GAO

Report to the Administrator, Food and Nutrition Service
Department of Agriculture

October 1986

FOOD STAMP PROGRAM

Statistical Validity of Agriculture's Payment Error-Rate Estimates

GAO/RCED-87-4

GAO

United States
General Accounting Office
Washington, D.C. 20548

Resources, Community, and
Economic Development Division

B-217883

October 30, 1986

Mr. Robert E. Leard
Administrator, Food and
Nutrition Service
Department of Agriculture

Dear Mr. Leard:

In response to a February 26, 1985, request by the Chairman, Senate Committee on Governmental Affairs, we have reviewed the quality control system that the Food and Nutrition Service used to determine the fiscal year 1984 error rates in Food Stamp Program benefit payments and to assess (sanction) states for their errors. As part of that review, we examined the soundness of the Service's policy, guidance, and procedures for statistically estimating payment error rates. This letter transmits the results of that part of our review. Our other review results were included in a report to the Chairman, entitled Food Stamp Program: Refinements Needed to Improve Accuracy of Quality Control Error Rates (GAO/RCED-86-195, Sept. 19, 1986).[1]

The Service annually estimates each state's food stamp payment error rate in a two-step process, using state and Service reviews of statistically selected case samples. First, the Service computes an estimate, called a regressed error rate, that is based on its review of a subsample of the state's sample. Second, the regressed error rate is adjusted upward if the state did not complete the number of sample case reviews prescribed in its Service-approved quality control review plan. The resulting number is the official error rate, which is used to determine the state's sanction, if any. (App. I contains a more complete discussion of this process, along with our review objectives, scope, and methodology.)

A sanction involves reducing the federally funded share of the state's administrative costs by a predetermined percentage tied to a target error rate. The target error rate is established by law. The determination of target error rates and sanction amounts is discussed in appendix III.

Overall, the Service's statistical policies for estimating the regressed and official error rates generally conformed with accepted statistical theory.

[1]Our Sept. 19, 1986, report refers to this report as Food Stamp Program: Statistical Validity of Agriculture's Payment Error-Rate Estimates (GAO/RCED-86-188). Please note that the report number has been changed to GAO/RCED-87-4.

However, as discussed in appendix II, we noted minor problems with some of the guidance and procedures used to estimate these rates. Give that sanction amounts are sometimes affected by small changes in the official error rates, we believe it is desirable that the Service use the soundest practical statistical methods in developing the payment error-rate estimates.

Accordingly, we shared with your staff our suggestions for improvements that could be made in the Service's guidance and procedures for estimating error rates. While the staff has taken action to implement certain of the changes we suggested, we are making recommendations t you on the other changes that we believe would further improve the statistical validity of the Service's payment error-rate estimates. (See app. II.) These recommendations address

- the Service's method of calculating the error rate,
- the Service's guidance for determining a minimum sample size of a state's subgroups (strata), and
- the need to calculate the precision of the Service's payment error-rate estimate.

According to Service officials, the Service will decide about implementing the recommended changes to its method of calculating the erro rate after two studies—one by the National Academy of Sciences and the other by the Secretary of Agriculture—of the quality control syste required by the Food Security Act of 1985 are completed. The Service officials said that the Service will be in a better position to decide on these changes after the studies' results are available. The Service officials also said that they planned to routinely calculate the precision of the payment error-rate estimate and that they would consider adding explicit guidance on determining a minimum sample size of a state's subgroups.

The views of directly responsible Service officials were sought during our review and are incorporated as appropriate. We did not obtain offi cial agency comments on a draft of this report.

Copies of this report are being sent to the Chairman, Senate Committee on Governmental Affairs; other interested congressional committees; the Deputy Administrator, Family Nutrition Programs; and the Inspector General, Department of Agriculture. We also will make copies available to others on request.

Sincerely yours,

Brian P. Crowley
Senior Associate Director

Contents

Abbreviations

GAO	General Accounting Office
Hy	Hyattsville, Maryland
RCED	Resources, Community, and Economic Development Division

Background

The Food Stamp Program provides food assistance benefits to households that meet program eligibility requirements. Income, household size, and liquid assets, such as bank accounts, are the principal factors for determining household eligibility. Benefits are issued in the form of food coupons used by eligible households to purchase food and thus obtain a more nutritious diet. The program is administered nationally by the Department of Agriculture's Food and Nutrition Service with 100-percent federal financing of the food stamp benefits—$10.7 billion in fiscal year 1984 and $10.8 billion in fiscal year 1985. States are responsible for local administration and day-to-day operation of the program. States may either administer the program directly through state welfare agencies or supervise its administration by county or city welfare agencies. The federal government finances part (usually 50 percent) of the states' administrative costs. In fiscal years 1984 and 1985, the Service's obligations for state administration totaled $807 million and $912 million, respectively.

