasonableness.

The sampling variability of survey estimates on intended farrowings is slightly larger than that for inventories. More important, actual farrowings may differ significantly from reported intentions due to unexpected economic and environmental conditions. These differences have exceeded four percent for about one-third of the quarterly pig crops during the last seven years.

In setting the inventory estimates, the Crop Reporting Board constructed a U.S. balance sheet using estimates on births, deaths and check data on slaughter, imports and exports. This balance sheet provided an additional check on survey inventory estimates./5/

Some users have commented that these have been useful in analyzing data but the numerical sampling or forecast errors have generally not been used extensively in modeling. In fact, some data users have completely ignored the cautions about intended farrowings and assumed that they will always represent what will occur during the next 6-month period. The table on page 7 illustrates the preliminary estimates for the inventory of all hogs and pigs based on sample survey data and the final estimate that was established after reevaluating all data when slaughter records became available six months later.

If we obtained perfect data collection the sampling errors would indicate that about 2 out of 3 of the estimates would require revisions of less than 2 percent and 19 out of 20 would require revisions of less than 4 percent. For the 20 estimates during this period, 18 required revisions of less than 2 percent and 19 required revisions of less than 4 percent. Hence, the sampling errors are reliable measures of the accuracy of the estimates. The same sample of producers is used to obtain data on farrowing intentions. Hence, the same statistical analysis can be applied to farrowing intentions using the table on page 8 .

These intentions forecasts also have a sampling error of about 2 percent. Note the very large deviations for some 6-month periods cannot be explained by statistical measures such as sampling errors. Thus, these deviations must be associated with either problems in acquiring accurate data on intentions from producers or changes in plans made by producers due to such things as weather, feed cost, market prices, or as a direct result of the published intentions report. Hence, analysts should use these data with much more caution than the inventory data.

Comparison of Preliminary and Final Estimates of U.S. Inventory of all Hogs and Pigs, 1971-1980

Year and Survey	All Hogs and Pigs Inventory		
	Preliminary Estimate Based on Survey	Final Estimate Based on Slaughter	Change
	1,000 Head		Percent
1971:			
June 1	66,070	65,718	-0.5
December 1	62,972	62,412	-0.9
1972:			
June 1	61,556	60,626	-1.5
December 1	61,502	59,017	-4.0
1973:			
June 1	60,271	59,571	-1.2
December 1	61,022	60,614	-0.7
1974:			
June 1	59,437	58,878	-0.9
December 1	55,062	54,693	-0.7
1975:			
June 1	48,165	47,860	-0.6
December 1	49,602	49,267	-0.7
1976:			
June 1	52,643	53,930	+2.4
December 1	55,085	54,934	-0.3
1977:			
June 1	54,100	54,460	+0.7
December 1	57,587	56,539	-1.8
1978:			
June 1	54,930	55,240	+0.6
December 1	59,860	60,356	+0.8
1979:			
June 1	64,890	65,020	+0.2
December 1	66,950	67,353	+0.6
1980:			
June 1	65,930	65,255	-1.0
December 1	64,520	1/	

1/ Subject to future minor revisions.

Comparison of Farrowing Intentions Forecast and Actual Farrowings That Occurred, U. S. Total, 1971-1980

Year and Period	Intentions Forecast	Actual Estimate	Change
	---- 1,000 Head ----		--Percent--
1971:			
December-May 1/	7,222	7,237	+0.21
June-November	6,265	6,339	+1.18
1972:			
December-May 1/	6,544	6,498	-0.70
June-November	6,005	5,973	-0.53
1973:			
December-May 1/	6,980	6,438	-7.77
June-November	5,979	5,869	-1.84
1974:			
December-May 1/	6,491	6,315	-2.72
June-November	5,760	5,476	-4.94
1975:			
December-May 1/	5,385	4,973	-7.65
June-November	4,730	4,952	+4.02
1976:			
December-May 1/	5,353	5,777	+7.92
June-November	5,811	5,850	+0.67
1977:			
December-May 1/	6,109	6,050	-0.97
June-November	6,144	6,009	-2.20
1978:			
December-May 1/	6,620	6,034	-8.86
June-November	6,247	6,398	+2.42
1979:			
December-May 1/	6,903	7,179	+4.00
June-November	7,419	7,306	-1.53
1980:			
December-May 1/	7,176	7,231	+0.77
June-November	6,716		
1981:			
December-May 1/	6,780 2/		

1/ December previous year.
2/ Latest estimates - subject to future revision.

The SRS has concluded four years' experience in using the Root Mean Square Error (RMSE)/6/ statistic as an indication of the reliability of crop production forecasts made during the growing season. The Root Mean Square Error is calculated on the basis of past forecasting performance. It is derived by averaging the squared deviations between monthly forecasts and the final estimate over a given period. The square root of these averages is the RMSE./7/ The assumptions necessary to make this statistical measure valid are (a) a normally distributed series of forecasts compared to the final estimates, and (b) factors affecting the current year's crop after the forecast date are not greatly different from those influencing crop forecasts during the historic reference period. For crops, 20 years of data are used and a t-value of 1.725 is used to compute the 90 percent interval compared to the normal distribution value of 1.645. Its performance has exceeded expectations as shown in the table on page 10 .

These data show that about 69 percent of the 189 forecasts made during the period have been within the expected 67 percent confidence interval and nearly 98 percent of the forecasts have been within the 90 percent confidence interval. Since the RMSE uses actual performance over a 20-year period to measure reliability, any improvements made in survey systems during recent years should make the measure a bit conservative when evaluating current year forecasts. An attractive concept of the RMSE statistic is that it can be derived well ahead of its expected use. The following table gives the RMSEs that have been or will be used in 1981 crop reports:

Root Mean Square Errors (Percent)
For 1981 Crop Production Forecast

(Derived from 1961-80 Forecasts as published by the Crop Reporting Board)

Crop	May	June	July	August	Sept.	Oct.	Nov.	Dec.
	Percent							
W. Wheat	6.7	5.9	3.4	1.3	--	--	--	8.0
O. Spring Wheat	--	--	10.2	6.7	3.2	3.1	--	--
D. Wheat	--	--	14.5	8.1	5.2	5.3	--	--
All Wheat	--	--	3.4	1.9	1.3	1.3	--	--
Barley	--	--	8.4	4.9	3.5	--	--	--
Oats	--	--	7.6	4.8	4.0	--	--	--
Corn	--	--	9.1	6.8	5.1	3.9	2.6	--
Sorghum	--	--	--	6.4	4.7	4.7	4.4	--
Feed Grains	--	--	--	5.5	4.1	3.1	2.1	--
Soybeans	--	--	--	5.1	4.2	3.3	2.6	--
Rice	--	--	--	4.2	3.2	2.6	2.3	--
Cotton	--	--	--	8.7	7.2	5.4	3.8	1.6

As should be expected, these data show that forecasts improve consistently as the growing season progresses and the RMSEs begin to approach the level of actual observed sampling error for at harvest estimates.

Data on Performance of the Root Mean Square Error Statistic
as an Indication of the Reliability of Crop Production Forecasts (1977-1980)

Forecast Period 1/	Total Number of Forecasts	67% Confidence Interval		90% Confidence Interval	
		Forecasts Within	Forecasts Outside	Forecasts Within	Forecasts Outside
		1977 Crop			
Early Season	23	20	3	23	0
Mid-To-Late Season	22	19	3	22	0
Total	45	39	6	45	0
		1978 Crop			
Early Season	25	16	9	24	1
Mid-To-Late Season	23	14	9	23	0
Total	48	30	18	47	1
		1979 Crop			
Early Season	25	15	10	22	3
Mid-To-Late Season	23	15	8	23	0
Total	48	30	18	45	3
		1980 Crop			
Early Season	25	14	11	25	0
Mid-To-Late Season	23	18	5	23	0
Total	48	32	16	48	0
GRAND TOTAL	189	131	58	185	4

1/ Early Season Forecast Months: December, May and June for Winter Wheat; July and August for Durum, Other Spring and All Wheat, Corn, Barley and Oats; August and September for Soybeans, Cotton, Rice, Sorghum Grain and all Feed Grains (1978-80). Mid-to-Late Season Forecast Months: July and August for Winter Wheat; September and October for Durum, Other Spring and All Wheat; September for Barley and Oats; September, October and November for Corn; October and November for Soybeans, Grain Sorghum, Feed Grains, Rice and Cotton.

Sampling Errors for 1980 Production Based on
Enumerative and Objective Yield Probability Surveys

Crop	Percent Sampling Error
All Wheat	2.2
Corn	1.6
Soybeans	1.9
Cotton	5.0

For less sophisticated data users the following type of table appears in each monthly crop report during the growing season:

Crop Production Forecasts

Ten-Year (1971-80) Record of Differences Between
First Monthly Forecasts and Final Estimate After Harvest

Crop and Month	Units	Quantity			Number of Years	
		Average	Smallest	Largest	Below Final Estimate	Above Final Estimate
July 1981						
Corn	Million Bu.	510	2	1,276	5	5
Oats	Million Bu.	45	8	92	5	5
Barley	Million Bu.	32	0	71	6	4
All Wheat	Million Bu.	54	2	143	5	5
Durum	Million Bu.	10	4	19	4	6
Other Spring	Million Bu.	31	3	97	6	4
Winter	Million Bu.	29	1	55	6	4
August 1981						
Rice	Million CWT	3	0	7	4	6
Soybeans	Million Bu.	78	1	165	6	4
Cotton	Thousand Bales	796	149	1,690	4	6
Sorghum	Million Bu.	34	2	78	5	5
Feed Grains	Million Tons	9	1	22	7	3

These are actual data that will appear in upcoming 1981 reports and are derived from records kept by the Crop Reporting Board to assist in analyzing its forecast record. For forecasts to maintain credibility they must be both reliable and objective. Producers always assert Crop Reporting Board forecasts are always too high. A simple measure such as the number of years forecasts are below or above the final estimates helps establish whether the Board errs consistently in an optimistic or conservative manner.

One topic that received much discussion in 1980 was differences in numbers of farms published by the Department of Agriculture and the Bureau of the Census. A task force was appointed to review the data and investigate whether (1) the Department survey procedures might be missing operations that qualified as farms, or (2) the Census might be tabulating duplicate records for the same farm or including operations that did not actually qualify as farms. As the result of this study, some modifications were made in survey procedures used by both agencies and many of the major differences observed in initial reviews were eliminated.

The number of farms as published differed about 2 percent for the 48 contiguous States with SRS estimates above Census in 10 States, essentially the same as Census in 12 States and below Census in 26 States./8/ Reasons for these differences include (1) the Census procedures include any farm operated during the year while SRS procedures measure only those operating on June 1; (2) the Census survey procedure is more likely to result in classification of marginal operations as farms; (3) Census coverage of special farms such as mink, nursery, fish operations, is more complete; (4) SRS's area frame screening procedures may not identify all urban farm operators or special farms; and (5) Census has more opportunity for duplicating partnerships, or counting landlords or tenants as operators. For the major north central region, the two series are at almost the same level.

