Distr.
GENERAL

CES/AC.61/2003/5
23 April 2003

Original: ENGLISH

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE/EUROSTAT/FAO/OECD Meeting on Food and Agricultural Statistics in Europe
(Geneva, 2-4 July 2003)

STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

FOOD AND AGRICULTURAL ORGANISATION (FAO)

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD)

HOW TO MEET NEEDS FOR AD HOC AND FLASH STATISTICS

Invited paper submitted by the National Agricultural Statistics Service, United States of America*

I. INTRODUCTION

1. There has long been interest in discussing procedures and approaches for the creation and release of ad hoc and flash statistics. Since data systems and user interests vary considerably from country to country there probably is not one simple prescription which can serve for all countries or possibly even for all data requests within one country.

* Prepared by Mr. Rich Allen, National Agricultural Statistics Service, .

The author thanks Ray Bollman and Denis Chartrand of Statistics Canada, and Phil Fulton of the U.S. Economic Research Service for their contributions to the paper.

2. Instead of providing a formula or specific set of procedures, this paper will instead focus on description of desirable approaches. It will also highlight statistical and policy concerns that organizations should consider in creating and, particularly, in releasing new statistical reports. A few examples will be cited. Those are not intended to provide exhaustive coverage of possible approaches, but to illustrate how various considerations have been implemented in practice.

II. DEFINITIONS

3. It is best to start with basic definitions for the two types of statistics. AD HOC statistics, for this paper, are defined as new, unscheduled reports or specific data elements. Ad hoc implies something not originally planned. Ad hoc statistics could be created through a new survey or other data collection effort but also might be generated though analyses of already existing information. In general, ad hoc excludes creation of statistics through specific new funding.

4. FLASH statistics refer to the creation and release of specific answers in relatively short time frames. Flash statistics might be ad hoc in nature or they could include the creation of a system to release time sensitive results on an ongoing basis.

III. OPTIMUM DATA SYSTEMS

5. Since the title assigned for this paper implies a broad approach it might be helpful to do some wishful thinking. Below are described attributes that would be desirable for any statistical organization in order to provide good quality, timely ad hoc and flash statistics. Those attributes basically fall in three categories: baseline information, infrastructure, and policy guidelines. Specific organizations can then determine which attributes are presently available and can attempt to work toward other desirable features. It is assumed that no organization currently has systems in place which incorporate all desirable features for answering the full range of new requests which it receives.

6. The first consideration in order to provide quality ad hoc analysis would be to have all relevant information in a serviceable data base. Excellent, detailed meta data are needed for each data element so analysts can reclassify and resummarize data relationships to meet new requests. Each data item should have many dimensions labeled such as geographic location, size of operation, relevant date for the data element, etc.

7. In addition to well-defined information, it is vital to have fully functional infrastructures for the creation, approval, and release of the ad hoc analysis. Experienced analysts must be on board, along with technical support staff members who can perform special tabulations, summaries, and graphic depictions of the data. There needs to be proper mechanisms for reviewing and approving the release of the new results. Since country level statistical organizations must always be cognizant of their responsibilities for impartial, non-political official statistics, it is important to define the proper approaches for releasing ad hoc statistics. In an optimum system there will also be alternatives for utilizing different release

formats such as hard copy, on-line special releases, posting to on-line data bases, etc., with appropriate review and approval procedures in place for each mode of release.

8. The truly ideal ad hoc statistics system would have such complete data bases and such experienced staff capabilities that analyses could be completed on line and results returned almost immediately. However, even if that capability is created the statistical organization would need to determine if release of new statistics to only the requestor is appropriate. One option to meet the public service obligation would be a system to announce all new ad hoc data releases.

9. The optimum system for creating flash statistics focuses heavily on the statistical organization's methodology staffing. The first determination in creating answers is whether an appropriate sampling frame or data set is available. Is there an ongoing survey effort which can be augmented to provide valid answers? Are opinion data sufficient for meeting the data request or are objective data required? Are some special analysis techniques such as post stratification or re-weighting of results based on other statistical efforts needed?

10. Experienced analysts are particularly important for flash statistics. Not only do they need to summarize any new data collection efforts, they must quickly interpret indications in light of data from other sources or from earlier time periods. There may be reasons for assuming biases in the present results which analysts are able to adjust for. It is also advisable to ensure that a second level of review is built into the approval and release procedures.

