YRB 63-04

Forecasting Yields With Objective Measurements

By

Walter A. Hendricks

Statistics Research Division

of

Research Triangle Institute

Reprint Series No. S-31

Reprinted from

Journal of Farm Economics Proceedings Issue

1963

FORECASTING YIELDS WITH OBJECTIVE MEASUREMENTS

BY

WALTER A. HENDRICKS

Reprinted from JOURNAL OF FARM ECONOMICS
PROCEEDINGS ISSUE
Vol. XLV, No. 5, December, 1963

FORECASTING YIELDS WITH OBJECTIVE MEASUREMENTS

WALTER A. HENDRICKS
Research Triangle Institute

As EVERYONE present at this session is aware, crop forecasts have from the beginning held a much more prominent place in official agricultural statistics of the United States than in those of other countries. Regardless of the value that may attach to estimates of final production when the season is over, it is safe to say that they are about as dead as yesterday's newspaper in the minds of people most intimately concerned with production and marketing. Although there is nothing peculiarly American about the need for crop forecasts, government statisticians in most other countries seem to have held the view that this was a form of soothsaying in which they did not wish to become involved. Happily, that view seems to be changing; the introduction of objective methods into forecasting procedures has brought an aura of statistical respectability to it, along with its other accomplishments.

Actually the forecasting process has always been more objective than was generally recognized. Crop reporters do look at their fields and those of their neighbors throughout the season to arrive at appraisals of yield prospects as compared with previous seasons, and their appraisals over the years usually show a reasonable correlation with the final outcome. However, the increasing specialization of agriculture and the growing complexity of marketing processes has brought with it ever increasing demands for more precision. What we now refer to as "objective forecasts" consist of nothing more than a sharpening of the sampling tools used in collecting the data, making detailed counts and measurements on specified plant characteristics that are pertinent to yield, and substituting arithmetical computations in the office for the mental processes that the crop reporter is expected to invoke in arriving at his appraisals.

Obviously, such an approach increases costs and it is necessary to decide whether the increased precision is worth the cost. That decision must be made by the users of the results who are also called upon to pay the bill. The increasing extent to which these procedures are now being applied by both governmental and private agencies, is by itself rather convincing evidence that the answer is at least a qualified "yes."

Definitions and Concepts

When we speak of a crop forecast, we need to specify what we are talking about. An early-season estimate of acreage planted to a crop is, in a sense, a crop forecast because it provides one indication of prospective total production. Similarly, a preharvest estimate of yield, obtained

from sample harvestings before farmers harvest the crop, can be considered a yield forecast even though the crop is already mature. In this discussion we will limit the term "forecast" to a prospective yield per acre, or per plant, computed from plant observations before the crop is mature. But we have to be careful to define what we mean by yield per acre or yield per plant.

The definition of yield is important because it specifies what we are trying to forecast. Furthermore, as the forecast date moves closer and closer to harvest time, the forecast becomes an objective preharvest yield estimate. When such forecasts and objective estimates are compared with census data and other statistics based upon farmers' reports, we are in bad shape if they cannot be reduced to some common denominator. When objective yield estimates were first tried out in the Statistical Reporting Service, we actually found ourselves in that unhappy predicament. We knew how the objective estimates were defined but the definitions underlying reported data with which we were comparing them were pretty nebulous. The growing application of objective forecasting and yield estimating procedures has had the salutory side effect of forcing us to examine more critically what is meant by "yield" and how that term is interpreted by farmers, among others.

For one thing we have to distinguish between a "biological yield" which refers to the total produce present in the field at harvest time and a "net yield" which refers to the amount actually removed from the field under normal farm harvesting conditions, which involve a certain amount of loss. Some individuals want to go a step farther and also deduct losses that occur during storage between harvest time and the time the crop is sold, or otherwise utilized, to arrive at a "net" yield.