Data Needs Identified for the 1980's

Numerous agricultural data needs have been identified by data users at Department of Agriculture and Bureau of the Census data users workshops, commodity organizations, advisory committees, and national or State policy officials./9/ These include:

(a) Earlier planting, farrowing or calving intents,
(b) Current feeding and feed conversion rates,
(c) Data on marketing trends,
(d) Calving rates for beef and dairy cows,
(e) Monthly sow farrowings,
(f) More detailed data on double cropping,
(g) Expanded sub-State and county data for livestock and poultry,
(h) Forecasts and broader geographic coverage for sunflower production,
(i) Data on quality and amount of forage production,
(j) County estimates for acreage, yield and production for double cropped and other cropping practices by commodity,
(k) More detailed data on farm and migrant labor,
(l) Improved coverage of aquaculture production, and
(m) Improved data on transportation costs.

All of these are very legitimate needs but must be prioritized with all existing series to match resources and needs. In making these decisions, a number of factors are considered. These include quality (reliability), timeliness, frequency, geographic and commodity coverage, data user support and industry acceptance of responsibility for providing the basic data.

Theoretical work in several areas has shown that social benefits accrue rapidly for commodity data until sampling errors reach about 2 percent. It also considers that sampling errors are not the only source of survey errors and that considerable

effort must be exerted to keep the level of nonsampling errors, which are difficult to measure, within the 2 percent bound. However, we must not overlook the fact that this generally becomes a threshold value in survey planning. The value of accuracy for certain types of information, such as prices received by farmers for crops covered by the Food and Agriculture Act of 1977, is more easily quantified. A change of 1 cent per bushel for the aggregate 5-month price period for corn could mean a $50 to $60 million difference in deficiency payments to farmers. Statistical data on such items as production, inventories, marketing, etc., can be used to develop a balance sheet for checking the reliability of data. The Crop Reporting Board develops these balance sheets also for use in generally reviewing statistical survey data. However, it does not discount survey data in establishing production or inventory in order to force the components to balance. To do so would imply an accuracy that doesn't exist. The Board policy is to accept an imbalance or residual of up to one percent which it feels reflects statistical (or non-statistical) errors that are present in each of the components used in the balance sheet.

For surveys during the growing season that relate to a specific date, such as August 1, the release date is set about 10 to 12 days later. Although there are numerous requests for earlier release of such data, its practicality and efficiency from the standpoint of costs are questionable. First, centering data collection on the first day of each month requires that data be collected the last few days of the previous month and the first few days of the new month. This leaves, at most, 7 or 8 working days to summarize, analyze and publish the results. Reviews of month-to-month changes in production forecasts over a period of 5 years show that they average less than 2 percent. This would imply that, on the average, changes that have occurred in the 10-day period between data collection and publication are generally very small. For larger surveys such as the hogs and pigs report and cattle or grain stocks reports, which must wait a few days after the first of each month for firms to close their books, the release date is about 3 weeks after the reference date. These surveys obtain inventories as of a point in time so the data are not subject to changes for the same reasons cited for crop production forecasts.

Forecasts of crop production are done on a monthly basis during the growing season with qualitative assessments on growing conditions provided by the Weekly Weather and Crop Bulletins. The Crop Reporting Board policy requires that survey data be collected from producers for each forecast. Hence, the added cost and the inability to measure changes that might occur for shorter intervals almost precludes the issuance of more frequent reports. Since it takes several days for each forecast to be totally reflected in the market and there is oftentimes 1 to 2 days' inactivity in the market immediately preceding major reports, there is some question whether more frequent reports would be beneficial. The established frequency of other weekly, monthly, quarterly or semi-annual reports is generally associated with the time required to produce the commodity. Much of the pronounced seasonality that once existed for milk, eggs, pigs, etc., has been somewhat eliminated by specialization and significant portions of the month-to-month changes are associated with differences in length of months rather than changes in actual production levels.

During recent years the Crop Reporting Board has followed a policy of providing current estimates of forecasts at the State level for approximately 95 percent of

production for each of the major crops and combining all other States into one aggregate total. A further refinement is to also include any State having at least 1 percent of U.S. total production. The geographic coverage is reviewed every 5 years and changed to reflect geographic shifts in production that occur through time. For the minor producing States, a single production forecast is made and carried forward through the growing season until an annual survey, using both probability and nonprobability surveys, is used to establish annual levels of production after harvest. This approach is cost-effective since it requires significant resources to collect data for commodities that are produced on a relatively small portion of all farms. It also carries some risks. This occurs when conditions or changes for the 5 percent omitted are considerably different from the other 95 percent.

New Federal Crop Insurance Corporation (FCIC) programs and rapid changes in cropping patterns are both increasing the need for county level data. Much of the current county data is financed by cooperative State funding and is based on nonprobability survey techniques. The current reliability of such data, although not scientifically measured, is thought to be 10 to 15 percent for major producing counties.

Cooperation in providing the basic data for the industry being surveyed is of utmost importance in maintaining the quality of data. Nonresponse rates of more than 10 percent are considered critical especially if they are confined to a particular segment of the universe. Special efforts are made to work with industry representatives to improve cooperation when nonresponse rates increase to this critical level.

Improvements Accomplished or Being Implemented

Area frame modifications made over the past decade that follow current land uses and provide for the control of segment size, have reduced sampling errors for major crops by about one-third even after sample sizes have been reduced about 15 percent. New work in this activity will focus on keeping these frames up-to-date in areas like the Mississippi Delta where significant land clearing continues and the Great Plains where rapid expansion of irrigation has introduced significant new cropland. Landsat imagery has been introduced into current area frame update procedures to compensate for the lack of up-to-date aerial photographs. Landsat data and technology are being researched as a method for refining area frame construction activities. /10/ Initial research results indicate that it might be possible to develop crop-specific stratification of intensively cultivated cropland by identifying special crops for areas 8 to 10 miles in size. /11/

Currently, efforts are underway to expand the list sampling frame (used in multiple frame sampling) to include as many farm operators as possible, with supplementary information on the commodities they produce and measures of their size. This frame will be more complete for large and specialized operations. Lists of farm operators from many sources have been assembled for developing as complete a list as possible. Despite extensive efforts to (1) identify duplication, and (2) eliminate names or operations with no agricultural operations, lists still generally contain more total names than there are farms and as many as one-third of those surveyed no longer operate farms. Thus, our experience in list development

is no different from that of the Census. It is very difficult to maintain a list frame that currently includes a large portion of the very small farm operators; however, the list frame with supplementary data is of utmost importance in developing more efficient data collection techniques and in improving the quality of statistical series at sub-State or county levels.

Research continues for making more extensive use of satellite data in improving crop acreage estimates at the sub-State and county levels. /12/ Relative efficiencies in terms of variability were about 2.5 for corn and soybeans compared with regular survey methodology. Relative efficiencies have to reach 4.0 or better to begin achieving the desired results. Although some county estimates had sampling errors as low as 7 to 9 percent, errors for a large number of counties still exceeded 20 percent. Recent research results on new estimators for using these same data show county estimates for major crops with mean square errors ranging from 5 to 18 percent with most being in the 5 to 10 percent range. /13/

Two very important data series that have not been shifted to a probability basis are the "on-farm" grain stocks survey and the fall acreage and production surveys designed to collect acreage utilization and yields for a large number of field crops. These will get careful consideration for future budget submissions.

* * *

LITERATURE CITED

/1/ Kibler, William E., "New Area Sampling Frame Being Developed," Statistical Reporter, Office of Management and Budget, Washington, D.C., August 1971.

/2/ U.S. Department of Agriculture, Statistical Reporting Service, Scope and Methods of the Statistical Reporting Service, Misc. Publ., 1308, July 1975.

/3/ Hartley, H. O., "Multiple Frame Surveys," 1962 Proceedings of the American Statistical Association, The Social Statistics Section.

/4/ Bullock, J. Bruce, "Social Costs Caused by Errors in Agricultural Production Forecasts," American Journal of Agricultural Economics, Vol. 58, No. 1, Feb. 1976.

/5/ U. S. Department of Agriculture, Economics and Statistics Service, Crop Reporting Board, Hogs and Pigs, MtAn 4 (6-81), June 22, 1981.

/6/ Kish, Leslie, Survey Techniques, John Wiley and Sons, pp. 566-571.

/7/ U. S. Department of Agriculture, Statistical Reporting Service, Crop Reporting Board, Crop Production, June 1977.

/8/ For detailed inventory refer to such documents as:

Proceedings of Workshop on Agricultural and Rural Data, Series A and B, ERS and SRS, USDA, Washington, D.C., May 1977.

"Changing Rural Development Data Needs", James T. Bonnen and Glenn L. Nelson, Proceedings of ASSA Meeting, Denver, Colorado, Sept. 1980.

Report of Economic Information and Statistics Data Users Workshops, Economics and Statistics Service, USDA, Washington, D.C., March 1981.

"Future Data Needs for Agricultural Decisions," Claud L. Scroges, Feed Stuffs, May 1981.

/9/ Vogel, Fred, "Comparison of Farm Numbers - Bureau of Census and USDA," USDA, Economics and Statistics Service, Estimates Division, Feb. 1981.

/10/ Wigton, William H., "Use of Landsat Technology by Statistical Reporting Service," Purdue LARS Symposium Proceedings, West Lafayette, Indiana, July 1976.

/11/ Hanuschak, George, and Morrissey, Kathleen, "Pilot Study of the Potential Contributions of Landsat Data in the Construction of Area Frames," Research Division, Statistical Reporting Service, Oct. 1977.

/12/ Hanuschak, George, and others. "Obtaining Timely Crop Area Estimates Using Ground-Gathered and Landsat Data," ESCS, USDA, Technical Bulletin 1609, August 1979.

/13/ Battese, George E., and Fuller, Wayne A., "Prediction of County Crop Areas Using Survey and Satellite Data," Statistical Laboratory, Iowa State University, Research Report, April 1981.

CENSUS OF AGRICULTURE

Shirley Kallek*

Present Methodology

The census of agriculture was taken every 10 years from 1840 to 1920 and each 5 years from 1925 to 1974. Presently, the census is undergoing a change in reference year to coincide with the 1982 economic censuses. Two 4-year censuses of agriculture, taken for 1978 and 1982, will accomplish the adjustment. After 1982, the census will revert to a 5-year cycle.

Starting with the 1969 Census of Agriculture, censuses have been conducted by mail. Prior to 1969, censuses were conducted by personal interview. The present procedure includes an initial mailing followed by a series of 6 followup letters and/or forms along with a telephone followup of all large farms. A sample of the final nonrespondents is then taken and adjustments are made in census totals at the county and state levels to include all farms on the mail list.