IV. HOW TO PROGRESS TOWARD OPTIMUM PROCEDURES

11. As stated earlier, no statistical organization presently has totally optimal data bases, organizational structure, and staffing to fully meet all possible ad hoc and flash requests. There are a number of steps organizations can take to improve their readiness. One is to analyze past requests which had to be deferred or rejected. What would it have taken to meet those requests? Has the organization made any changes which would now allow it to perform those analyses?

12. One approach that can often help organizations to prepare for ad hoc statistics is linkage to census of agriculture or other large scale general purpose surveys. If current crop, livestock, or demographic surveys are selected as sub samples of the census or larger base surveys it may be possible to link backwards for analyses. Similarly, if basically comparable data are collected in a large scale, multipurpose survey each year it might be feasible to summarize across multiple years to extract indications for specific types of operations and make more definitive conclusions than possible from only one year.

V. AD HOC STATISTICS EXAMPLES

13. One example of building a data system for generating ad hoc answers is provided by the ability of the National Agricultural Statistics Service (NASS) of the United States to be able to evaluate the amount of final spring plantings in the case of late planting seasons or to

measure the amount of prevented and lost plantings in the case of serious flooding during planting. NASS conducts two major probability surveys during the first half of June each year that collect planting information. One is an area frame survey in which all farming operations within the boundaries of random segments are contacted, all fields are drawn off on base aerial photographs, and field area and crops or land uses are recorded. The second is a large list survey, conducted mainly by mail and telephone, which collects total area in individual farms planted to specific crops. NASS publishes planted area estimates at the end of June each year, but detailed harvested area forecasts are not published until about August 10, based on smaller scale surveys conducted about August 1. In the event of considerable delayed plantings or serious flooding, NASS can fairly easily implement follow-up surveys of June survey respondents, tailoring surveys to just areas affected by the adverse weather conditions. If the concern is late planting, only operations that had not completed planting at the time of the June interviews will be re-contacted. In the event of wide scale flooding, as occurred in several major U.S. states in 1993, re-contacts are made of operations which had completed planting as well as those which were still planting at the time of the interviews. This capability to re-contact operations and publish updated planting information at the same time as the first harvested area forecasts has been very appreciated by data users. Since there is a human tendency to overestimate the extent of weather problems, being able to create probability-based estimates on an ad hoc basis has been important.

14. Another example of establishing a data system for ad hoc statistics is provided by the Agricultural Resource Management Survey (ARMS), conducted by the United States Economic Research Service (ERS) and NASS. This now serves as the umbrella statistical program for the collection of production expenditures, cost of production, farm income, and chemical use data. An important aspect of ARMS is collection of household income data, as well as farm income, in order to evaluate the full capacity of the farm operation to make financial decisions and adjustments. Even though many aspects of the ARMS data collection change from year to year, key economic and demographic data are collected annually. As a result, ERS now has a powerful data base for ongoing research efforts and answering ad hoc inquiries from the U.S. Department of Agriculture, the U.S. Congress, and data users. ERS has also developed a farm characteristics topology which classifies farms not only by farm income levels but by family or non-family structure and by demographic factors such as retirement and residential lifestyle rural residences. ARMS data are now used extensively for special studies and formed the underpinning for much of the special detailed analyses in a 2001 U.S. Department of Agriculture publication entitled Food and Agricultural Policy, Taking Stock for the New Century.

15. An example of establishing a staff to work on ad hoc statistics requests is provided by Statistics Canada. The Small Business and Special Surveys Division specializes in cost recovery projects and has the survey expertise to quickly organize and conduct new surveys. Statistics Canada Agricultural Division called on this special surveys staff when it received inquiries for new information for policy decisions. The Agricultural Division has also now created a small Marketing and Customer Relations unit which can handle many ad hoc inquiries, particularly internal requests from regional office staff members.