The Food Stamp Act of 1977 authorized the Secretary of Agriculture to establish standards of performance for efficient and effective administration of the program and required state agencies to establish procedures for monitoring and reporting on program performance. In response to the act, the Food and Nutrition Service established the current quality control review system, which measures the percentage of benefits issued to ineligible households and overissued to eligible households. In 1980 the Congress established a sanction system that made states financially responsible for a portion of their erroneous payments, on the basis of the results of quality control reviews. To implement this system, the Service established quality control error-rate targets that states had to meet each fiscal year to avoid sanctions (reductions in the federal funds states use to administer the Food Stamp Program). In 198 the Congress revised the sanction system by setting more stringent quality control error-rate targets that state agencies must achieve to avoid sanctions.

The Quality Control and Error-Rate Sanction Systems

The 1982 amendments established the Service's current procedures for applying sanctions against states with error rates in excess of specified target goals. The sanctions are based on the official payment error rate determined by the Service. The payment error rate is the percentage of all food stamp allotments issued in a fiscal year that are either issued to ineligible households or overissued to eligible households.

To determine the payment error rate, the Service relies on information obtained from its quality control system. Under the quality control system, a state is required to take a statistically representative sample of active food stamp cases on the basis of a state sampling plan approved by the Service and determine whether the benefit amounts paid to the households were correct. The Service then independently reviews a subsample of the state's sample to determine if the state review determinations were accurate.

The Service uses information from the completed state and Service quality control reviews to estimate the state's official payment error rate for each fiscal year in a two-step process. First, using statistical formulas, an estimated error rate, called the regressed error rate, is computed. Then, this rate is adjusted upward if the state did not complete its review of the number of sample cases prescribed in its approved sampling plan. The Service makes this adjustment to encourage states to complete quality control reviews. The resulting number is the official (or adjusted) payment error rate on which the amount of a state's sanction, if any, is based.

The official error rate is used to determine if a sanction is required and the amount of the sanction. Each state has a predetermined target error rate established in accordance with the 1982 amendments. If the state's official payment error rate is below the target figure, the state is not sanctioned. If the error rate is above the target, the sanction amount is a percentage of the state's federally reimbursed administrative costs for the fiscal year. As the difference between the official and target error rates increases, so does the amount of sanction.[1] (See app. III for a discussion of target error rates and sanction amounts.) The sanction amount can be very sensitive to small changes in the official payment error rate. While in some cases an increase in the official error rate of 0.99 percentage point will not increase the sanction, in other cases an increase as small as 0.01 percentage point can change the sanction amount by as much as 10 percent of federally reimbursed administrative costs. Thus, the Service needs to ensure that the official error rate is computed as accurately as possible.

[1]The Food Security Act of 1985 prohibited the imposition of sanctions during the 6 months beginning Dec. 23, 1985, and required the Secretary of Agriculture to both conduct a study of the quality control system and contract with the National Academy of Sciences for a concurrent independent study of the system. During the 6-month moratorium on sanctions, the Secretary and the states were required to continue to operate existing quality control systems and calculate error rates.

As of July 1, 1986, 42 states had been assessed 95 sanctions for about $138 million on the basis of their quality control error rates for fiscal years 1981-84. Only three sanctions have been paid—all by Connecticut. The remaining sanctions have been challenged by states in administrative hearings and/or federal court, have been waived by the Service, or were pending as of July 1, 1986.

Objectives, Scope, and Methodology

In response to a request from the Chairman of the Senate Committee on Governmental Affairs, we reviewed the quality control system that the Service uses to determine payment error rates and to assess sanctions.[2] As part of our effort, which this report discusses, we reviewed the statistical validity of the Service's official payment error rates for fiscal year 1984 (the latest year for which complete data were available). Our objectives were to determine the statistical validity of (1) the Service's policy on calculating payment error rates from quality control review data, (2) the Service's policy and guidance on selecting cases for quality control reviews, and (3) the quality control samples actually taken in three states.

To accomplish the first two objectives, we (1) reviewed relevant laws, regulations, and Service guidance, (2) discussed policies and procedures with Service statisticians and program officials, (3) obtained information from relevant reviews and held discussions with the Department's Office of Inspector General, and (4) reviewed statistical literature. We then compared the Service's policies and guidance on sampling, computing the regressed error rate, and computing the values used in the adjustment for incomplete reviews with standard statistical texts to determine if the Service's procedures conformed with accepted statistical theory.[3]

To accomplish the third objective, we obtained information from state program and regional Service officials on how samples for three states—New York, Illinois, and Wisconsin—were taken and compared the procedures actually used with accepted statistical sampling procedures. Through this comparison, we were able to determine the statistical validity of the samples taken in the three states. In selecting these

[2]Results of the overall review were included in a report, Food Stamp Program: Refinements Needed to Improve Accuracy of Quality Control Error Rates (GAO/RCED-86-195), which was issued to the Chairman on Sept. 19, 1986.

