UNITED STATES DEPARTMENT OF AGRICULTURE
AGRICULTURAL MARKETING SERVICE

THE POSSIBILITIES AND LIMITATIONS OF OBJECTIVE SAMPLING IN STRENGTHENING AGRICULTURAL STATISTICS

By Charles F. Sarle, Principal Economist,

Address, Research Conference,
Cowles Commission for Research in Economics,
Colorado Springs, Colorado, July 3-28, 1939

There is nothing new or startling in the fact that much better agricultural statistics are needed than can be secured with the facilities and techniques now available. The type of agricultural statistics demanded, at any one time has reflected the nature of the current agricultural problems in much the same way as the economic theories in the past have reflected the current business and economic conditions of the times.

During the several decades prior to the 1860's representative farmers of the nation began to demand (1) agricultural colleges where farm boys could obtain an education in agriculture and related and supporting sciences, (2) agricultural experiment stations where improved agricultural practices could be developed under the expert guidance of scientifically trained men and (3) timely agricultural statistics showing the "condition" of the growing crops in the various parts of the country, production of crops at harvest, and numbers of livestock on farms.

They wanted these agricultural statistics and they wanted them to be issued by a neutral agency such as the Government rather than by trade agencies. Farmers believed that crop reports as then issued by trade interests were juggled for the benefit of the buyers and speculators. They believed that reports of a big crop were put out before a crop was sold by the farmers, only to be followed by reports of a much smaller production after the crop had left the farmers' hands.

The State and County agricultural societies that flourished during this period before the Civil War attempted (without success) to make and issue current crop reports. Coordination was difficult and it was impossible to get a sufficient number of farmers to report regularly and consistently. One of the major tasks in 1862 of the newly established government agency that later became the Department of Agriculture was the issuance of official crop reports.

During the half century ending with the World War this nation was a heavy exporter of farm products--cotton, wheat, cattle, and hogs. The demand for agricultural statistics appears to have been reasonably well satisfied with the annual estimates of crop production and inventory numbers of livestock on farms January 1. The need for forecasts of crop production during the growing

season was met by the monthly reports of the "condition" or appearance of the crop. These rather crude statistics presumably reflected the direction of year-to-year change in agricultural production and gave a rough approximation of the extent of the change. The trade early developed a simple method for translating the condition reports into forecasts of crop yield and production--a method that was later (1915) adopted by the Department.

Prices received by farmers for crops as of December 1 each year were estimated and used in calculating the value of crop production, for each crop and for the total of all crops. Likewise the prices of livestock as of January 1 were estimated and used in calculating livestock values. Furthermore, an inquiry concerning wage rates paid farm laborers--monthly rates with and without board and day rates, with and without board--was made 19 times during the period from 1866 to 1909 inclusive, and annually thereafter.

It was not until about 1910 that Nat C. Murray, a foresighted statistician in the Department of Agriculture succeeded in starting a monthly inquiry on prices received by farmers for 42 agricultural products sold from the farm. In fact, one of the reasons for starting this price-reporting service was to provide useful employment for the clerks during the slack time between monthly crop reports. Annual inquiries relating to prices paid by farmers, land values, rents, etc., were began about this time. One is inclined to wonder just what post-war economists, statisticians, and legislators would have done without these pre-war data in calculating price indexes and "parity prices."

During that first half-century of official agricultural statistics the methods of sampling employed were relatively simple. Voluntary correspondents were asked to report on the situation in their immediate locality. The sample data represented their judgment as to a given phenomenon--yield per acre for a crop at harvest, condition of a growing crop in percent of "normal", the current year's acreage of a crop or numbers of a given class of livestock as a percentage of last year's, prices received or paid, prevailing wage rates, and so on. The best and most public-spirited farmers were encouraged to serve as voluntary unpaid reporters.

Accumulative bias in the reports on acreage and livestock numbers became apparent when current estimates in census years were compared with census data. The year-to-year estimates of change applied to the previous census as a base failed to keep abreast of the rapidly expanding agricultural plant during the decades from 1870 to 1900.