As with any large-scale mailing list, there is undercoverage, especially of the smaller farms. In order to provide more complete statistics at the state level for the 1978 census, the mail census totals were supplemented with estimates for farms not included on the mailing list based on a direct enumeration area sample of 6,393 area segments which was large enough to provide reliable state estimates for rural areas.

As part of most of the recent periodic censuses of agriculture, a series of sample follow-on surveys is taken for the year following the census. These surveys provide state- and national-level estimates and contain data on high priority subjects that were not included on the census questionnaire. Samples for those surveys are selected from the census records from the prior year which allow for the selection of efficient samples, each of which can be stratified to fit the particular subject surveyed. In addition to the sampling efficiency, data from the census record for the individual farm can be related to the survey data, thus expanding the total data available from the survey. The 1978 census follow-on surveys were conducted on farm financial characteristics, irrigation, and energy consumption.

---

* Associate Director for Economic Fields, Bureau of the Census, Department of Commerce.

Reliability and Completeness of Census Results

The Census Bureau since 1945 has included an evaluation program as part of each census of agriculture. Results of these evaluations of coverage along with estimates of sample error have been published along with the census results. In the past, few data users made use of reliability and coverage information, or did only in a casual way. However, as users become more and more sophisticated in their use of data, their use of such information increases.

In addition to reliability estimates, the Bureau publishes measures of the effect of the major statistical adjustments made to census totals during the processing.

Statistical Adjustments - For all mail surveys or censuses, there are nonrespondents as well as some refusals for which some statistical adjustments must be made. For the 1978 census, a canvass of a sample of nonrespondents was made for each state. Based on results from this sample, an estimate of a number of nonrespondent farms was made county-by-county and a sample of respondent records was then selected and weighted to represent the nonrespondent farms. Table A shows the effect of the adjustment for 1978.

Table A. Percent of United States Totals Represented by Adjustment for Whole Farm Nonresponse: 1978

Item	Percent of Total
Farms	8.5
Land in farms	4.6
Value of farm products sold	3.9
Harvested cropland	6.5
Corn harvested (acres)	6.3
Cattle and calves (inventory)	5.1

The inclusion of the direct enumeration sample estimates for farms not included in the mail list in state and national totals for 1978 was a major change in census procedure and resulted in a sizable adjustment to census mail totals, especially in regard to counts of farms. Essentially, the inclusion of the area sample estimates significantly reduced the net undercoverage rate from 10.7 percent in 1974 to about 3.4 percent in 1978. As most farms that are missed from the census mailing list are small, the area sample had a sizable effect on farm counts with much less effect on acreages, values, and livestock numbers. The impact of this adjustment is shown in Table B.

Table B. Contribution of Area Sample to Selected United States Totals: 1978

Item	Percent Represented by Area Sample
Farms	8.9
Land in farms	1.4
Value of farm products sold	1.0
Harvested cropland	1.1
Corn harvested (acres)	1.0
Cattle and calves (inventory)	1.8

Underenumeration--While the inclusion of the area estimates substantially improved census coverage, there was still some undercoverage of farms. This was measured by a two-part evaluation program which consisted of (1) a post-enumeration survey of the area sample to measure undercoverage in the original enumeration, and (2) an independent household survey to measure farms missed in urban areas not covered by the area sample frame and to identify farms not included or erroneously classified as nonfarms in the census.

Preliminary estimates now available indicate that 3.4 percent of the farms were missed by the 1978 census. Table C shows the preliminary results.

Table C. Net Undercoverage of United States Farm Numbers 1/

Item	Net percent undercount
Farms	3.4
with sales under $2,500	6.5
with sales $2,500 or more	2.5

1/ Preliminary estimates
More detailed data on census coverage will be available in a special report on Coverage Evaluation which will be published in the fall of 1981.

Sampling Variability--Sampling variability in census data arises from several sources. First, sampling involved in the imputation for nonrespondents relates to all data items particularly for mid-size and smaller farms. Second, estimates from the area sample contribute to variability of state and national totals. Third, there is variability for those report form items which were enumerated only from a sample of farms (approximately 26 percent). Estimates of sampling reliability are published for selected items for each state and county in the Volume 1 State and County Data series. Table D shows estimates of relatively standard errors for 1978.

Table D. Estimates for Sampling Reliability for Selected Items: 1978

Item	Standard Error (Percent)
Farms	.2
Land in farms	.1
Value of farm products sold	.1
Harvested cropland 1/	
Value of land and buildings	.1

1/ Collected from sample of farms (26 percent)

Nonsampling error--In addition to sampling variability, errors arise from incorrect or incomplete reporting on the part of the respondent, processing procedures, and the inability to obtain reports from all eligible reporting units. Census processing procedures are designed to minimize these types of errors to the degree possible within limits of resources and time. Quality control, verification, and check measures for specific operations are used to keep processing errors within acceptable limits.

Detailed computer checks are made to detect and correct inconsistencies in reporting. Missing data, along with obviously inconsistent data, are imputed or corrected during the computer processing based on data relationships for similar farms in the county. Large discrepancies or gross incompleteness on large farms are reviewed by the technical staff and often telephone calls were placed to the respondents to resolve problems. Likewise, all identifiable large nonrespondent cases were called to obtain reports.

While these procedures reduce nonsampling errors, they do not eliminate all errors. Extensive statistical measures will be published in the 1978 Coverage Evaluation Report on census coverage of farms, but no measures are available on other types of errors.

Improvements Accomplished or Being Implemented

Substantial changes were made in the 1978 census procedures in an effort to improve the completeness, quality, and usability of the data over those of the 1974 census. First, response burden was recognized as a major factor affecting the quality of data in any collection effort. Thus, the questionnaire content was significantly reduced from the 1974 level and the use of sampling of data items was reintroduced in the census. These changes reduced response burden by 30 to 40 percent over that for the 1974 census. This reduction had a most favorable impact on respondents with an obvious improvement in the quality and completeness of the data reported for 1978.

Improvements were also made in the quality of the mail list used for 1978. This included the incorporation of some rough lists from SRS's list frame for 26 states along with improvements to the list handling procedures. These changes also had an impact on the completeness of data for 1978. As an additional means of improving the quality of the mailing list, a Farm Identification Survey was taken early in 1978, the main purpose of which was to identify active farm operators from other persons on the mailing list who were only land owners or were no longer involved in agriculture. Those who were identified as not being active farm operators were removed from the final census mailing list.

The major improvement for 1978 was the implementation of the direct enumeration area sample which supplemented the mail census by providing state estimates for farms not included on the census lists. The inclusion of these estimates in census state totals significantly improved the reliability and usability of the data.

It should be noted that because of the major improvements made to the census, data for 1978 are not directly comparable to totals for 1974 or 1969. Some lack of comparability will result from any improvements made in a statistical series.

The primary efforts for the 1982 Census of Agriculture are aimed at providing more comparability in coverage while minimizing the reporting burden on the farmer. Additions or changes to the data content of the census form have been severely limited so as not to increase respondent burden. To minimize the appearance of burden, the report forms will be tailored to reflect crops grown in various regions of the country. The 12 versions of the report form will have crops ordinarily grown in the region prelisted on the report form to ensure more complete reporting of crop data. Sampling of selected data items within the report form will be continued. The direct enumeration of an area sample will also be used as in 1978 to provide state-level estimates for farms not on the mail list.

Sources of the mail list will be similar in general to the 1978 census, although the unduplication and list handling process will be refined to remove a higher proportion of nonoperators. The availability of improved computer facilities and historical records should produce an improved quality list. The major emphasis will be on improvement in publication schedules with the preliminary data series completed within 14 months and the Volume 1 series in 20 months of the start of the data collection. Timeliness was the major complaint from users of census data.

Emerging Data Needs for the 1980's

As the structure of agriculture becomes more complex and more interrelated to other sectors of the economy, data needs from census programs are increasing beyond just the basic farm counts, acreage, and livestock number counts historically provided. The availability of computer data bases is allowing data users to absorb larger and larger amounts of data. For example, there has been a marked increase in the use of county-level data in the private sector for market analysis along with an increase in demand for small area data. In addition, increased concerns on land use at local levels and various kinds of rural development programs are requiring greater and greater detail.

Recent requests for data and comments gathered from the Bureau's advisory committee, data user conferences, other government agencies, and private groups center on the following:

1. Need for data on a more timely basis.
2. More need for comparability within and between data series.
3. Increasing need for size distribution and farm classifications to analyze particular segments of agriculture.
4. Need for more detailed data to be available on computer tapes, microfiche, etc.
5. Increased interest in longitudinal studies.
6. Need for improved linkage of production sector data with the remainder of the food and fiber industry.
7. Need to reevaluate many of the measurement concepts used in data collection.
8. Need to provide measures of data quality (accuracy and completeness).
9. Specific areas of expanded current need
   a. land ownership and control
   b. irrigation as related to water use
   c. farm labor input (paid and unpaid), wage rates, etc.
   d. more detail on value of commodity sales (necessitated by changes from historical market channels)
   e. continuing interest in financial information, off-farm income, etc.

There is a growing realization that farms are increasingly different, they have different resource characteristics, needs, and goals. Part-time farms differ from full-time farms as do dairy farms from cash grain farms. Even the same type of farm differs greatly geographically. These increasing differences are requiring greater detail of data and increased use of farm classifications.

The availability of historical census records at the Bureau now provides an opportunity for longitudinal studies. The Economic Research Service of the USDA is presently working with the Bureau in proposing some special tabulations looking at matched census records from the 1974 and 1978 censuses. If funds are available, a series of tabulations will be made and a plan designed to allow for linkages being made to the 1982 census.

With the 1982 census being conducted concurrently with the economic census for the first time, the opportunity exists to improve the linkage of data for the total food and fiber industry. Louis Upchurch and James Bonnen have been articulate spokesmen in expressing the need for these improvements. The Economic Statistics Committee of AAEA has established a task force on Joint Data Systems for Agriculture and Agribusiness which will work with the Bureau in developing and improving the linkage between these statistical series. Hopefully, this group will provide the focus and priorities to the Bureau which then can improve the description of the input and output flow within the food and fiber sector, primarily through special tabulations of the two data files. However, this is a long-term goal and will have to be approached a step at a time.

Many of the definitions and concepts used for data collection in census programs are becoming increasingly out-of-date as the structure of agriculture changes. For example, continued increases in vertical integration, forward pricing, pool-type marketing, and direct marketing all are making the "farm gate" concept increasingly difficult to apply. Other concepts are becoming equally difficult to apply in practice.

More and more users are raising questions on the completeness and accuracy of individual data items. While the Bureau has been a leader in promoting evaluation studies, most studies have concentrated on the coverage of farms, not the quality of reported data. While such studies have been discussed in recent years, lack of funds have prevented the implementation of any such evaluation survey.