[3]William G. Cochran, Sampling Techniques, 3rd ed. (New York: John Wiley and Sons, 1977) and Morris Hansen, William Hurwitz, and William Madow, Sample Survey Methods and Theory (New York: John Wiley and Sons, 1953) were the primary texts used.

states for review, we considered the value of benefits issued, the type of state management structure, and whether the state had been sanctioned because of higher-than-allowed error rates.

To test the effect on official payment error rates of some changes we proposed, we recalculated fiscal year 1984 rates for eight states—Colorado, Florida, Georgia, Montana, New York, South Carolina, South Dakota, and Wisconsin. We selected these states because they took samples that would be most affected by our proposed changes and because the data needed for the test had been collected as part of our broader review. The eight states are in four of the Service's seven regions. The regions represented are the Northeast, Southeast, Midwest, and Mountain Plains.

We made our review from April 1, 1985, to August 15, 1986, and in accordance with generally accepted government auditing standards. The views of directly responsible Service officials were sought during our review and are incorporated as appropriate. We did not obtain official Service comments on a draft of this report.

Details of Findings on Statistical Validity of Payment Error Rates

The Service's policies relating to estimating the states' fiscal year 1984 official error rates generally conformed with statistical theory. We noted that minor improvements could be made, however, in the guidance and procedures used to estimate official error rates. We shared our observations with Service staff and suggested appropriate changes. The Service has implemented or is planning to implement most of the changes we suggested; however, it is deferring decisions on other changes we recommend until after the two studies of the quality control system required by the Food Security Act of 1985 are completed. According to Service officials, the Service will be in a better position to decide on the other changes we recommend after the National Academy of Sciences' and the Secretary of Agriculture's study results are available.

Our findings and the Service's responses are summarized below and discussed in more detail in the remainder of this appendix.

- In calculating the regressed error rate for states using stratified samples,[1] the Service did not always appropriately combine the estimates from the strata. In addition, the Service assumed that the same proportion of cases from each stratum was subject to review (eligible for inclusion in the payment error-rate estimates) and that cases for which reviews had not been completed in the quality control review had the same error rate as cases for which reviews had been completed. Such assumptions would not be necessary under the error-rate calculation procedures that we suggested. The Service has revised its guidance on combining estimates from the strata and has issued guidance to Service regions that will eliminate the assumption about cases subject to review. However, it has deferred a decision on the assumption about the error rate of cases for which reviews had not been completed until the two studies called for by the Food Security Act of 1985 are completed.
- In adjusting the regressed error rate for sample cases for which reviews had not been completed, the Service did not use information it gained from its review of the state sample and did not account for differences in the percentage of cases completed in each stratum. The Service has deferred its decision about using the results of its state sample reviews until the two required studies are completed, but the Service has agreed

[1]Stratified samples are developed by dividing the universe (food stamp caseload) into two or more parts, or strata (e.g., cases active in the first or last half of the fiscal year or households also receiving or not receiving benefits from the Aid to Families with Dependent Children Program). From each part, a random sample is selected. A sample estimate is then determined separately for each part, and the sample estimates from all parts are combined to form an estimate of the universe.

to use a weighted completion rate to account for differences in the completed-case percentages among strata.

- In approving sampling plans for states where stratified samples are drawn, the Service did not require that an adequate number of cases be selected in each stratum. The Service will consider adding guidance on an acceptable minimum sample size for a stratum.
- In computing states' official error rates, the Service has not routinely calculated the rates' measure of precision, which is usually computed for statistical estimates. The Service is planning to modify its computer programs to automatically calculate the rates' measure of precision.

Using quality control review data from eight states for fiscal year 1984, we tested the effect of our suggestions for changing the error-rate calculation. The test showed that our suggested changes would have had little effect on the states' fiscal year 1984 error rates—changing them by 0.11 percentage point at most. These error-rate changes would not have led to changes in the fiscal year 1984 sanction amounts for the eight states we analyzed. However, use of the methods we suggest could lead to sanction amount changes in other situations. The results of this test are discussed in detail later in this appendix.

mproving the Accuracy of Regressed Error Rates

The formulas the Service used to estimate the states' fiscal year 1984 regressed error rates were appropriate, according to statistical theory. However, we identified two problems that occurred in using the formulas that affected those states that took stratified quality control samples. The problems related to the weighting procedures[2] prescribed by Service guidance for combining estimates from the strata and to the Service's assumption that the same proportion of cases in each stratum was subject to review. As discussed in the following subsections, the Service has taken action to correct these problems.