Improvements in methods of sampling and estimating during this pre-world war period were made primarily by, (1) increasing the size of the sample, (2) increasing the number of State agents and extending their employment from part-time to full-time, (3) and employing a few crop specialists on a regional basis. The township list of reporters was set up in 1897, thereby greatly increasing the size of the samples over the size of samples obtained from county reporters. That is, the objective was to have an active reporter in each agricultural township reporting for his locality in addition to a reporter in each county reporting for the entire county. The original plan called for several sub-reporters to report to the county reporter on conditions in the parts of the county where they lived--a plan that never seemed to work out very well in practice. Furthermore, it was impossible to persuade more than 50 to possibly 75 per cent of the county reporters to respond in any one month. There was obviously a limit to the amount of service these people would give without financial compensation.

The part-time State agents frequently were men who were on the agricultural staffs of agricultural colleges and experiment stations whose regular duties kept them in touch with agricultural conditions in their respective States. Others were men with political connections. The regional men usually were full-time reporters who traveled over several States. Both State agents and regional men had small lists of crop correspondents reporting directly to them. It was not until 1913 that State crop and livestock estimators were brought under Civil Service.

The entire system of sampling and estimating was highly subjective in nature, in that the observations were opinions expessed in quantitative terms and the sample was obtained from voluntary crop correspondents who were willing to serve without pay.

The World War with its problems of producing food for civilian and army consumption, as well as for export to the Allies, gave rise to the demand for more accurate estimates of agricultural production and for quantitative forecasts of crop production from the monthly condition reports. In April 1918 an inquiry on the supply of and demand for farm labor in percentage of normal was started.

The economic and social mal-adjustments in American agriculture growing out of the World War further increased the demand for more and better statistics, especially those relating to economic aspects of agriculture. Estimates of the annual production of livestock and livestock products and estimates of farm income were started in the 1920's. Estimates of farm employment were begun from samples of the individual farms of crop reporters. The agricultural statistics of the Department began to be used in many ways never anticipated by those who compiled them and who were responsible for their accuracy. Statistical economic analyses were made to determine the factors that influence prices, supply, consumption, and exports of agricultural products. It was during the 1920's that "Big Business" began to demand agricultural statistics of production and income by counties as a basis for determining sales programs, and quotas.

Acute distress arising from the economic effects of the world-wide depression of the 1930's has given rise to many "action" programs by Federal and State agencies, designed to directly alleviate distress conditions of agriculture. The administration of these programs has called for unexpected uses of available agricultural statistics.

County quotas for cotton and wheat acreages had to be established almost over-night and to determine them average yields per acre for the country had to be ascertained. A little later, county quotas for corn and hogs had to be established. Crop and livestock statisticians were drafted for service in connection with the operation of these programs. Other programs, authorized by Congress, called for more and better statistics concerning farm tenure, taxation, credit, and movement of the farm population.

Within the last year the estimates of employment on farms have been used--along with estimates of non-farm and city employment--for making estimates of monthly employment by States. These estimates of employment have been used for making estimates of unemployment by subtracting the number employed from estimates of the total labor force. These estimates of unemployment are one of the factors used in ascertaining the work-relief needs among States. Billions of dollars are allocated among States largely on the basis of these crude estimates of employment and unemployment.

With all this increase in the uses made of statistics during recent years little or no provision has been made for improving the existing methods and facilities available for making agricultural statistics. What can be done to strengthen agricultural statistics to meet these ever-increasing demands? As in the past, the answer lies in the further development and use of objective, as contrasted with subjective methods of sampling and estimating.

The terms objective and subjective apply (1) to the nature of the individual observations, (2) to the method of sampling followed in taking the observations, and (3) to the methods of estimation used.

Three kinds of observations are made in obtaining sample data relating to agriculture: (1) the judgment type of data--the reporter is asked to estimate for his locality such items as the condition or yield per acre or change in acreage from last year for a particular crop, or the change in number of different kinds of livestock, or the prices being paid or received by farmers, (2) the individual farm type of data--the reporter is asked to report for the farm he operates, the acreage in the various crops, production, or numbers of livestock, and (3) the direct count or measurement data--where the phenomenon is sampled by either the crop reporter or the State crop-estimator himself. The measurement of the acreage or field frontages of the various crops along selected highways by the crop-meter, stand and boll counts for cotton, and head sampling of wheat, are all examples of direct observational data.