16. An example of determining proper release procedures for ad hoc statistics is provided by recent NASS experience with a request for marketing contract data. Those data had been

collected from corn, soybeans, and wheat producers as part of the ARMS program. Data had been edited but were not intended to be released as stand alone estimates. Staff members from one U.S. Senate office felt that early release would be valuable to producers and requested the data be published without the planned detailed analyses. NASS agreed that the data were publishable but would not provide the data just to the Senator's staff. Instead, NASS created a specific statistical publication and advertised the release two weeks in advance. The Senator's staff members were allowed to review the new publication one hour before release in a secured area which allowed no communications during that hour. At the time of release, the staff members were thus able to quickly communicate with important constituents. There was one other significant consideration in preparing that report. This was a first time effort with no historic comparisons available. NASS made several statistical defensibility decisions such as merging subcategories which had few positive responses. If the marketing data collection is repeated on a frequent basis it might be possible to provide additional subcategory data - - if those subcategories prove to be consistent over time.

VI. FLASH STATISTICS EXAMPLES

17. One example of using an existing survey as the basis of flash statistics was provided in 2001 by the NASS Weekly Weather Crop Survey. That survey is conducted weekly from April through the end of harvest in November. It is an opinion survey which asks knowledgeable people, such as local Extension agents, to evaluate crop progress and crop conditions in their area. Reporting is by electronic mail or regular postal mail - - with the questionnaire for the upcoming week in each state being created based on results for the current report each Monday. (The whole Weekly Weather Crop program might be considered as a flash statistics system in itself.) In 2001, there was widespread concern about the availability of nitrogen fertilizers which are vital for planting the important corn crop. NASS added questions, every four weeks until planting was well underway, on nitrogen fertilizer availability as a percent of normal to provide a measure of expected supplies. Those results were properly weighted to the state and national levels and released in the already scheduled reports.

18. Statistics Canada Agricultural Division has created the ability to convert many requests from Agriculture and Agri-Food Canada which they considered to be ad hoc responses (1 to 2 weeks to answer) to a "flash" response (½ to 1 day) basis by adding staff to their Whole Farm Data Program unit. These staff members have the expertise and systems knowledge to produce new summary outputs very quickly. Since the requests are coming from another governmental unit, Agricultural Division can consider most responses as "pre approved", thus minimizing the review and release time.

19. Another example which somewhat blurs the distinction between ad hoc and flash is provided by some of the work performed in the NASS Data Lab. The Data Lab was established to allow approved non regulatory government and university researchers to access otherwise confidential data sets under tightly controlled confidentiality procedures. NASS Data Lab staff also perform a number of cross tabulation analyses for requesters. Some requests in the past year required extremely quick (3-hour or less) turn around as the U.S. Congress considered provisions for new farm program legislation. In one such example, 1997

Census of Agriculture data were used to quickly estimate how many farms in each state might produce various combinations of crops being considered for program benefits (because of provisions for monetary caps on total benefits). That data set was provided quickly to Congress but then added to the listing of data sets which could be obtained from NASS.

20. The final flash statistics example is related to a data base approach which comes close to providing some on-line responses. NASS has created a "Quick Stats" data base, accessible on the Agency Internet Homepage, which contains nearly all published National, State, and County estimates. Files are presently downloadable, but with limited capabilities to perform mathematical operations within the on-line offering. One interesting use of the new feature occurred when a U.S. Senate staffer wanted average state by state production percentages for the past five years for four major crops. The staffer commented she could download the data and create the calculations, estimating it would take her six hours. Because policy debate was already underway, she asked if NASS would run the calculations. A NASS crop statistician was able to send the completed calculations for all crops in one hour.

VII. SUMMARY

21. Hopefully, the examples above have provided illustrations which clarify the information, infrastructure, and policy needs for creating effective ad hoc and flash statistics. Individual statistical organizations can evaluate their own capabilities and examine what staffing and policy changes might be needed for enhancing their capabilities.

References

Agricultural Statistics Board, "Corn, Soybeans, and Wheat Sold Through Marketing Contracts", USDA/NASS, Washington, D.C., USA, February 2003. (www.usda.gov/nass/pubs/pubs.htm)

Allen, Rich, "Forecasting Crop Acreages and Yields in the Face of and In Spite of Floods", USDA/NASS, Washington, D.C., USA. 1994.

United States Department of Agriculture, "Food and Agricultural Policy, Taking Stock for the Next Century", Washington, D.C., USA, September 2001. (www.usda.gov).