The major limitations to meeting many of the new data needs are the respondent burden and availability of detailed records on the part of respondents. Regardless of the data need, meaningful and useful data cannot be collected from the census or related surveys in all cases. Some kinds of detailed data will have to continue to be collected as part of research projects and other small-scale surveys. Data from these detailed surveys often then can be extrapolated to totals from other data series.

REFERENCES

Lin, William, George Coffman, and J. B. Penn, "U.S. Farm Numbers, Sizes, and Related Structural Dimensions: Projections to Year 2000," Washington, D.C., USDA-ESCS Technical Bulletin No. 1625, July 1980.

Nicol, Kenneth J., "Farm Sector Data: Presentation and Improvement," American Journal of Agricultural Economics, Vol. 63 No. 2, May 1981.

Penn, J. B., "The Changing Farm Sector and Future Public Policy: An Economic Perspective," Agricultural-Food Policy Review, Washington, D.C., USDA-ESS AFPR-4, April 1981.

Upchurch, M. L., "Steps Toward Better Data for the Food Industry," Washington, D.C., U.S. Department of of Commerce, Bureau of the Census, Economic Research Report 2, February 1979.

U.S. Department of Commerce, Bureau of the Census, Census of Agriculture, Volume 1, United States Summary and State Data: 1978.

U.S. Department of Commerce, Bureau of the Census, Census of Agriculture, Special Reports on Evaluation of Coverage, 1974.

Events and Activities Impacting on Agriculture Data--Sources of Concern and Implications

Gaylord E. Worden and John H. Berry 1/

The managers of the two major Federal units which provide most of our basic agriculture statistics gave us a glimpse of their aspirations for continued program improvements. Such planning is a continual process partly because of the long lead time needed to implement a new or changed data program. But they face, like other program managers, a decision environment that includes several important exogenous factors.

Our attempt in this paper is to briefly describe some of these exogenous factors, hopefully the more important ones, and to conclude what implications these factors have for the agricultural data system. The factors we will discuss are limitations and reductions in the Federal budget, reductions in respondent burden, and pressures to respond to special needs in and outside Government. None of these is new; however, their possible cumulative impact at this time is reason for concern.

Federal Budget Limitations and Reductions

Resource limitations are a fact of life for any manager--whether he or she is a farmer, business manager or Federal statistical program administrator. For the past few years, however, the resource climate for most Federal statistical programs, including agriculture, had been a slight upward trend--enough resource increases to cover increasing costs of the existing programs and to also allow a few program improvement initiatives.

That has changed. Some statistical agencies have taken cuts in travel, absorbed pay increases and have had to find ways to contend with inflationary cost increases for printing and other services. The

---

1/ Director and Economic Statistician, respectively, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce, Washington, D.C.

extent of budget impacts, even near term, on statistical programs will not be clear until the President's Budget for FY 1983 is completed several months from now. It is likely, however, that maintaining our present statistical programs, and certainly expanding any program, will be far more difficult than any time in the recent past.

The impact to date on agriculture statistics programs has been a general belt tightening and the beginning of some reduction in existing programs. The Census Bureau has had to eliminate some of the tabulations and publications of U.S. level statistics for selected subjects and also some special reports from the 1978 Census of Agriculture. The Statistical Reporting Service (SRS) has proposed eliminating the quarterly farm labor program in order to meet cost increases and budget reductions. One of the more significant impacts to date, however, affects demographic data needed for rural areas. Legislation was passed in 1976 which authorized a mid-decade population census for 1985. The administration decided to eliminate planning funds in 1981 because of the large projected cost of conducting such an activity, even in the form of a large sample survey.

In addition to tighter fiscal restraints on statistical programs, the reductions taken by entitlement program agencies may impact on the statistical and analytic community. These agencies have traditionaly contracted with both Federal and non-Federal units for program evaluation data and other types of analyses. As a result of that work, statistical data bases have been augmented with special surveys and adminstrative records data. As resources are reduced in these agencies, some of these administrative data sources are also placed in jeopardy.

The more severe resource constraints probably lie ahead in fiscal years 1983 and 1984. Agencies have been instructed to turn in budget requests for 1983 that will allow higher level managers to make decisions on where to cut back on programs rather than approving new initiatives. So the direction appears to be changed, and there will be much more difficult resource allocation decisions to make in this new budget environment.

Public Reporting Burden Reduction

Every President since Franklin D. Roosevelt has followed his lead with various attempts to reduce unnecessary paperwork and reporting burden. In their 1977 final report, the Commission on Federal Paperwork estimated the annual cost of Federal paperwork at $100 billion or more. [1] They concluded that some part of this was unnecessary and the hidden tax this burden placed on the economy was too often ignored by managers of Government programs.

Last December, President Carter signed into law the Paperwork Reduction Act of 1980. This law established organizations and procedures that had been recommended by the Commission on Federal Paperwork to control reporting burden.

The paperwork target of the law is much broader than the collection of statistics. All forms of information collection are covered including the public burden imposed by the Internal Revenue Service and independent regulatory agencies. This action expanded considerably the information collection approval authority of the Office of Management and Budget (OMB). Previous to the Act, OMB only had control over slightly less than one-half of the total burden.

The Act also set a goal of reducing reporting burden by 15 percent by October 1, 1982 and another 10 percent in the following year. One of the new procedures set up by OMB to help meet these objectives is the Information Collection Budget. The principle is that burden imposed on the public is a limited resource that should be budgeted, allocated and controlled by OMB, just as the other important resources--funds and personnel positions--are managed.

During the current fiscal year the Federal government is expected to impose 1,228 million hours of reporting burden on the public. Three-fourths of that burden is mandatory reporting to the Government while

only 6 percent is voluntary. The remaining 19 percent is required to obtain benefits. The mandatory portion is so high because a little over one-half of the total burden is imposed by the Internal Revenue Service.

Only about 2 percent of the total burden is caused by statistical surveys. Even the nearly 14½ million hours imposed on farmers as a respondent class is primarily for regulatory and compliance, applications for benefits, and program planning and management purposes. About 4 percent of what farmers are asked to furnish the Government in FY 1981 is for general purpose statistics.

The philosophy of burden reduction is also being used to justify some actions of others. In at least one instance, the action is counter to the position of the American Agricultural Economics Association Committee on Economic Statistics. In its 1972 report, the Committee discussed improved data concepts needed to describe the agricultural industry. [2] One achievement pointed to was the effort by the Bureau of the Census to enumerate the farm service sector of the food and fiber industry. Yet in the FY 1980 Senate appropriation bill, language was added which prohibited the use of funds for planning any survey of agricultural services. The Senator who introduced the language said that constituents had complained about burden of the Census of Agriculture Services form. Similar language was included in the FY 1981 bill, and the Farm Finance Survey was added to the prohibition.

What might this renewed emphasis and tighter controls on response burden mean for agriculture statistics? Certainly if SRS and the Census Bureau have to reduce their programs by 25 percent in order to

achieve the burden reduction goals, the impact will be very significant. Few of us will argue with the goal of keeping respondent reporting burden to a minimum. However, we need to guard against that becoming a single-minded goal. The policy of the central statistical office is that statistical surveys should not have to take an equal or porportionate share of the burden reduction since they are not the cause of the problem.

Agency Priorities and Special Interests

We have a very decentralized statistical system in the Federal government. For example, there are important price, consumption and international statistical programs for food and agriculture that are conducted by agencies other than the Census Bureau or SRS. And the many sources of data available to those of you who are users of statistics on rural population and communities or on natural resources and the environment illustrate the fragmentation even more vividly.

After studying the Federal statistical system for about a year, Jim Bonnen concluded that much stronger coordination of budget development for statistical programs is necessary to ensure that national, as opposed to separate departmental and agency, statistical priorities are identified and prevail through the planning process. [3] High among national level priorities would be increased relevance of statistical information for addressing policy issues that more and more frequently cut across departments and policy decision structures. These broader policy questions often require new statistical data or complex new combinations of data that may not be viewed as high priority by an agency trying to fulfill its own mission.

Historically, Federal statistical programs have generally been developed from the bottom up, one program at a time, to support the mission of an agency. Departments and agencies normally resist spending their budgets to serve the needs of users external to that

mission. This tendency is reinforced by many of the organized data users that communicate with the managers of statistical programs and legislators about the agency's mission. Such organized data users frequently have only a narrow or special interest in mind such as concern for a single commodity or expenditures on a particular type of manufactured input.

Pressures to reduce reporting burdens and fiscal budgets that we discussed earlier only magnify the impact that these organized interest groups have on setting priorities for statistical programs. The resource pressures and the lobbying by organized data user groups also make it more difficult to get resources committed to needs for data that are a broader public good. Richard Perrin, chairman of the AAEA Economics Statistics Committee, addressed this type of problem in a response to USDA's proposal to discontinue the quarterly farm labor work. He argued that these data are a public good that will not be collected except by the Federal Government. His concern is that in times of restricted funding there is a great danger that Government data collection may cease where there is the clearest argument for a public good, while collection of other data will continue because the interest of other constituents is sufficiently concentrated to mount an effective lobbying effort.

There is another potential for special interest pressure from within the Federal government which has implications for the statistical system. President Reagan issued an Executive Order in February 1981 and Congress is working on legislation to reduce the burdens of existing and future regulations and to provide for presidential oversight of the regulatory process. For every major rule, each agency is required to prepare a regulatory impact analysis. These analyses are to provide adequate information concerning the need for and consequences of proposed Government action for administrative action. Therefore, while there are pressures to reduce information collection, there likely will also be increased pressures to collect specialized kinds of data for impact analyses.

Implications for the Profession

Hopefully by now, you have reached the conclusion that there is a real need for data users who have a broad, public interest perspective on the agricultural and rural data system to get organized and actively seek to influence priorities and future directions of that data system. This is an important task that the agricultural economics profession should undertake.

The pressures we have discussed can have rather significant implications for the agricultural and other data systems. They can make it difficult to plan and implement a well-coordinated continuing statistical program. They may distort priorities in resource use including public burden. And they may cause statistical programs to be directed by needs perceived by political decision-makers rather than professional considerations of concepts and measures.

The Economic Statistics Committee, which the AAEA has now had for 11 years, has made some useful contributions to recommending future directions for some agricultural data series through studies and workshops. However, these efforts focused only on improving concepts underlying the data system. Such a focus was consistent with the Committee's 1972 report which identified conceptual obsolescence as a major problem of agricultural data. But the work of the Committee stopped short of actively trying to influence priorities for the data system.

We are in no way saying that the conceptual obsolescence problems have been solved or are no longer important. In fact, we are concerned that more progress has not been made in introducing the improved concepts. In critique of these efforts, it appears that there was insufficient follow through by the Committee and others to effectively influence program plans and priorities. Certainly now, in a time of more limited resources, it is going to be even more difficult to get changes made that introduce improved concepts or make other improvements. Only a well organized effort by a broad, professional group can hope to have much influence toward getting changes made in the existing programs.