The method of sampling followed may be subjective in that the selection of persons reporting or farms obtained in the sample is controlled largely by psychological factors as is the case of the returns from unpaid voluntary crop correspondents. The method may be objective in that the selection of observations is pre-determined by the statistician according to sound principles of sampling design.

The method of estimation may be subjective when the statistician depends largely on his (more or less snap) judgment in giving proper weight to two or more indications in making an estimate or objective when the statistician sets up pre-determined weights based on evidence as to past behavior statistical precision of the sample indications or other well reasoned criteria for combining two or more indications in making an estimate.

During the decade from 1910 to 1920 under the stimulus of demands for more accurate statistics arising from the World War, the idea of asking the crop reporter for information concerning his own farm was tried out in connection with acreages in the various crops and numbers of livestock on farms--the individual farm inquiry as contrasted with the judgment inquiry applying to the locality. By the middle 1920's the judgment sample on year-to-year change in acreages and numbers of livestock, was replaced entirely by the individual farm sample.

At first the reporter was asked to state the acreage of the different crops, or numbers of different kinds of livestock on his farm for the current year and the year previous. It was soon found that this current-to-historic comparison was subject to both memory bias--a forgetting of the acreages in some of the minor crops, and to cash crop bias--a tendency to so state the acreages for the 2 years as to minimize the current year's acreage in relation to the acreage in the previous year. The annual June acreage survey, now used in estimating planted acreages, is an inquiry of this kind.

In 1924 the rural-carrier acreage and livestock surveys were started. The rural mail carriers distribute schedules to farmers along their routes. The size of sample was greatly increased. Only data for the current year were obtained. The ratios of the acreages of the various crops to all cropland or to all farm land were used from year-to-year, or from the current year to the census or base year, in making annual estimates of crop acreages. In this way the influence of any constant bias found in the current-to-historic comparison was avoided.

In the case of the livestock surveys, the year-to-year relationship of the average number per farm was used in estimating changes in numbers of livestock.

By 1929, the method of matching reports from the same farms for 2 successive years was adopted to get an additional indication of year-to-year change for both crop acreages and numbers of livestock. Stratification of the sample by size of farms and weighting by census distribution of farms by size also were tried out, but the results did not appear to justify the increased work required.

It took approximately 60 years of experience in crop reporting to prove the very limited utility of the judgment sample of year-to-year change in acreage and livestock numbers and to adopt methods of using the individual farm sample. The judgment type of inquiry, however, has proven satisfactory as a basis for estimating prices and wage rates and fairly satisfactory in estimating yield per acre of crops at harvest and condition of the growing crops. The more intelligent, well-informed, and observant the voluntary reporter, the more reliable are the observations in the case of a judgment type of sample. This selection of better-than-average farmers for reporting on judgment inquiries is an advantage.

When these same reporters, however, are asked to report crop acreages and livestock numbers for the farms they operate, this characteristic of selectivity results in a seriously biased sample. It is selective, or representative, only of those classes of farmers who will voluntarily reply to a questionnaire they receive in the mail or a schedule that is left in their mail box by the Rural Carrier. A large number of farmers--non-resident or "suit-case" farmers in parts of the wheat belt, tenants that move frequently, farmers with a poor education or who are not familiar with the English language, and those of a highly suspicious nature--seldom become regular crop reporters and are not likely to respond to the Rural Carrier surveys. Farms operated by such farmers are seldom represented in the individual farm samples relating to acreages and livestock numbers.

It is true that drought, floods, and powerful economic forces are likely to have somewhat similar effects on all classes of farms and farmers but the fact remains that the established methods of sampling that are available to the Department of Agriculture are subjective rather than objective; consequently the samples obtained are selective and not representative of all kinds and classes of farms.

What has been done and what can be done in the way of developing objective methods of sampling?