Guidance on Weighting Procedures for Stratified Samples

In a November 1984 memorandum, Service headquarters instructed its regions that states taking stratified samples could weight the regressed error rate in each stratum either by caseload size or by dollars issued, depending on how the state elected to weight its reported error rate. In a meeting with Service staff on September 23, 1985, we pointed out that statistical theory requires that the weights be based on cases rather

[2]When cases in different strata have different probabilities of being included in the quality control sample, the sample results have to be adjusted in order to arrive at the correct estimate for the entire caseload. The mathematical process used in this adjustment is called weighting.

than dollars. We said that because some states with stratified samples had based their weights on dollars, their regressed error rates had been incorrectly calculated.

The Service staff later agreed with us and revised the policy guidance on weighting procedures. Service officials also told us that they had recalculated the fiscal year 1983 and fiscal year 1984 payment error rates that had been affected by the old policy guidance. We were unable however, to determine the effect that correct weighting procedures had on these estimated rates, because other changes (such as correcting data because of the settlement of disputed cases) were made at the same time.

Assumption About Number of Cases Subject to Review

To properly use the Service's formulas for estimating the regressed error rate for states with stratified samples, the number of cases subject to review in each stratum must be known. However, because some of the data needed to correctly use the existing formulas were unavailable, the Service assumed that, in stratified samples, the same proportion of cases in each stratum was subject to review. This assumption could have led to incorrect estimates of the regressed error rate, which the Service defines as the rate of only those cases subject to review.

According to Service regulations, certain types of cases are not subject to review. Such cases, which are excluded from the Service's error-rate estimate, include those in which

- a household was eligible for benefits during the review month but did not receive benefits,
- all members of a household died or moved out of the state,
- the household is under investigation for food stamp fraud, and/or
- reasonable attempts to contact the household in order to perform a quality control review would not or did not succeed.

Although the total number of cases subject to review is required for proper weighting, states usually have information on only the number of cases eligible for benefits or the number of cases that received benefits. Lacking the information needed for weighting, the states generally provided, and the Service used for weighting, the number of cases either eligible for or receiving benefits—implicitly assuming that the proportion of cases subject to review was the same in each stratum. However, when the proportion of cases subject to review varies among the strata, a statistically incorrect estimate may be produced. (The potential for

error is illustrated by an example in app. IV.) If the proportion is similar in each stratum, then the amount of error may be small.

It is not uncommon statistical practice to make assumptions when needed data are unavailable. However, because different views may exist on what assumption is appropriate in any particular situation, we believe that assumptions should be made only when no other practical alternative exists.

Our research indicated that, statistically, this problem might be overcome by slightly modifying the method used to make the estimate for stratified samples. We discussed with Service staff a possible approach that involves estimating the number of cases subject to review from sample data. At our suggestion, the Service tested whether this approach could be adopted for use in calculating the regressed error rate. It concluded that the approach could be used, and it has issued guidance to Service regional staff that will eliminate the assumption about cases subject to review.

Revising the Adjustment Used to Derive the Official Error Rate

In calculating a state's official error rate, the Service adjusts the state's regressed error rate upward if the state did not complete its review of all cases for which reviews should have been completed. The Service does not view this adjustment as statistically based but rather as an incentive to encourage states to complete reviews of sampled cases. In computing the adjustment for fiscal year 1984, the Service used (1) a measure of the variability of the payment error rate, called the standard error or, in the Service's terminology, standard deviation, and (2) the proportion of required sample cases for which the state completed reviews (i.e., completion rate).

Because the Service does not view the adjustment as being statistically based, we did not review the appropriateness of the formula used to adjust the regressed error rate. We did, however, review the Service's method of calculating the factors used in the formula. As discussed in the subsections that follow, we found that (1) the formula the Service used for fiscal year 1984 to compute the standard deviation for stratified samples differed slightly from that in statistical textbooks and (2) states with stratified samples would have had more incentive to increase the number of cases represented by completed reviews if the Service had used a weighted completion rate for such states. Appendix V provides details on the Service's fiscal year 1984 method and our suggested method for calculating the adjustment. Service officials told us

that they had corrected the standard deviation formula and would begin using a weighted completion rate.