Although we agree with continued effort by the Committee to improve the conceptual base of the agricultural data system, this environment of burden and budget reductions raises the need for another broad effort by the profession. That additional need is to look at some broad trade-offs in statistical program resource use that have not been examined in any comprehensive way during the past ten years. There are questions, for example, about reducing commodity detail or county level data in order to fill other data needs. Among the many needs that have been identified are finding out who really receives the income from farming or the benefits from commodity programs, how much goes to corporations whose primary business is not farming; improving our information on structure and control of the farming sector so we can more adequately assess trends and the influence of public programs on these trends; improving our socio-economic data related to our land resources; and the need to have an annual survey of farm firms so that we no longer are as dependent on commodity statistics to answer complex questions about the sector or groups of firms. These are major, "direction setting" questions that have received little attention.

In addition to an organized effort by the profession, individuals can also influence the statistical system. There are at least four contacts that individuals can make to express opinions about priorities and future directions for the agricultural data system. We have discussed one which is the AAEA Economic Statistics Committee. Three other contacts are (1) the agency responsible for collecting the type of data you have in mind, (2) the agricultural economics representative to the Agricultural Census Advisory Committee, and (3) the Statistical Reporting Service representatives at one of its data users meetings. Individuals can also have an indirect, but important effect on how respondents react to Federal data collection activities. Many of the fact sheets and newsletters distributed by universities and businesses draw heavily from statistics published by the Federal government. Making the clientele aware of this important source of data can have a positive effect on response rates.

Our attempt in this brief paper is to convince the profession and others that it is time to back off and take a look at the way we now spend money on agriculture data and determine if it should be used to move us in other directions. We think the agriculture economics profession should play a leading role in conducting this comprehensive review because these questions need to be addressed within the framework of longer term policy issues.

Literature Cited

1. Commission on Federal Paperwork. Final Summary Report. Washington, D.C., October 3, 1977.

2. American Agricultural Economics Association, Committee on Economic Statistics. "Our Obsolete Data Systems: New Directions and Opportunities." Amer. J. Agr. Econ. 56 (1972): 867-75.

3. Bonnen, James T. "Improving the Federal Statistical System: Report the President's Reorganization Project for the Federal Statistical System," Statistical Reporter. 80-8 (1980): 197-211.

CHANGES IN THE QUALITY OF AGRICULTURAL STATISTICS--INPUTS, FARM INCOME, OUTPUT, AND PRICES

Bruce Gardner*

Farm Inputs and Income

I would like to distinguish two problem areas: problems of accurate measurement and problems of meaningful measurement. The first category is usually narrowly statistical, the second more broadly conceptual.

With respect to the narrower issues, many farm inputs are more difficult to measure than prices or outputs. Moreover, USDA devotes less effort and resources to generating statistics on some of these. Indeed, with farm inputs we come up against measurement problems that can't be resolved simply by better sampling procedures on a more complete list frame or more accurate statements by respondents. Most notable are questions of the quality of inputs, measuring the service flows from land and capital goods, and measuring nonconventional inputs. Proper measurement of cash outlays for inputs is essential to measuring costs and hence farm income but services of owned inputs are part of a residual. Hence, the farm income task forces of AAEA and USDA (1972, 1975) emphasized problems other than input measurement. But the issue of full and complete measurement of input service flows, even for inputs supplied by the farm operator, is fundamental for productivity measurement. Therefore the AAEA task force on productivity measurement (1980), amplifying criti-

* Professor, Department of Agricultural and Resource Economics, University of Maryland.

cisms originally put forth in the 1960s, emphasized measurement problems.

On the broader conceptual side, in all the task forces, and in the early AAEA Economic Statistics Committee (1972) report and other reports such as Upchurch (1979), conceptual issues of data organization were given prominence. In both measurement and conceptual organization USDA has recently undertaken notable innovations. This section examines the progress made in improving the statistics during the 1970s.

The occasion for such an appraisal is especially ripe because USDA has just published some substantial revisions of its economic data publications. In terms of format, the former Balance Sheet of Agriculture and Farm Income Statistics have been combined in Economic Indicators of the Farm Sector, the first issue of which appeared in December 1980. In terms of conceptual organization of data, there is a substantial effort in the new report to more clearly distinguish data on farm households from data on agricultural business enterprises. For discussion of details, see Nicol (1980). In terms of the substantive quality of the underlying data, no changes have been made that I could detect, but substantial improvements were made earlier in the 1970s, notably the Farm Production Expenditures surveys.

In the AAEA Productivity Task Force report, the allegedly most serious measurement problems involved: (a) the service flows of owned inputs, notably land and labor; (b) the contribution of nonconventional inputs, such as research, infrastructure, and the environment; (c) the quality of inputs. None of these

generates any notable difficulty in measuring income, because income is a residual between receipts and expenses. The problem in productivity measurement is to allocate the residual to various sources, but for calculating income this is irrelevant.

The input data problems more serious for income measurement are measuring depreciation--capital "used up" to generate current receipts--and capital accumulation--inputs used to produce wealth rather than current receipts.

The USDA measure of capital accumulation, "total net investment in farm plant and equipment" in their terminology, has been much improved in the 1970s. Bhatia (1971) provided in my view a quite devastating critique of the estimates of that time. Investment was estimated by using a cross-sectional study from 1955 to extrapolate for later years by the formula

$$I_t = \beta \Delta Y_{t-1}$$

where β was about 4.8. Bhatia discusses the biases that are likely to arise from this approach, and points out that the errors may be very large indeed. Since gross capital accumulation amounts to about half of net farm income, each 10 percent error in estimating capital accumulation generates about a 5 percent or about $1 billion (1960 dollars) error in net farm income.

However, since the early 1970s, USDA has been utilizing the Farm Production Expense surveys to get annual data on investment in capital equipment and structures, which answers Bhatia's objection and should provide much improved data. In recent publications the pre-1970s data on capital expenditures have been extensively revised also. This is good because historical data

are often used for analytical work to aid our understanding of the economics of agriculture. For example, someone might want to use the USDA data to study investment behavior by farmers, to see if it is sensitive to interest rates. If this were done on the pre-1970 data as described by Bhatia, the student would find that a 1-year lagged accerelator model would fit very nicely, and that interest rates made no difference. This is because the data were created by a lagged-accelerator model in which interest rates made no difference.

Of course, the student could have been forestalled if the created data had been labelled as such. A general gripe I have about both the old and the new farm income statistics (and most of the other USDA statistics) is the absence, in the publication, of description of how the data were generated. Especially for constructed measures, there are serious traps for the unwary. I couldn't tell from the publications of the 1980 revised series why or how the revisions were made. For example, the 1962 Farm Income Situation shows differences from the 1980 publication in gross capital expenditures that I don't understand at all, and they are quantitatively substantial (Table 1). For service buildings and structures, the revised figures[1] are about 60 percent higher in the early 1940s, move to approximate equality in the early 1950s, and increase to about 40 percent higher in the early 1960s. Why? (I assume that the 539 figure for 1959 in the revised 1980 data is a typo. I noted also that the value of 1207

---

[1]These are not revisions made along with the changes in format in 1980, but are revisions made earlier in the 1970s.

Table 1

Farm Gross Capital Expenditures

Year	Farm operators' dwellings		Service buildings and other structures	
	A	B	A	B
1940	139	139	165	108
1941	147	147	183	123
1942	126	126	217	135
1943	109	109	269	164
1944	111	111	347	203
1945	125	125	362	249
1946	409	409	752	621
1947	554	554	880	760
1948	702	702	938	877
1949	683	683	887	777
1950	642	739	880	841
1951	665	788	934	897
1952	665	885	949	1,008
1953	619	848	908	965
1954	572	788	853	896
1955	532	766	853	872
1956	529	740	863	842
1957	537	737	874	840
1958	514	700	841	796
1959	539	728	539	829
1960	485	700	1,207	797
1961	592	735	1,156	837

A: USDA estimate of 1980

B: USDA estimate of 1962

for 1960 is the only other "backcast" that was changed between the 1979 and 1980 publications, so I infer a typo there, too. There should not be typos in officially published tables of government statistics. My small smaple indicates more problems in 1980 than earlier. This looks like a decline in productivity in the government sector.)

Revisions can cause special problems when only part of a series is changed. Recently the series on investment in farm operators' dwellings was revised substantially for the 1950s but not at all for the 1940s. This would cause problems for anyone trying to explain this series by a regression equation, especially if one worked with annual changes. The old data show a 10 percent decline in investment between 1949 and 1950, while the revised data show a 6 percent increase!

Depreciation continues to present serious measurement problems even after the 1980 revisions (which did not change the procedures used). Essentially, USDA substracts a fixed percentage of the capital stock which varies from category to category (4.87 percent for dwellings, 7.22 percent for service buildings and equipment, 12 percent for tractors, 14 percent for other machinery, 21 percent for trucks). But it is not adjusted for age of the existing capital stock or changes in its quality. The latter point is important not so much because the durability of specific items has changed but because of heterogeneity within the categories. In particular, it seems likely that structures and equipment have changed over time in that the share of simple buildings, which are relatively durable, has fallen, while the

share of complex equipment like milking parlors or automated grain-handling equipment, which is less durable, has increased. So the depreciation rate should rise. Also, there is the question of obsolescence due to technical change. In short, I don't have much confidence in the depreciation statistics and recent reforms have not taken steps to improve them.

On the issue of conceptual obsolescence, the main issues are: (a) provision of inputs through contractual arrangements that preclude either identification of the price paid pertinent to farm income or identification of returns to inputs as part of farm or nonfarm income, (b) the growth of nonconventional purchased inputs such as legal services or tax advice, (c) counting nominal interest costs as expenses when an expected-inflation premium accounts for most of these costs, and (d) identification of an appropriate set of economic agents whose income to measure.

The new revisions give most attention to item (d), and end up with two alternatives: farm production establishments and farm households. Income of the former is "net farm income" and of the latter "farm operators' income." The old net farm income is essentially "farm operators' income <u>from farming</u>," a smaller number than either of the two new measures. The new net farm income includes returns generated by the activity of farming but accruing to people who do not live on farms, such as nonfarm landlords. (But hired farm workers or providers of custom services who do not live on farms are still excluded.) And the new farm household income excludes income in kind not generated from farming, in particular housing services, which used to be

included.