In the days before good roads formed a network over the agricultural areas and before automobiles became cheap enough for State agricultural statisticians to own, travel was limited to the railroads. A resourceful statistician conceived the idea of getting a measure of the frontage of the fields in the different crops by counting and recording the number of telegraph poles for each crop along the railroad right of way as he rode along on the train. By going over the same route for 2 successive years he obtained data that would furnish an indication of change in acreage of the various crops found along the route. These were objective observations and although the routes might not be representative, he was able to obtain a sample of all farms along the route--provided of course he could identify the crops in the field as he sped along.

Later the automobile was substituted for the train and telephone poles along the highways were counted or the number of fields in the various crops were counted. Later, about 1925, one of the field statisticians, D. A. McCandliss, invented the "crop-meter" for measuring the frontages in the various crops on one side of the highway; later a "double-bank" machine was perfected that permitted both sides of the road to be "metered" at the same time. Two men in a car are required to measure both sides of the road simultaneously--one to punch the buttons on the crop-meter and the other to drive the car and to check the identification of the crops made by the first man.

The measurement of the frontage of a given crop along a "representative" route taken in the current year is compared with the measurements over the same route for the same crop taken in the previous year and/or the census year, in obtaining an indication of change in acreage. The observational data are objective but the selection of the representative route involves the use of judgment and consequently is subjective and is liable to bias the sample.

A study is being made to test the basic assumptions involved in using the crop-meter data as a basis for estimating acreage change. Aerial photographic survey data for acreage in 19 counties located in 10 North Central States are being used for the purpose of this and other sampling studies relating to acreage. Data for only 1 year have been available for study thus far.

An intensive application of the crop-meter method of sampling is being made in a number of cotton counties where complete AAA acreage measurements will be available for checking the sample.

The crop-meter sample has proved to give a valuable indication of year-to-year change in the acreage of a crop that changes greatly from year to year in an important commercial-producing area, such as potatoes on the Eastern

Shore of Virginia. The crop-meter sample has given a reliable indication of the acreage of cotton in South Carolina and has been used for a number of years for cotton in other States. But in some areas, such as in the low lands of river valleys, the cropping practices and selection of crops along the highways might not be at all representative of those away from the roads.

The crop-meter cannot be used in estimating changes in livestock numbers. The samples of individual farms obtained from voluntary crop correspondents and by the help of the Rural Mail Carriers are not sufficiently representative to meet present-day demands for accuracy and for geographic detail involved in county estimates.

The logical way out of this difficulty in getting selectivity of acreage and livestock samples is to provide for an enumerative sample or sample census each year. Paid enumerators would obtain acreages of the various crops and the number of the different kinds and classes of livestock from a sample of individual farms by personal interview. Dependable statistical methods could be used in designing a sample that would be representative of all kinds of farms in an area. This method of conducting an enumerative sample has been tried experimentally with good results, but is is much more expensive than obtaining a sample by mail or with the help of the Rural Carrier.

When samples are used, instead of a complete enumeration, there is always the problem of having a sample large enough to give "serviceable accuracy" to the average or other statistics calculated from the sample, assuming, of course, that the sample is not selective and is free from bias. Data relating to the acreage of a given crop or numbers of a particular class of livestock on individual farms vary greatly from farm to farm even in the same locality and from year to year on the same farm. For statistical purposes the acreage of a crop on some farms may be "zero," on others it may be hundreds or even thousands of acres. Some farms may not grow the crop at all one year and may have a substantial acreage of that crop in another year. Variability in acreage of a crop has no well-defined limits established by nature as is the case with yield per acre of a crop.

The greater the variability in the data, the larger the sample needed to obtain a desired degree of precision. If one kind of data has twice the variability of another kind, the sample must be four times as large to give equal precision to the averages.