Standard Deviation Formula Used for Stratified Samples

The standard deviation formula the Service used for fiscal year 1984 for stratified samples differed slightly from the formula for such estimates found in statistical texts. Appendix VI gives both the Service's and the suggested formulas. When a large percentage of case reviews is completed, the official error rate is relatively insensitive to a change in the estimate of the standard deviation. For example, for New York State, which had a high case-completion rate in fiscal year 1984 (96 percent), the use of the suggested formula changed the official error-rate estimate by 0.01 percentage point. In the 24 other states for which we obtained completion rates, the lowest completion rate was 92.9 percent. However, as the percentage of case reviews that were not completed increases, the value of the standard deviation has more effect on the official error rate. We discussed this problem with Service officials who, on August 15, 1986, told us that they had revised their estimating method so that an appropriate formula for computing the standard deviation would be used.

Accounting for Differences in the Completed Case Rates Among Strata

Although the adjustment for incompletely reviewed cases is intended to be an incentive to encourage states to complete reviews of sampled cases, the Service's fiscal year 1984 method for adjusting error rates did not encourage states taking stratified samples to maximize the proportion of the caseload represented by completely reviewed cases. The Service's method defined the completion rate as the number of cases completed divided by the number of cases in the original sample. This method assumed equal completion rates for all strata. To the extent that this assumption was incorrect, the adjustment for incompletely reviewed cases may not have been an incentive to improve completion rates. If a weighted completion rate had been used in calculating the adjustment for states taking stratified samples, these states would have had an incentive to maximize the proportion of the whole caseload for which a statistically valid error-rate estimate could be made. Service officials agreed and said that they were currently testing procedures for using a weighted completion rate in the future.

Completing reviews of all sampled cases is important because statistically valid error-rate estimates can be obtained only for that part of the caseload represented by completed sample cases. If an error rate for the whole caseload is needed, the error rate for those cases not represented

by the completed sample must be obtained from other information or be assumed. For example, if a state using a nonstratified sample completed reviews of 95 percent of the cases in its quality control sample, then a statistically valid error-rate estimate could be obtained for only 95 percent of the caseload. No statistically valid estimate for the remaining 5 percent of the caseload would be available. For an error-rate estimate to be made for all cases in the state, an error rate for the remaining 5 percent of the cases not represented by the sample results would have to be assumed.

The Service's fiscal year 1984 method of calculating the adjustment for incompletely reviewed cases assigned the same amount of penalty to every case whose review had not been completed. In states taking nonstratified samples, the Service adjustment method provided the states an incentive to maximize the proportion of the caseload for which statistically valid estimates could be made because each sample case represented the same number of cases in the caseload. Therefore, completion of each sample case added the same amount to the proportion of the caseload represented by the sample results. This was not true, however, for states with stratified samples.

In states with stratified samples, a case sampled from one stratum may represent more or fewer cases than a case sampled from another stratum. For example, assume a state divides its caseload into two strata, on the basis of which of two offices processed the case. Further assume that the state's approved plan calls for sampling 1 case for every 1,000 cases processed at Office A and sampling 1 case for every 100 cases processed at Office B. Thus, each case in the sample from Office A represents 1,000 cases in the caseload, and each case sampled from Office B represents 100 cases. If the state must decide whether to complete the review of one case from Office A or to complete the review of one from Office B, and if the objective is to maximize the proportion of the caseload represented by sample results, the case from Office A should be chosen for completion because it represents 10 times more cases than the case from Office B.

In fiscal year 1984 a state taking a stratified sample was given no incentive to complete its review of the case or cases that represented more of the total caseload. Under the fiscal year 1984 system, the potential penalty for incompletely reviewed cases was reduced by the same amount regardless of which case from a stratified sample has been completed. If the Service's adjustment had been based on a weighted case-completion rate, an incentive would have existed for the state in the example above

to complete its review of the case from Office A instead of that from Office B. When a weighted completion rate is calculated, the fact that the sample case from Office A represents 10 times as many cases as the sampled case from Office B is appropriately considered. The adjustment for incompletely reviewed cases would be reduced 10 times as much if the review of the additional case from Office A was completed, compared with the reduction from completing the review of an additional case from Office B.

Service officials agreed that a weighted case-completion rate would provide more incentive to complete reviews of cases representing more of the total caseload in states taking stratified samples, and they said that they were testing procedures to calculate a weighted completion rate and would routinely use weighted completion rates once the tests had been satisfactorily completed.

Alternatives to Two Procedures the Service Used in Computing Payment Error Rates

We identified alternatives to two Service procedures that, in our opinion, would improve the payment error-rate calculation. One alternative relates to the handling of incomplete case reviews in computing the regressed error rate. The other alternative relates to the calculation of the standard deviation used in the incentive adjustment for incomplete case reviews.