Schematically, the data choices can be depicted as follows:

	income of farm residents (U.S. Census definition)	income of nonfarms residents
income from farming activities	A	B
income from nonfarming activities	C	D

All would agree that income in cell D should be excluded, but from there things become less clear. The new definitions are, roughly, net farm income = A + B, and farm operators' income A + C, compared to old net farm income = A. The AAEA Economic Statistics Committee (1972) concluded that the farm was obsolete as a basic unit of account for agricultural statistics, and the AAEA task force (1975) and Upchurch (1979) recommended its replacement by establishments, of whatever kind, that produce agricultural products. This is essentially accomplished with the new net farm income concept (Nicol 1980).

The conceptual problems that remain involve items (a) and (b) above. At the practical level, they involve drawing the line between farm and nonfarm residents, and farming and nonfarm activities. Since any boundary will be arbitrary, we might say just pick one and stick with it. Unfortunately, this is impossible. With respect to the residential criterion, we rely on sales of $1,000 to draw the line between rural farm and rural nonfarm.

Even if this is appropriate today, it won't be in a few years. We have to keep moving in order to stay in the same place, as Alice said. Everybody knows this, and USDA and Census try to make appropriate adjustments. More subtle are changes along the other boundary: what activities count as "farming." For example, if a farmer increases his income by intelligent management, the returns are counted as returns to the farming activity. But if the management services are hired, it is not counted as a return to the farming activity. But, a third twist, if the services are provided free by an extension agent, then the returns are again counted as a return to the farming activity. Similar problems arise with respect to legal services, artificial insemination or veterinary services, fertilizer or pesticide application, hired versus unpaid family labor. As with the residence dichotomy, if a boundary were constant, the problem might not be serious. But in fact the economic functions, and methods of paying for them, keep moving back and forth across the "farm gate." This raises again the questions about the farm, however defined, as an appropriate criterion for farming activity. What we really want are returns to certain economic functions, wherever performed. At the same time, the population living on farms as the Census defines them is a set of people whose economic well-being is of interest. Their income is essentially measured in the new USDA household income concept. Thus, with respect to a conceptual basis for both rural economic and social statistics, USDA has made notable progress.

With respect to item (c), the problem of full costing of

interest payments under inflation is related to the issue of counting capital gains as income. The AAEA farm income task force recommended abandoning USDA's old "realized" farm income concept, and this has now been accomplished to the extent that inventory accumulation is included in income. However, capital gains due to price appreciation are still not counted as income. Instead, the new USDA publication provides balance sheet information along with the income flow data. This seems to me the sensible approach, and parallels the data system favored by the accounting profession for economic reporting by publicly-held corporations. Nonetheless, a real problem of meaning of the income measure arises when interest costs include a substantial inflation premium. For this premium will never be offset by corresponding flows of returns in the current year. Therefore, farm income (and corporate profits) are understated when there is anticipated inflation. Gardner and Hottel (1980) estimated the understatement of farm income at about $4 billion for 1979.

Prices of Farm Commodities

With respect to obsolescence caused by institutional change, the key issue in farm prices is the disappearance for some commodities of an observable market price at a time and place approximating the "farm gate." This issue has been covered by Glenn Nelson. I want to mention one technical problem that arises even assuming that USDA has measured the appropriate farm price or shadow price. This is the problem of appropriate aggregation of prices for individual commodities to construct an

overall price index. USDA constructs a chain-linked Laspeyres index, which tends to overstate (or understate) price increases when relative prices are changing, because consumers (producers) gain utility (income) by adjusting their quantity bundles instead of maintaining the base-period quantities. However, one cannot tell a priori if this will be a serious problem. In order to obtain some evidence on this question, I constructed discrete-approximation Divisia indexes (recommended by Christensen 1975 and the AAEA productivity task force 1980) of the basic farm crops and compared them with Laspeyres indexes. The Divisia index is chain-linked annually with crossed value weights, and so is not subject to the Laspeyres bias. Without going into the details here, indications are that any bias in the USDA price series is very small--probably less than 1/2 of 1 percent per year in the estimated rate of change of farm prices in the 1970s.

Farm Output

The flow of farm output is inherently more difficult to measure than farm prices in that the relationship between a particular sample result and the population statistic is less straightforward. A random sample of 10 farmers' prices received for corn can be expected to tell us more about the U.S. market price of corn than a random sample of 10 farmers' outputs will tell us about U.S. corn output. However, USDA's substantial efforts to construct aggregate statistics from farm data for production, farmer-held stocks, and feeding, and from commercial sources for commercial stocks, exports, and domestic disappearance, permit

consistency checks which, at least to the outside observer looking at the published figures, inspire a good deal of confidence. (I've always wondered, though, why rice is the only commodity with a published statistical discrepancy between the supply and demand sides.) Particularly impressive is that USDA keeps revising its estimates even after the "final" annual estimates are in, yet the adjustments are rarely large. Presumably the main new source of information in "revised backcasting" is the Census of Agriculture, a wholly independent survey source and therefore a good source of data for checking.

Speaking of the Census, there is the issue of how much the quality of the benchmark data was harmed by the new procedures and nonresponse in the 1974 Census. One hears griping about this, but I don't know of a full assessment of the problem.

Finally, I want to mention a conceptual problem with farm output that I hold no hope of seeing solved. This is the aggregation of crop and livestock output. USDA makes efforts to exclude from the aggregate that volume of farm crop production (quite large) which is fed to animals, but this cannot be accomplished with great accuracy. And even if it were, we would still have an aggregate output which is a combination of two production levels like "houses and lumber output." This aggregate is economically dubious even if we subtract out the lumber used in houses from the aggregate. However, the fact that crops and livestock are jointly produced on many farms creates problems with separation, also. However, perhaps we can hope for relief from the current practice of counting fish as part of farm output

if produced on farms but not if caught in rivers, lakes, or oceans, or counting timber as farm output if sold from a farm, but not otherwise. (Simunek, p. 38).

References

AAEA-ERS Task Force on Farm Income Estimates. "Report" USDA, mimeo, 1975.

AAEA Task Froce on Productivity Measurement. "Measurement of U.S. Agricultural Productivity," USDA-ESCS, Tech. Bul. No. 1614, Feb. 1980.

AAEA Economic Statistics Committee. "Our Obsolete Data Systems," Am. J. Agr. Econ. 54(Dec. 1972):867-75.

Bhatia, K.B. "The USDA series on Net Investment in Farm Real Estate--A Critique," Jour. Am. Stat. Assn. 66(Sept., 1971):492-95.

Christensen, L.R. "Concepts and Measurement of Agricultural Productivity," Am. J. Agr. Econ. 57(1975):910-15.

Gardner, B.L., and J.B. Hottel. "The Rate of Return to Investment in Agriculture and the Measurement of Net Farm Income," presented at AAEA meetings, 1980.

Nicol, K. "Economic Information for the U.S. Farm Sector: A Revised Format," USDA-ESCS, NED Staff Report, August 1980.

Simunek, R.W. "The Relationship of the Farm Balance Sheet to Sector and National Income and Product," in USDA-SRS, _____, 1977.

Upchurch, M.L. "Steps Toward Better Data for the Food Industry," U.S. Dept. of Commerce, Bureau of the Census, Econ. Res. Rpt. No. 2, Feb. 1979.

FARM TO RETAIL PRICE SPREADS AND EXPENDITURES FOR FOOD

Glenn Nelson [1/]

The quality of the information available on farm to retail price spreads and expenditures for food has improved over the last five years. The purposes of this paper are to review this progress, identify selected areas in need of further improvement, and reflect on the reasons for the progress. A period of five years is convenient since in 1976 a task force jointly sponsored by the Economic Statistics Committee of the American Agricultural Economics Association and by the Economic Research Service of the U.S. Department of Agriculture (USDA) issued its review and evaluation of this information (Brandow, et al.).[2/] Their report provides a benchmark for this paper. The material is organized within four subtopics: (1) price spread information, (2) marketing bill information, (3) other closely related data series, and (4) concluding observations.

Price Spread Information

Price spread series are computed to shed light on questions related to the sources of changes in food prices. The farm value, retail value, and farm to retail price spread are calculated for a market basket of domestically produced food (not including fish), for selected food groups, and for selected individual food items. The farm to retail price spread is further divided into costs cross-classified by input (i.e., labor, packaging, etc.) and marketing function (i.e., processing, retailing, etc.) for a few food items. Monthly data on farm values, retail values, and price spreads are reported in the "Statistical Indicators" portion of Agricultural Outlook (USDA). Annual data, including the cost component data, are published in an annual report which contains considerable textual discussion of developments in the preceding year (e.g., see USDA, 1981).

The manner in which lags within the food marketing system should be incorporated into the price spread series was a hotly debated issue in the early and mid-1970s. The choice of methodology can have major effects on short term changes in the price spread as farm prices change. USDA practice has generally been one of not attempting to reflect marketing lags, which are

---

1/ The author is Associate Professor, Department of Agricultural and Applied Economics, University of Minnesota, St. Paul. Harry Harp, Economist with ERS, USDA, was very helpful in discussing this paper and identifying relevant sources. The comments of Paul Baumgart, Denis Dunham, and the feedback of symposium participants were useful inputs to the revision of the original paper.

2/ The members of the task force were G.E. Brandow (chm.), D.W. Barrowman, P.A. Baumgart, J.W. Hammond, G.L. Nelson, R.W. Ward, and G.E. Worden.

variable among commodities and even within a single commodity over time, when calculating price spreads. The task force supported this position (Brandow, et al., p. 16). A problem which persisted through the mid-1970s, however, was that a backward lag was implicit in the matching of Bureau of Labor Statistics (BLS) retail price data with USDA farm price data for the same month. Retail prices were gathered in the first week of the month while farm prices were usually average monthly or mid-monthly prices. BLS now collects retail food prices throughout the month which is a marked improvement for purposes of calculating price spreads which continue to be computed contemporaneously, i.e., without explicit insertion of a lag. However, the USDA market basket series,which is closely related to but nevertheless distinct from the price spread series for individual products, now incorporates a two week lag in the comparison of farm and retail prices for fresh beef and pork and a one month lag for processed beef and pork. While this latter practice violates the general rule noted above, the lag is somewhat defensible in view of historical patterns and the high profile of the meat group. Emphasizing the obvious, analysts using price spread data for the meat group in the market basket over several years should be aware of the change from a one-to-two week backward lag to a two-to-four week normal lag within the historical series beginning in 1978.

Another major concern in the mid-1970s was the use of conversion factors which had not been subject to recent review and thus were possibly outdated (Brandow, et al., pp. 17-18). Conversion factors represent the quantity of farm product needed to produce one unit of the related retail product. The use of fixed conversion factors when actual factors are changing leads to fallacious elements in the price spread. Recent updates of conversion factors by USDA have made important contributions to more accurate price spread statistics (see, especially, Duewer, and USDA, 1979). USDA has not found sufficiently stable and important seasonal patterns in conversion factors to warrant seasonal adjustments, which was an avenue of investigation suggested by the task force (Brandow, et al., p. 17). A task force suggestion that retail sales of beef cuts may not be in proportion to yields from carcasses due to the selective purchasing of the institutional food market has not led to changes (Brandow, et al., p. 18). USDA has explored the suggestion somewhat but concluded that a satisfactory effort would require more resources than could be justified. The problem is probably becoming more significant as trade in carcasses declines relative to box beef.