One advantage of the acreage sample of the Rural Carrier is its large size: 10,000 to 12,000 reports are received in some of the larger and more important agricultural States. This number is adequate to give serviceable accuracy for acreage estimates for the entire State, and is reasonably satisfactory in making estimates by crop-reporting districts, of which there are 6 to 10 in the State of average size. But when a sample, even as large as 12,000 is broken down among, say, 100 counties there are only 120 reports for a county; in estimating either acreage or livestock numbers a sample of this size is not sufficient for serviceable accuracy. This principle of statistics explains the apparent paradox that it is easier to make a reliable estimate for an area as large as a State than for territory as small as a county.

The problem of making reliable estimates of yield per acre for various crops does not involve so many serious difficulties as the problem of estimating acreage. The well-informed crop correspondent is usually well qualified to estimate the yield per acre for a given crop on his own farm and usually on a number of farms in the neighborhood. It is entirely practicable, therefore, to use these "judgment" (for the locality) reports on yield as the direct basis for annual estimates of yield per acre. However, the figures for yields per acre derived from census data serve as a useful check in census years and as a measure of bias that is applied to the current year's sample. Any incompleteness in the census coverage of farms is not likely to be serious, insofar as yield per acre is concerned, unless the enumerator consistently fails to get either the lower or the higher-yielding farms in his district. But the nearer the time of taking the census to the completion of harvest, the more reliable are the census data on yield per acre. Late fall would be the ideal time to take Federal census of agriculture--before farmers have moved to other farms.

The problem of selectivity of the sample is much less serious with the judgment inquiry on yields than with the individual-farm inquiry on acreage. Selectivity, however, must be guarded against. In recent years the regular judgment inquiry on yields has been supplemented by an individual-farm inquiry on acreage and production, from which a figure for yield per acre is derived. As might be expected, such an inquiry is selective of the more productive farms in most States and the yields obtained are used primarily in a relative way to indicate year-to-year changes. Judgment yield inquiries obtained from operators of grain elevators also are used in some States to supplement the returns from regular crop reporters.

The yield-per-acre data, whether from the judgment inquiry for the locality or from the individual-farm inquiry, are much less variable than are acreage data. Nature sets up limits to the variation in yield per acre. Consequently, the size of sample need not be nearly so large for a given degree of precision as with samples of crop acreages or livestock numbers on individual farms.

The greatest weakness or source of error in the data and methods now in use in estimating yield per acre is the cash-crop-bias (under-statement) that appears in the sample data for certain, if not all, commercial crops. This cash-crop bias is not consistent from year to year in the same State or among States in the same year.

To meet the problem of cash-crop bias in yield sample data the Department of Agriculture is experimenting with pre-harvest field surveys for cotton, wheat, corn, and apples. There is a need for further development and use of such methods based on physical measurements of yield taken from representative fields whereby the factor of human bias is held to a minimum or is eliminated. Additional trained crop estimators are needed to obtain a sample of sufficient size, as these surveys involve actual counts and measurements--size of ear in the case of corn, head samples for wheat, boll counts, and size of bolls for cotton--made in commercial fields over wide areas.

In the summer of 1938 a survey of wheat was made immediately preceding the harvest in five crop-reporting districts in Eastern North Dakota. East-and-west and north-and-south routes were selected in each district. (There

are nine crop reporting districts in North Dakota.) Wheat fields were selected every 2 miles of wheat frontage on one side of the road along these routes. A crop-meter was used to measure the wheat frontage. Two head samples were taken from each field from a strip not over 125 paces wide that paralleled the highway. In all 232 fields were sampled. Seven varieties were found, although 444 of the total of 464 samples were of three varieties--Ceres, Thatcher, and Durum.

The head samples were threshed later and the yield data were analyzed to determine sources of variability. The greatest source of variability was variety, the next largest source was the geographic districts. The variability among fields was significantly greater than the variability within fields, indicating that maximum efficiency in sampling could be obtained by increasing the number of fields rather than by increasing the number of observations within fields.

The yield obtained from the objectively taken sample was generally higher than the yield reported by agricultural correspondents. This difference is no doubt due in part to losses in grain that occur in harvesting and threshing or combining. A similar pre-harvest survey is now being taken (1939) extending from Oklahoma to the Canadian border. Quality characteristics of protein percentage and test-weight-per-bushel also are being ascertained.