The Service stresses the importance of completing reviews of cases selected in the quality control sample, but it acknowledges that certain circumstances may prevent state reviewers from completing reviews of all cases. Such circumstances include those where (1) the household or case record cannot be located or (2) all participants in the household refuse a personal interview or in some other way do not cooperate with the quality control reviewer. In computing the regressed error rate, the Service assumes that the incompletely reviewed cases have the same likelihood of error as those for which reviews were completed. We believe that the Service could simply eliminate this assumption.

In adjusting the regressed error rate upward as an incentive to encourage states to complete reviews, the Service uses a measure of variability called the standard deviation, which it calculates solely from information collected by the states. We believe the Service should revise the calculation of the standard deviation so that it includes the information gained in the Service's review of state quality control cases.

Service officials said that they would consider acting on these suggestions after they receive the results of the two quality control studies required by the Food Security Act of 1985.

ssumption About Error ate of Cases Where eviews Were Not ompleted

In fiscal year 1984 the 25 states we reviewed completed reviews of 92.9 to 99.8 percent of the sample cases found to be subject to review. To compensate for incomplete error amount information, the Service assumed that incompletely reviewed cases had the same error rate as those with completed reviews. Service officials said that this assumption would most likely produce error rates favorable to the states because incompletely reviewed cases were probably more error-prone than were completed cases. However, we believe that the Service should eliminate this assumption and produce a regressed error rate that would represent only completed cases.

As we pointed out in our September 19, 1986, report (GAO/RCED-86-195), the error rate for incompletely reviewed cases may be substantially higher than for completed cases. This would cause the regressed error rate as calculated by the Service to be understated. The amount of understatement may not be large, however, when reviews of a high percentage of the cases have been completed. In New York and the 24 states in the Service's Midwest, Southeast, and Mountain Plains Regions, fiscal year 1984 completion rates ranged from 92.9 to 99.8 percent.

Nevertheless, because of the potential sensitivity of sanction amounts to small changes in the official payment error rate, we believe the Service should consider changing its method of estimating the regressed error rate so that it would represent only completed cases. This change would eliminate the need to make assumptions about the incompletely reviewed cases. If this change is adopted, the weighting procedures would need to be changed for states using stratified samples, because the proportion of completed sample cases could vary among the strata. The approach the Service adopted for eliminating the assumption about cases subject to review could be used to produce a regressed error rate representing only completed cases.

According to Service officials, methods other than the one we propose might be available to handle the problem of incomplete data. They said that the Service will decide about implementing this change after the results of the legislatively mandated studies of the quality control system are available.

Adjustment for Not-Completed Cases Does Not Include the Results of the Service's Reviews

When adjusting the fiscal year 1984 error rates for the cases not completed by the 25 states in our review, the Service increased the states' error rates by generally less than 0.1 percentage point. As we pointed out previously, the Service does not view the adjustment as statistically based but rather as an incentive to encourage states to complete reviews. In calculating the adjustment, the Service used a standard deviation calculated solely from information collected by the state. Therefore, none of the information gained in the Service's review of a subsample of state cases was used in the standard deviation calculation.

We believe that the Service should use the standard deviation of the regressed error rate instead of the standard deviation computed solely from the state results because the regressed error rate includes corrections to errors that occurred in the state reviews. Changing the method of computing the standard deviation would affect only the proportion of the cases that are incompletely reviewed. Generally, this proportion is small.

Service officials said that they used the standard deviation from the state sample because, when they set up the adjustment, they did not have standardized computer programs that calculated the regressed error rate or its standard deviation, while standard deviations from state samples were available from the states. Because of the Service's progress in automating the quality control system, using the standard deviation of the regressed error rate could now be achieved with minimal effort. The standardized programs used to compute the regressed error rates in fiscal year 1984 could be modified to automatically calculate the standard deviation of the regressed error rate. An equation for such a calculation is contained in a June 5, 1984, letter from James L. Solomon, Jr., Commissioner, South Carolina Department of Social Services, to the Service. This equation assumes that the number of cases in each stratum is known. It would need revision if the number of cases is estimated. We have discussed possible revisions with Service officials.

According to Service officials, the Service's decision on this matter will be deferred until the results of the two studies required by the 1985 act are available.

ffect of Our Suggested hanges on Official ayment Error-Rate stimates for Eight tates

Using quality control review data from eight states for fiscal year 1984, we tested the combined effect of making five changes we proposed: (1) appropriately weighting the estimates from each stratum by using cases rather than dollars when computing the regressed error rate, (2) using a statistical approach rather than an assumption about the number of cases subject to review, (3) using a standard deviation formula for stratified samples that incorporates the results of the Service's reviews of subsamples of state cases, (4) using a weighted case-completion rate in calculating the adjusted error rate for states taking stratified samples, and (5) eliminating the assumption that incompletely reviewed cases had the same error rate as those with completed reviews. Table II.1 shows the results of the test.