Appropriate adjustment for the different quantities which are sold depending upon how items are priced and promoted, and especially whether the item is or is not an advertised "special", has been an important concern. The continuous sampling adopted by BLS, noted earlier, has somewhat alleviated the problem. USDA proposed a cooperative program with industry which USDA analysts felt would give improved data on quantities sold under special for beef, pork, and broilers but the program was not adopted. An imprecise adjustment continues to be made for beef and pork, but no adjustment is made for broilers even though this was recommended by the task force (Brandow, et al., p. 20).

A major concern of the task force was that USDA was attempting, and being pressured further, to compute and report data on a too detailed

level. The task force felt that in some cases, especially that of allocating indirect costs to specific food items when many food items are handled, the detail lacked any conceptual foundation and that in other cases the data were not sufficiently accurate for more detailed analysis. Retail costs for specific food items are no longer allocated to specific cost components, and the number of commodities for which detailed functional and/or cost components details are reported has been reduced from 22 in 1973 to 11 in 1980 (see USDA, 1981, pp. 29-41). On a more negative note, USDA publications still do not contain an explanation of the conceptual problem of partitioning the price spreads (as opposed to margins) into cost and profit components as urged by the task force (Brandow, et al., p. 23).

Weekly estimates of prices and price spreads were computed for beef and pork until recently (USDA, weekly). The calculation and distribution of these weekly spread data were discontinued in September, 1981, which was consistent with the task force's doubts as to the usefulness of weekly information (Brandow, et al., p. 16). Pressures to compute weekly price spreads for commodities other than beef and pork were successfully resisted, as were pressures to report functional and cost component data for time periods of less than one year.

Several other points warrant brief mention. USDA has nearly completed full implementation of a task force recommendation that first point of assembly upon completion of farm production be used to determine farm value (Brandow, et al., p. 16). Consideration of shifting to an agency rather than functional basis for disaggregation of spreads was rejected, as recommended (Brandow, et al., p. 24). Task force recommendations that price spread data be tied to the all urban rather than wage earner price index and that the market basket not be expanded to included imported food and fish were both accepted by USDA (Brandow, et al., pp. 14 and 20). The term "farmer's share", which the task force found unnecessarily value laden, has been changed slightly to "farm value share of retail price"; the latter, however, is not as neutral as the task force's recommended term, "farm value/retail price ratio" (Brandow, et al., p. 37). USDA has chosen not to implement a task force recommendation that updated seasonal indexes be computed for the farm to retail spread for the market basket and selected food groups (Brandow, et al., p. 17). The controversy surrounding the methodology now used by BLS in computing seasonal indices is one of the factors underlying the decision not to move ahead on this topic. Finally, there is a continuing need to eliminate errors and inconsistencies, e.g., the inconsistency between the retailing and wholesaling data for butter in 1980 in two tables on the same page (USDA, 1981, p. 37), through more careful editing and proofing.

Marketing Bill Information

The food marketing bill is the estimated annual difference between the farm value and retail cost of domestically produced food, not including fish. Since the marketing bill is based upon actual total expenditures rather than upon a fixed set of quantity or price weights, the bill incorporates both price and quantity changes.

Several of the task force recommendations and USDA responses with respect to the food marketing bill are analogous to the case of price spreads. Pressures to develop a finer disaggregation have been resisted, pressures to compute a quarterly marketing bill have been resisted, and USDA has changed from an agency to a functional classification of components.

USDA has developed annual estimates of the total value of all food consumed (Manchester and King). This series is a useful complement to the marketing bill data, which include the total value of U.S. farm-produced foods consumed by civilians, and the data on consumers' expenditures for food as shown in the national income statistics, which include the expenditures for food by households and individuals from their own funds. The new, total expenditures series allows the reconciliation of measures as recommended by the task force (Brandow, et al., p. 14) but commendably goes well beyond what would have constituted a minimal effort to satisfy this recommendation.

Other Closely Related Data Series

The development of a new set of price indexes for measuring changes in the costs of processing and distributing food, the "Food Marketing Cost Index", is a major new development (see Harp). These indexes enable more insight into the reasons for changes in food prices and particularly in farm-to-retail price spreads. The task force recommended that such indexes be developed (Brandow, et al., p. 34).

A comprehensive study of the food service industry by the International Food Service Manufacturers Association is nearing completion. A final report is being planned for publication in late 1981. The report will contain basic statistics on 1) the quantity of foods received by establishments, 2) the use of equipment and services, and 3) structural and organizational characteristics of the industry. 3/ This cooperative study by industry and USDA is consistent with task force recommendations (Brandow, et al., pp. 24 and 30-1). The data will be especially useful in improving the quality of the food marketing bill and total food expenditure series.

Two important areas needing additional work and emphasized by the task force remain much as they were five years ago. Data on margins, as opposed to spreads, are still lacking. Occasional studies of earnings on total assets in branches of the food industry would be a useful supplement to the more conventional figures on profits as percentages of stockholders' equity.

Concluding Observations

Much has been accomplished in the last five years. Analytic studies have been completed, as evidenced by a perusal of the "New Publications" section of the National Food Review, in addition to the improvements in

---

3/ For additional information, readers should contact Michael G. Van Dress, Food Economics Branch, NEAD, ERS, USDA.

the published data series. Now would seem an appropriate time to reflect on the role played by the joint AAEA-USDA task force report.

Many, if not most, of the problem areas identified by the task force were known to USDA staff at the time the report was prepared. A major contribution to the task force was to reinforce the staff in their internal efforts to continue ongoing analyses, to implement some changes, and to resist others. In some cases the task force seemingly had the time and the fresh perspective that enabled a clearer statement of the arguments than could have been accomplished by USDA staff alone. The mid-1970s were years in which food prices were highly visible and even the data issues were very politicized (e.g., see U.S. General Accounting Office). A carefully reasoned statement of the problems and preferred solutions from a long term analytic perspective was especially useful in this context.

USDA analysts used the task force report professionally rather than politically. The analysts did not use the report primarily as a blind appeal to authority, i.e., did not flaunt it in disputes with other agencies and with Congress on the basis of having gathered expert opinion and claiming that such opinion was obviously correct. The reasoning in the report was used by USDA analysts as they evaluated alternative procedures, responded to pressures, and proposed procedures. This manner of use is conducive to continued, future use of task forces in a professional and constructive fashion.

The time and material resources which were devoted to this task force have repaid handsome dividends in improved information on food processing and distribution. In my opinion, the most pressing need at this time in this subject area is a careful review of research and policy issues in order to identify those most in need of further analysis, followed by an appropriate analytic program. A joint AAEA-USDA committee might be an appropriate vehicle to organize a conference of public and private sector representatives to conduct such a review. The new types of information discussed in this paper should facilitate the analysis. Eventually, of course, we should expect the analysis to lead to new views on the needed data base -- continuing a cycle which is not discouraging but, to the contrary, a reflection of an area of study which is undergoing exciting developments.

Analysts must be careful, of course, to be vigilant rather than complacent in monitoring the statistical information discussed in this paper. Useful data series may be discontinued if they are not defended, especially in an atmosphere of stringent budgets. Further improvements, some of which are noted in this paper, can and should be made in existing data series and presentations. The gains of the recent past will almost certainly disappear if this area becomes subject to benign neglect.

The discussion in the AAEA symposium subsequent to the presentation of this paper revealed clearly that members of some other task forces feel their efforts were much less successful. Thus, while the experience of the Brandow, et al. task force shows clearly that such a group can have a constructive impact, the experiences of others evidently show that the task

force approach is not uniformly successful. A comparative study of successes and failures might lead to lessons which the Economics Statistics Committee would find useful in establishing a future agenda. Such a study is unfortunately broader than the experience of this author and the scope of this paper.

References

Brandow, G.E., et al., Review and Evaluation of Price Spread Data for Foods, report of a task force jointly sponsored by the AAEA and USDA; ERS, USDA, Washington, D.C., January 1976.

Duewer, Lawrence A., "Changes in Price Spread Measurements for Beef and Pork", Livestock and Meat Situation, LMS-222, ESCS, USDA, August 1978, pp. 33-40.

Harp, Harry H., The Food Marketing Cost Index: A New Measure for Analyzing Food Price Changes, Technical Bulletin No. 1633, ESCS, USDA, August 1980.

Manchester, Alden C., and Richard A. King, U.S. Food Expenditures, 1954-1978: New Measures at Point of Sale and by Type of Purchaser, Agricultural Economic Report No. 431, ESCS, USDA, August 1979.

U.S. Department of Agriculture, Agricultural Outlook, ERS, monthly, except for the January/February combined issue.

_____, Conversion Factors and Weights and Measures for Agricultural Commodities and Their Products, Statistical Bulletin No. 616, ESCS, March 1979.

_____, Developments in Farm to Retail Price Spreads for Food Products in 1980, Agricultural Economic Report No. 465, ESS, April 1981.

_____, National Food Review, ERS, quarterly.

_____, Weekly Estimates of Prices and Price Spreads for Beef and Pork, NED Economic Indicators-Statistics, ERS, weekly through August 1981.

U.S. General Accounting Office, What Causes Food Prices to Rise? What Can Be Done About It?, CED-78-170, Washington, D.C., September 8, 1978.

ECONOMIC STATISTICS FOR AGRICULTURE:

REACTIONS AND COMMENTS

James T. Bonnen*
Michigan State University

Since we are reviewing the results of ten years of effort to improve the agricultural data base, let me start with a little ancient history. The Economics Statistics Committee which has managed this process for the AAEA, was created in late 1969 as a result of the concern and persistent efforts of Joe Ackerman of the Farm Foundation and Harry Trelogan, then Administrator of the Statistical Reporting Service. There was also in existence then a joint American Statistical Association-AAEA Committee on Agricultural Statistics which was advisory to Harry Trelogan and the Statistical Reporting Service. It's long time chairman was Joe Ackerman. I followed him in 1969. The Economics Statistics Committee was created in the same year and I was also appointed its chairman. From 1970 to 1975 we struggled with two broad objectives. One was to sort out and define the problem of an apparently deteriorating data base. This we attempted to do in our 1972 report to the AAEA during its annual meetings. Secondly, we tried to get the profession intellectually aware of and involved in the problems of the agricultural data base. From 1972 through 1975 we worked in task forces and workshops, generally with the USDA, on a number of problems. We also organized several sessions for AAEA annual meetings. In 1975 I left the Committee and made the state of the agricultural data base the topic of my presidential address to the American Agricultural Economics Association. Subsequently, Jim Hildreth, Luther Tweeten, Bruce Gardner and now Dick Perrin have chaired the Economics Statistics Committee. In the period since 1976 this committee has sponsored a very productive series of seminars, conferences, AAEA sessions and joint USDA/AAEA conferences and task forces. The profession is today concerned and involved in debate about many aspects of the data base for which it is professionally responsible.