Estimates of production are made by multiplying the estimates of harvested acreage by estimates of yield per acre. In the case of a number of commercial crops, marketing or processing records are obtained and used in truing up estimates of production for the previous year. A complete record of the quantity of cotton ginned each year is made by the Census Bureau. Reasonably complete information is obtained on total annual sales of such crops as tobacco, peanuts, rice, sugar, and crops for canning. Railroad shipments and mill-door receipts are obtained for wheat, rye, and flax in a number of States.

When the total quantity of a crop sold is considered along with estimates of farm utilization for seed, carry-over on the farm into the next season, and other farm uses, an estimate of production can be made for a crop that is independent of the production estimate based on acreage and yield per acre. These estimates based on sales are useful in assisting the crop estimator in finding and measuring bias in the sample data and methods used in making the current estimates. Every effort is made to obtain data on sales, market movement, and processing of commercial crops.

Forecasting Crop Production: The first official forecasts of crop production for most of the major crops are made in connection with the July crop report. The production of fall-sown winter wheat and rye is forecast several months earlier, and cotton on August 1 each year.

The forecasts of crop production, exclusive of tree and bush crops, are made in two or three steps--(1) estimates of acres planted, (2) current estimates of abandonment of acreage since planting for certain crops, and (3) forecasts of yield per acre.

Forecasts of yield per acre before harvest are based primarily on reports of the condition, or promise of yield, of the crop "in percent of normal" obtained on the first of each month from the regular crop correspondents of the Department. Later in the growing season a report on the probable yield per acre also is obtained. Both of these samples are judgment inquiries based on the mass opinion of the reporters and they apply to the locality in which the crop reporter lives.

The condition reports are not so likely to be selective as are the reports on yield per acre or acreage. The condition data also are less variable. The very nature of the condition inquiry leads to less variability in the individual observations. Normal yields may vary considerably from one area to another, even within a county. A condition of 50 percent would mean a yield of 20 bushels of corn in one area and perhaps 25 or 30 bushels in another. In fact, the greatest variability ordinarily shown by a condition sample is about equal to the smallest variability found with judgment-yield sample data (the reports on yield). The most variable yield samples seldom exceed the variability of the least variable samples of individual farms on acreage or livestock numbers.

A forecast of yield per acre is made on the basis of the statistical relationship between the reported condition for a given date and the final yield per acre shown for a period of 15 or more past years. The relationship between November 1 condition to final yield per acre for cotton for the period from 1922 to 1936 for the United States is shown in figure 1. The average relationship in this case is best measured by a straight line, called a regression line. If the currently reported condition is, say, 60 percent, then a forecast of a yield of 172 pounds per acre is indicated by the regression line. The years 1937 and 1938 do not fit in with the relationship shown for the previous years. This interesting point will be considered a little later.

One of the fundamental difficulties in using the reported condition of the crop as a basis of forecasting yield lies in the subjective nature of the sample data--the mass judgment of a large number of crop reporters. There are times when one or more serious insect pests or plant diseases are present in a crop but are not sufficiently in evidence to be readily recognized by anyone not well trained in entomology and plant pathology. Stem rust in wheat is a good example of a disease that is not readily recognized in its early stages of infestation. Under such condition, the crop reporters' estimates of condition are likely to be entirely too high in terms of the final yield of the crop.

One way to meet this difficulty is to have crop estimators, trained in the identification of insect pests and plant diseases and experienced in estimating damage to yields resulting therefrom, make frequent field inspections of the growing crop in important commercial producing areas.

The crop reporter's concept of a normal or full crop--condition is reported in percent of normal--remains reasonably constant over a long period of years. In fact, the crop correspondents are slow in changing their concept of normal when the introduction of improved cultural practices or higher yielding varieties call for an upward revision. For example, the increasing use of soil-building crops and the concentrating of reduced cotton acreage

on the more fertile land on the farms of the South have resulted in substantially higher yields of cotton. The reported condition or appearance of the growing crop apparently tends to reflect the vegetative more than the reproductive aspect of the cotton plant. The reported condition for cotton has failed to allow for the increase in potential yield--the level of the line of relationship between November 1 condition and yield for 1937 and 1938 has apparently been raised some 20 pounds. This apparent change in the level of relationship in the last 2 years is shown graphically in figure 1. This difficulty is partly overcome by asking the crop correspondent to report in terms of probable yield per acre.