ble II.1: Comparison of Existing rvice Procedures and Our Suggested ocedures for Estimating Official yment Error Rates

State	Estimated official payment error rate		Increase (+) or decrease (−) using GAO method
	Current policy	GAO method[a]	
Colorado	10.66	10.56	−0.10
Florida	8.95	8.97	+ .02
Georgia	9.56	9.45	−.11
Montana	8.77	8.78	+ .01
New York	10.14	10.13	−.01
South Carolina	10.80	10.80	.00
South Dakota	3.59	3.60	+ .01
Wisconsin	9.60	9.62	+ .02

Note: Comparisons based on fiscal year 1984 data available June 6, 1986, for eight states with stratified samples.

[a]Because of data availability problems, we could not use our suggested formula for standard deviation (see p. 18). We cannot determine if the formula we suggest would yield higher or lower results than presented here. However, we would not expect the results to differ dramatically.

As table II.1 shows, the changes we suggested would have changed the fiscal year 1984 final payment error-rate results by, at most, 0.11 percentage point for these states. The suggested changes would not have led to changes in the fiscal year 1984 sanction amounts for the states we analyzed. However, it is possible that using the methods we suggest would lead to sanction amount changes in other situations. Therefore, we believe that because of the potential sensitivity of the sanction to a small change in the error-rate estimates, the Service should use the soundest practical statistical methods.

Need to Ensure That State Samples Include a Sufficient Number of Cases in Each Stratum

The Service's written guidance on selecting the state quality control samples and the Service's subsamples is generally consistent with soun statistical sampling methods. In addition, implementation of the approved sampling plans was generally good in the three states—Illinois, New York, and Wisconsin—where we reviewed conformance witl Service guidance. However, we found that the Service's guidance does not explicitly address acceptable minimum stratum sample size. Providing such guidance would be useful to help ensure the statistical validity of results from stratified samples.

Guidance on Acceptable Minimum Stratum Sample Size Is Needed

The Service must approve a sampling plan before it is implemented by state. Although the sampling plans in the three states we reviewed gen erally conformed with sound sample design principles, the lack of explicit Service guidance on acceptable minimum stratum sample size led to the approval of one sampling plan that contained few cases in some strata. The Service approved a fiscal year 1984 sampling plan for New York that resulted in few cases being selected in three of six strat In these 3 strata, 5, 11, and 18 cases were selected for review. In one stratum, no error cases were found. Because the plan allowed few case to be selected in some strata, the statistical validity of the error-rate estimate was affected.

When the sample size within a stratum is small, it is possible that no cases with erroneous payments will be found in the sample even thoug the stratum may contain many such cases. When this occurs, no statist cally valid way exists to estimate the error rates and sampling errors f the affected strata. Even when one or more errors are found in each stratum, small sample sizes tend to increase the bias in ratio estimates, such as the food stamp payment error rate.

Service staff agreed that small strata sample sizes caused estimating problems and should be avoided whenever possible. They said that the believed their current guidance almost always guarantees adequate strata sample sizes but that they would consider adding explicit guidance on an adequate stratum sample size.

Sample Selection in Three States Was Generally Good

After approving a state's sampling plan, the Service reviews the state's implementation of the approved plan to ensure that the resulting samp is statistically representative. In the three states where we reviewed th implementation of fiscal year 1984 sample plans, we found that Illinois and Wisconsin's samples were statistically representative, while New

York had some minor sampling irregularities. However, the Service was already aware of these problems through its own review process and was working with the state to correct them.

In New York, clerical errors had been made in selecting the cases for review. Much manual effort was required to draw the sample. Long lists of cases were manually numbered sequentially, listed cases were manually checked against another list to see if they received Aid to Families with Dependent Children benefits, and case sequence numbers were manually compared with a list of numbers that represented those cases that were to be included in the sample. We did not find any evidence to indicate that the clerical errors occurred as a result of a deliberate attempt to bias the state's quality control sample.

Another irregularity we observed was New York's failure to conform with the Service's guidance on drawing systematic samples. Each month, the same numbered case on the list was the first case included in the sample. The effect of using the same starting point is that cases at the very top and bottom of the lists are underrepresented in the sample. If these cases have a different error rate than other cases, then the sample results will be somewhat biased. The amount of bias is apt to be small when a small proportion of the total cases is underrepresented in the sample and such cases have an error rate similar to the rest of the cases. We did not attempt to determine the amount of bias, if any, that was introduced by New York's sampling procedure.