When speaking of yield per acre, we also need to consider what is meant by an acre of a crop. Are we talking about the size of the field in which the crop is grown or are we talking about the net area in the field on which plants are actually growing? If the latter, which blank areas are to be deducted and which will be ignored? In addition to affecting the concept of what is meant by yield per acre, these matters have a bearing on the portions of sample fields which are excluded when sample plots are laid out in those fields for observation. When dealing with crops grown under acreage allotments it is also necessary to take account of legal definitions such as the rule that when cotton is planted with two blank rows between every two rows of cotton only one-third of the measured area is counted as cotton acreage.

With respect to the net yield concept, it seems logical to define that yield in terms of amount of produce removed from the field by the farmer at harvest time in the course of his normal harvesting operations.

It does not seem logical to deduct subsequent losses from the yield estimate; such losses should be covered in estimates of crop disposition. I believe this concept has been adopted rather generally in most countries. West German agricultural statisticians, I understand, are deducting from their objective bread grain yield estimates those losses that occur between the time the unthreshed sheaves are taken from the fields and stored in barns and the time they are threshed, which may be several months later. I would prefer not to charge those losses against yield, but, as a practical matter, the most important thing is to have a clear definition of yield regardless of what that definition is, so that anyone using the data knows what is included and what has been deducted. It seems desirable to let that definition come as close as possible to farmers' own ideas of what constitutes yield.

With respect to acreages, it also seems desirable to use definitions that come as close as possible to farmers' own concepts and to compute estimates of yield per acre in conformity with those concepts. That policy was followed by the Statistical Reporting Service when I was in the organization, and I believe it is still in effect. The thinking behind that viewpoint was that acreage estimates are largely based upon acreages reported by farmers and it appears more sensible to put yields on that basis than to make them fit some other acreage definition. For one thing this avoids the necessity for tampering unduly with census acreage data.

Experience has shown that for crops grown under rigid acreage controls, with accompanying field measurements to verify compliance, farmers generally report on a net basis—that is, acreage on which plants are actually growing—and in conformity with official regulations on what is chargeable to their allotments. For other crops reported, acreages tend to be gross areas of the fields.

Forecasting Models

After the definition of yield is established, we are ready to attack the forecasting problem. This involves translating counts and measurements made in the field into an objective indication of prospective yield. I do not wish to take time here to go into the sampling problems involved in selecting the plants or field plots where those detailed counts and measurements are made. The various techniques employed are familiar enough to everyone present here today. It will be of more interest to consider the resulting data and the way they are used.

With respect to field crops, a forecast made early in the season when plants have just emerged is based largely on stand—that is, the number of plants per acre. When the plants begin to fruit, the number of cotton bolls, number of heads of wheat, number of ears of corn, and the like, per acre enter into the picture. But here we must be careful. As of any

date we must know whether all fruit has already appeared or whether there is more to come and how much. This makes it necessary to take account of the stage of maturity of the crop as of that date. As a general principle, the state of maturity of the fruit already present is related to average plant maturity on that date and serves as a guide to the amount of additional fruit, if any, that is likely still to appear. The particular plant and fruit characteristics that serve as indices of maturity vary from crop to crop. The Statistical Reporting Service has made considerable progress in identifying them for a number of our principal crops, but there is still much to be done on others. With cotton, for example, the average plant carrying only squares has about one-fourth of its full fruit load, the average plant with blooms or small bolls but no large bolls has about three-fourths of its full load, and the average plant on which large bolls are already present is carrying its full load. With wheat we can observe the relative numbers of heads in boot, milk, and dough stages as maturity indicators. With corn we have no serious problem because the overall average number of ear shoots per stalk is pretty much of a predetermined constant.

In addition to numbers of fruit per acre, we need to predict the average size or weight per unit of fruit at harvest time. Here too, measurements or weights observed on immature fruit must take account of the state of maturity of the fruit in order to relate those observations to size or weight at full maturity. Here we can also take advantage of such things as the fact that the length of a head of wheat or an ear of corn, which is related to the weight of the mature grain that it will ultimately carry, reaches its maximum long before the grain itself is mature enough to weigh. We must also take account of the expected fruit mortality that occurs between the forecast date and harvest time as well as the expected losses that will occur during the harvesting operation itself. Average losses observed during a number of crop seasons are about the only basis for making such allowances.