The rapid introduction of hybrid corn in the Corn Belt States has had much the same effect in raising the line of relationship for corn condition and yield. The rapid expansion of soybean production in States like Illinois, during recent years, accompanied by the selection of the higher yielding varieties, has resulted in a sharp upward trend in the condition-yield relationship for that crop.

Not only is it possible for the level of the condition-yield relationship to change but the relationship can even change from one with a positive slope (the higher the condition the higher the yield) to one with a negative slope (the lower the condition the higher the yield). Such a change occurred in the case of winter wheat in Maryland soon after 1900. The disease known as septoria nodorum became widespread in Maryland at about this time. It is not readily recognized during the growing season and does its greatest damage under weather conditions that are highly favorable to the vegetative growth of the wheat plant. Consequently a high condition is then associated with a low yield.

Potatoes and peanuts are crops that do not show a high degree of relationship between the reported condition of the growing crop and the final yield. This is not surprising in view of the fact that potatoes and peanuts are both developed underground and the reported condition of the crop tends to relate largely to the appearance of tops or vines. Obviously it is essential that more precise methods be developed for forecasting yield per acre in advance of harvest for these two crops. Some research has been undertaken (financed by the special research funds allotted by the Secretary of Agriculture for this work) to develop methods of forecasting crop yields of a more objective nature. The experience gained in making boll counts of cotton and the results obtained in studies of the relationship of yields to weather indicate three lines of development: (1) methods involving the use of weather data along with condition in forecasting yield, (2) methods based on the relation of yield to structural counts and measurements of the plant characteristics associated with yield from representative samples of the growing crops, and (3) methods involving the use of the direct or indirect influence of weather on crop yields.

The reported condition as of the first of a particular month may not fully reflect the full influence of weather that has prevailed up to that time. A deficiency in subsoil moisture, accompanied by adequate surface soil moisture may give a temporarily high condition for a crop at a given date. If rains are ample and well distributed, a good yield may result, but normal or light rains may not furnish sufficient moisture to carry the crop through to a good yield.

The yield of a crop depends on certain plant characteristics such as stand, size of the plant, the size of the head or ear, and the plumpness of the kernels. Some of these characteristics can be measured well in advance of harvest. The crop-weather research studies regarding wheat, carried on in Great Britain, indicate that height of plant just prior to heading is positively correlated, and the maximum number of tillers early in the season is negatively correlated, with the yield of wheat.

The analysis of head samples of spring wheat representing the eastern half of North Dakota, taken immediately preceding the 1938 harvest mentioned above, shows that height of the wheat plant, number of heads, and length of heads, are all associated with yield per acre. In the case of Ceres multiple correlation was R = .84, with Durum R = .86 and for Thatcher R = .95. These are plant characteristics that could be measured several weeks before harvest, and at a time when the prevalence of insect or plant disease could be detected by a well-trained observer.

For about 10 years, during August and September, the Department of Agriculture has been making boll counts and measurements on cotton in commercial fields along several thousand miles of routes laid out through the Cotton Belt. For the last 2 or 3 years, field observations have been made on corn in the major Corn Belt States during September. The results obtained suggest the feasibility of developing objective methods of forecasting the yield of these two crops based on actual field observation.

In a number of States, studies have been made of the relationship of weather factors to the yield per acre of various crops using averages of weather data and yields for the entire State. Although some of these studies have shown a high relationship for the period of years covered, the forecasting formulas developed have not proved very satisfactory when actually used for forecasting in seasons not included in the study.

During the last 2 years, research of a more intensive nature has been undertaken using weather and yield data from the Agricultural Experiment Stations. Moreover, special field-plot experiments designed for the purpose of studying the effect of weather and soil moisture on growth and yield of wheat, corn, and cotton have been started--one on winter wheat at Manhattan, Kansas, one on corn at Ames, Iowa, and another on cotton at Florence, South Carolina. Methods are being developed for doing this type of crop-weather research.