The papers presented in this symposium make it clear that the agenda of research and action established between the 1972 Economics Statistics Committee statement of the agricultural data base problem and the 1976 Brandow task force sponsored by the Economics Statistics Committee have been seriously examined, many acted on and progress made in improving the data base.

A number of goals have not been achieved. The social and demographic statistics, as well as those for rural development, remain in about the state they were in 1970. In fact there is less information available for small rural areas today than there was in 1970. The improvements are concentrated primarily in

---

* Professor, Department of Agricultural Economics, Michigan State University.

agricultural data. Even here, however, the problem we face, even with the progress that has been made, is that demands on the data system continue to outrun improvements and are still overwhelming the system. There are several reasons for this.

One cause is the changed value of agricultural information. From the 1950s through the early 1970s, we had agricultural programs which, by administrative action, controlled price and often many other conditions of production. From the early 1970s to date, agriculture has operated in a nearly free domestic market. The consequence is a very substantial increase in price variability and other forms of uncertainty. When uncertainty rises, the value of information increases and the expectation as we face the next decade, is one of continued if not higher uncertainty in agriculture. This growing uncertainty arises out of energy and various domestic resource problems and out of the great growth in dependence on exports to foreign markets. The U.S. operates an essentially open market while most of the rest of the world, especially the soviet bloc countries, operate through state trading or some form of socialized purchasing or sales in international trade. As a consequence the worldwide impacts of weather as well as national policies for agriculture are transfered from abroad into the unprotected U.S. market.

The second reason for the increased value of information is that domestically we are moving into an information revolution in agriculture that is proceeding at a very rapid pace. This can be seen in the changes in government and private firm uses of computers and especially software capability and in the construction of interactive capability where none existed previously. Even more profound will be impacts of large commercial farmer's use of computers and computerized information systems. This is growing at a very rapid pace. It used to be said that if you were not a good farm manager, you were not likely to survive. Then in the '60s we began to say that if you were not a good financial manager you would not survive as a commercial farmer. Today it is quite clear that if you are not a good information manager you are not likely to survive. Market uncertainty and the necessity to manage information to control decisions in the context of uncertainty is powering this information revolution.

The third dimension which has increased the value of information is the changing context of public policy decisions. There has been a very substantial increase in the use of statistics to allocate public resources. In the fiscal 1980 federal budget, conservatively 30 percent of the expenditures were indexed and 20 percent were allocated through the use of statistical formulas of various sorts (DeMilner, Emery et.al.). In short, over half of the federal budget (by some measures over 70%) is now allocated by statistical measures. This forces statistical policy and public policy into an interdependent embrace from which neither can escape. The danger of politicizing statistics and statistical policy is obvious.

Compounding this is the increased degree of interdependence and interaction between various policy areas which once were quite separate. I usually describe this by saying that the issues which dominate various public policy sets today are increasingly cross-cutting. In agriculture we used to make domestic agricultural policy decisions without a thought for either international trade or international political or financial implications, environmental degradation, health or safety policies. Yet all of these today in some degree interact with the various farm policy decisions that are supposedly made within the purview of the Secretary of Agriculture. When issues interact in this manner, it often requires new and

complex combinations of data. These data were usually designed initially to support specific and separate areas of policy. Since the issues and decisions now interact, the data are expected to be integrated. But by and large they are not. The need for integrated data increases with interdependence and decision complexity. This is the same as saying the value of information increases with such complexity.

The other dimension of the changing policy context was referred to in Worden and Barry's paper. This is the increased specialization of agriculture which fragments economic and political interests in agriculture, including the interest in data. This now constitutes a major threat to the coherence and comprehensiveness of the agricultural data base as a system. Current budget pressures will undoubtedly call forth a major effort on the part of various interests in agriculture to defend those data that they use, without any concern for the coherence and totality of the system and the social values generated by the fact that we do have in some degree an integrated and partially coherent system today. Indeed, the basic complaint of the Economics Statistics Committee has consistently been that agricultural data is not adequately integrated and coherent. It is not possible for example, today, to describe in a consistent manner the total food and fiber system; while there are detailed statistics for farms, we have very inadequate descriptors for the marketing subsectors and no integrating paradigm. The interaction of tight budgets, inflation and the fragmented interests in agricultural data is likely to push us backward over the next few years as far as the coherence of the data system is concerned.

The papers by Kibler and Kalleck do a fine job of reviewing briefly the structure and methodology of the major data sets produced by the Statistical Reporting Service and by the Census of Agriculture. The first thing that must be recognized is that there have been major changes in methods over the last decade, not only in statistical design but also in the collection and information processing technologies, organization and design. Sample design and census methods clearly have been improved. However, in the process we have traded one set of problems for another. High rates of change in methodology tend to destroy the continuity of time series and to create new sources of error in data. Today, given the shift toward the use of list frames, mail out and mail back collections, new computerized processing methods, automatic editing and other matters, I worry a lot more about nonsampling sources of error and much less about sample error. I am increasingly concerned about the necessity to improve the construction and maintenance of list frames. With the greater necessity to integrate data bases, we also have substantial problems with confidentiality statutes that get in the way of the cooperation of the Census and the USDA, especially in the construction of list frames. For example, the Tax Reform Act excludes the USDA from access to IRS address lists to which the Census does have access.

The other reaction I have to these papers is one I have had many times before. Statisticians do a far better job of providing a description of their methodologies than economists do. Economists not only do not usually provide clear descriptions of their analytical methodologies, but they are prone to use statistical data without reviewing the methodological base or the analysis of the limitations of the data which are normally provided by statisticians.

The paper by Nelson and Gardner reviews a number of ERS statistical series focused primarily on the 1976 Brandow Task Force recommendations concerning farm retail price spreads, marketing bill information needs, farm inputs and farm

income and output measures. This is an excellent assessment of a number of specific data series. There are many quite interesting matters that could usefully be discussed but I will limit my comment to one minor aspect of the paper. Gardner points out that ERS has designed a new format for much of its income and balance sheet data (now published in the Economic Indicators of the Farm Sector). I should like to extend this point. Information system theory tells us that the informational content of data lies not only in its conceptual and measurement structures but in its coding and formating for communication purposes. I consider these format changes to be improvements. There are several examples of format changes in reporting ERS data which reflect improved conceptual integration and revision that is laudable.

The Worden and Barry paper is a very good overview of many of the larger problems of agricultural statistics of which agricultural economists are not usually aware. They point out we are entering a period in which not only the dollar budgets but the paperwork burden budgets (a new control concept) are going to place considerable pressure on statistics. I quite agree with them that the government-wide average reduction in paperwork budgets should not fall on the statistical data base, since it is not the source of the great increase in paperwork burden that has occurred over the last decade or decade and a half. Most of this growth has been in regulatory and tax forms, yet the administration appears to be using an across the board budget reduction technique in implementing the new paperwork act.

The Senate prohibition to taking the follow-up surveys on agricultural services and on farm finance for the 1982 Census of Agriculture has great significance. Increasingly, censuses worldwide are multiple frame vehicles providing complete enumeration for a limited data field combined with surveys of other fields with sampling rates tailored to useage and accuracy requirements. Censuses now commonly combine list frame and area frame universes. When done well this can result both in lower costs as well as improved accuracy. It will not do simply to take a Census of Agriculture. One must combine the complete census enumeration with various kinds of simultaneous and follow-up surveys to do an adequate job today. The fact that one intellectual Luddite in the Senate can create so much damage for so little gain is testimony to the fragmentation and disorder in our political processes today.

After reading these papers and thinking about where we stand today, several thoughts occur to me about the future agenda of the Economics Statistics Committee. It seems appropriate at this point to review the past decade of activities for at least two purposes. One, is to revise and develop a more comprehensive definition of the problem. We should now have a better understanding of the heterogeneous set of problems which we face. These are not adequately described by saying, as we did in 1972, that there is a problem of statistical obsolescence. While this is still true, there are other major difficulties which need to be understood and acted on. In addition, the Committee probably needs to review the past decade of experience for the purpose of developing a specific agenda for action over the 1980's. The AAEA must establish priorities for its commitments to work with government agencies. I would also find useful a complete bibliography of the Committee's many published efforts. For all I know the Committee may well have already started on all of these tasks.

In thinking about where we stand today, my strongest single reaction is a belief that we need an immediate comprehensive evaluation of the agricultural

data base as a system, if we are to protect its coherence against the pressures of single interest data users who are now reacting to budget pressures. If nothing is done, the public good dimensions of the system will soon be riddled with more anomalies and failures to maintain data or to develop the data base adequately. In other words, the question that must be asked is not only, "what do individual users need," but "what are the essential elements and uses of the system?" I believe Worden and Barry's concern in this area is valid. There are several dimensions that perhaps should be elaborated. This task involves not just data base design, but the design of the entire information system, which includes research and analysis. One cannot specify the data base as a system without also specifying the analytical modes through which that data is to be processed. We need an evaluation of the type of analysis and research and that is currently in place and in prospect over the next decade in agriculture. It is possible to anticipate some data needs by examining the new analytical techniques and ideas and their implications for required data. Changes in the nature and power of analysis drives a substantial part of the need for new data or the modification of old data.

The structure of agriculture has changed drastically and continues to change. These changes have many clear implications for the policy agenda and thus for the changing nature of the data base which supports public and private decision making in agriculture. If, as I suggested earlier, we are moving into an information revolution in the management of agriculture, what constitutes a coherent data base for the next decade is a relatively high priority matter. There are many gaps in our knowledge which must be identified for research and which must be resolved before researchers can expect government to take any action. One of the primary reasons why there has been so little improvement over the last decade in the data base for rural development policy is that there has been little intellectual investment in developing an adequate conceptual framework. There are also substantial institutional deficiencies (Rural Society). Before data needs can be specified, one has to have either a clear analytical framework through which you intend to process that data or a comprehensive programmatic decision structure which requires administrative and other data, or both. We need to look at the activities and the structure of agriculture and of rural life and begin to reason about what our data needs are.

REFERENCES

DeMilner, Lawrence E. Indexing the Consumer Price Index: Problems and Alternatives, Washington D.C.: Congressional Budget Office, June 1981.

Emery, Danuta, Valencia Campbell and Stanley Freedman. "Distributing Federal Funds: The Use of Statistical Data," Statistical Reporter, No. 81-3, December 1980, pp. 73-90.

Rural Society in Transition: Statistics for Policy. Statistics for Rural Development Policy Panel, Committee on National Statistics, National Academy of Sciences, Washington D.C.: National Academy Press, 1981.