There is no magic formula by which a forecasting model can be derived. It can be established for a crop only by painstaking studies of plant growth and development and by accumulating data on all related factors that bear upon the ultimate harvest. Clearly anyone engaging in this form of research must know quite a bit about plant biology and agricultural practices and also be endowed with a good sense of perspective and direction.

In forecasting production of tree fruits, initial counts are usually made after all fruit has already appeared on the trees. The problem reduces to predicting the droppage and harvesting losses that will occur and, in most cases, predicting average fruit size at maturity. Data on droppage and losses are not too difficult to obtain, but predicting fruit size at harvest has presented some serious problems. Periodic size measurements during the growing season apparently serve as a good basis for predicting harvest sizes of oranges and grapefruit. Objective orange and grapefruit forecasts in Florida now have an excellent record over a number of years. For some fruits attempts have been made to relate size at harvest time to size at a specified maturity stage (where such a maturity stage can be defined). With cling peaches in California, for example, harvest size apparently is related to size at the time the pit begins to harden. But these attempts do not appear to have been too successful.

In the case of nut crops, sampling and forecasting problems are similar to those for tree fruits, except that the fruit-size problem is replaced by the problem of forecasting the number of blanks. I have been out of touch with developments in this area during the past few years, but up to that time it appeared to me that the people working with nut crops were having a somewhat better batting average than those working with most tree fruits.

One interesting area of objective forecasts in which little has been done to date is that of forecasting amounts of produce that will be ready for harvest by specified dates, for crops where plants exhibit a continuous fruiting habit throughout the year. The basic problem involved here is that of classifying fruit on the plants by size or maturity stage on the date of the forecast and predicting the changes that will occur in the resulting frequency distribution during the period covered by the forecast. That prediction involves an allowance for fruit growth, recognizing that there are random variations in growth rates of individual fruit, an allowance for mortality, and, for long-term forecasts, an allowance for fruit that may not yet be present on the plants at the time of the forecast. Studies of this kind on crops such as lemons in California and tomatoes in Florida would offer interesting possibilities. Some studies had begun on California lemons while I was still with the Statistical Reporting Service, but I am not acquainted with the direction those studies have taken during the past few years, if they have been continued.

Current Position of Objective Forecasts

At this point it seems appropriate to say a few words about the place that objective forecasts now seem to have reached. They had their inception in an expressed desire for greater precision from a number of different sources, but particularly from organized groups of producers (especially fruit and nut growers operating under Marketing Agreements) and from Congress (especially in the case of cotton). As a result of pilot studies conducted on a few of the principal crops grown rather generally all over the country, such as cotton, corn, and wheat, and on a number of specialized fruit and nut crops, mostly grown by organized

groups under Marketing Agreements, we now have a number of objective forecasting programs functioning at what can be considered an operational level. In my opinion they have fully justified their existence wherever forecasts are actually put to use in reaching important administrative or marketing decisions. Perhaps the best evidence that objective forecasts are here to stay, and that they will be adopted even more widely in the future, is the extent to which large private agencies are beginning to introduce those methods into their operations. Anyone willing to put his own money into a project must believe in it.

In conclusion, a few words of acknowledgment are in order. Those programs now being conducted by the Statistical Reporting Service and its official cooperators would not be where they are now without the initial and continuing support of many other interested individuals and organizations. Some of those people, including the distinguished chairman of this session, are present today. As a former member of the Statistical Reporting Service, perhaps I can take the liberty of speaking for that agency in expressing its appreciation to all of them. Those who have been active in this field have also drawn heavily upon the experience of some people in private industry, who have been making objective forecasts for their own purposes for many years. It is only a fair exchange if they, and others like them who are just now getting into this kind of work on their own, can derive some benefit from what we have learned.