Five or six years of crop-weather field experiments at some 6 to 10 stations for a given crop would form an excellent basis for ascertaining the direct and indirect influence of weather on crop growth and yield per acre and perhaps would make it possible to develop objective methods of forecasting yield per acre from weather.

One of the inherent difficulties in forecasting the yield per acre in advance of harvest is that the growing and maturing crop is always subject to the hazard of damage either directly from weather or indirectly from disease and insects which are, in turn, largely controlled by weather. On the other hand, the yield of a crop with indeterminate fruiting habits, like cotton, can be greatly increased even late in the season by unusually favorable weather.

The earlier in the growing season the forecast is made the greater is the hazard of subsequent weather. It is not surprising, therefore, that early season condition usually is not sufficiently related to final yield to justify its use in forecasting. For many years, the condition of cotton was obtained as of May 25. In figure 2 it will be seen that the relationship of condition with final yield was too low (T. = .37) to be statistically significant and that condition at that time would be valueless in forecasting yield. The dotted line on this chart is the regression line when the change in condition from May 25 to October is taken into consideration. The dashed line is the "par" line formerly used in interpreting condition. No condition data are now obtained nor are forecasts made of the cotton yield per acre before August 1.

If accurate forecasts of yield per acre are to be made early in the growing season, it will be necessary (1) to know the relationship between weather and yield per acre, and (2) to know how to forecast the weather, or at least the extremes of weather, for several months in advance.

A research program in meteorology that has practically doubled the scanty knowledge in this field has been under way during the last 2 years. It will take time and a great deal of investigation to develop dependable methods of long-range weather forecasting.

Concluding Observations: The ever-increasing demand for more and better statistics relating to agriculture in its many phases can be met by the two governmental agencies responsible for these statistics, (The Agricultural Marketing Service of the U. S. Department of Agriculture and the Bureau of the Census of the Department of Commerce) by the development and practical use of objective methods of sampling and estimation.

Three specific suggestions for using objective sampling methods in strengthening agricultural statistics were made by the writer in a previous paper, 1/ as follows:

1/ "Development of Partial and Sample Census Methods". Proceedings of the Farm Economic Association meeting December, 1938.

(1) Use a split schedule in Census enumerations to broaden the scope of basic agricultural statistics by from 50 to 200 percent.

(2) Take a rotating partial census in intercensal years in commercial fruit and truck areas to obtain more reliable and more useful basic statistics regarding these specialty crops than is possible in connection with a general enumeration.

(3) Take a nation-wide annual enumerative sample of individual farms covering crop acreages and production, livestock numbers and production, etc.

The need for these improved methods has been developed more fully in the present paper, in which a critical description has been given of some of the more important methods of sampling and estimating now used by the U. S. Department of Agriculture and of some of the results obtained in developing more objective sampling methods.

In looking into the future it is not difficult to visualize a return to the county statistical reporter as the source of much of the sample data on which estimates concerning agriculture would be based. He would not be asked to render his service free of charge. He would be given part-time employment and would lose his job if he failed to make the required reports regularly and well. Very little of the data he would submit would be based on his judgment or opinion alone.

Once or twice a year he would take an enumerative sample of individual farms in his county--an objectively designed sample of representative farms. He would obtain information concerning these farms as operating units--crop acreages production and sales, numbers and production and sales of livestock. Each month he would have variously assigned duties--the taking of head samples of wheat in one month, cotton boll counts or ear counts for corn in another month, and perhaps depth of soil moisture another month. Each month he would report on prices received and paid by farmers, numbers of persons employed, and perhaps rate of egg and milk production, as well, on sample farms within the county.

He also might obtain a record of purchases of farm products by local buyers and processors and shipments out of the county by rail and truck of agricultural products to be used with other data in estimating farm income by counties.

Such a program would be more expensive than the present system of voluntary unpaid correspondents, but the resulting statistics would be worth many times the increased cost in the more efficient administration of State and Federal action programs alone.

SHH