Calculating the Precision of Payment Error-Rate Estimates

The Service's estimates of payment error rates are based on statistical samples. As is true of any estimating method, the statistical sampling estimates may differ from the true value. One advantage of statistical estimates is that a measure of how far the estimated number may be from the true value can be obtained. Thus, the adequacy of sample sizes can be assessed. One measure of precision is called the sampling error of the estimate. The Service does not routinely calculate the sampling error of its estimates. As a result, information about the precision of the Service's payment error-rate estimates is generally unavailable.

Even though the sampling error of the official payment error rate does not affect sanction amounts, we believe that the Service should routinely calculate the sampling error for the following reasons:

- Accepted statistical practice requires that sampling error be calculated so that the precision of the estimate can be judged.

- Sampling errors, given current sample design and sizes, can be used to determine the need for or desirability of changing required sample sizes or sampling designs.
- Calculating sampling errors at the same time other estimates are made is less expensive and time-consuming than calculating them when a later specific need or request for the data occurs.

In addition, congressional interest has been expressed in the precision of the estimates, and legislation has been proposed to base sanctions on an error-rate estimate lowered by using the sampling error estimate. For example, the Chairman, Subcommittee on Domestic Marketing, Consumer Relations, and Nutrition, House Committee on Agriculture, asked us to calculate the sampling error for fiscal year 1983 food stamp error rates.[3] In addition, Congressman James Jeffords introduced H.R. 2621 on May 23, 1985, which, among other things, would change the basis for assessing state sanctions to a lower error-rate estimate that would be calculated using the sampling error of the current estimate. That bill has not been enacted.

The Department's Office of Inspector General recommended in a December 14, 1981, report that the Service calculate sampling error when the official payment error rate is used to sanction states.[4] The report pointed out that the sampling error of the payment error-rate estimates varied from state to state, yet this information was not included in the Service's reports. The Inspector General concluded that since the Service's payment error rate is a statistical estimate and all statistical literature is in agreement that both the estimate and its precision should be included in any analysis, the Service should calculate sampling error to determine that precision.

Service officials told us that the Service had not calculated sampling errors in the past because they do not affect the sanction amount and would have been difficult to calculate and review without standardized computer programs. The computer programs used in fiscal year 1984 to generate the regressed error rate, however, could be readily modified to automatically produce the sampling error estimates. The Service officials agreed that the standardized programs could be changed and said that they were planning to make this change.

[3]See Quality Control Error Rates for the Food Stamp Program (GAO/RCED-85-98, Apr. 12, 1985).

[4]Food and Nutrition Service Food Stamp Program Nationwide Audit of the Quality Control System (27627-2-Hy, Dec. 14, 1981).

Conclusions

In general, the Service used sound statistical procedures to arrive at its fiscal year 1984 regressed payment error rates. However, we noted some changes the Service could make to improve the statistical validity of the estimated error rates. For the states we reviewed, the changes we suggested had only a small effect on the error rate and implementing them would not have led to changes in the fiscal year 1984 sanction amounts. However, they could lead to sanction amount changes in other situations.

In line with our suggestions, the Service has taken action to revise its guidance so that cases rather than dollars will be used for weighting purposes when combining estimates from different strata to compute the regressed error rate. It also has issued guidance to eliminate the assumption that the same proportion of cases from each stratum was subject to review. The Service is testing, and said it will implement, procedures for using weighted case-completion rates and said it will use a textbook formula to compute the standard deviation in calculating the adjusted error rate for states taking stratified samples.

In two cases, the Service is deferring decisions about changes we suggested until the results of the National Academy of Sciences' and the Secretary of Agriculture's studies of the quality control system are available. In the first of these cases, we believe that in computing the regressed error rate, the Service could eliminate the assumption that incompletely reviewed cases have the same error rate as cases for which reviews were completed. Service officials believe that there may be other ways to handle the problem of incomplete data and said that they will decide about implementing this change after the results of the legislatively mandated studies are available.

In the second case, we believe the Service should use the standard deviation of the regressed error rate, which includes the results of the Service's review of a subsample of state cases, in calculating the adjusted error-rate rather than relying on state review results alone. Because of the Service's progress in automating the error-rate computations, calculating the standard deviation of the regressed error rate could be achieved with minimal effort. An equation that could be used as a starting point is available, although possible revisions, which have been discussed with Service officials, may be needed.

The Service's guidance to its regions and the states on sample selection procedures generally conformed to accepted statistical sampling theory. Although we did not find any major problems with implementation of

sampling plans in fiscal year 1984 in the three states where we reviewe such implementation, small strata sample sizes in New York somewhat affected the statistical validity of that state's error-rate estimate. Explicit guidance on an acceptable minimum stratum sample size, whic Service officials said they would consider adding, would help avoid this problem.

Because the Service has not routinely calculated the sampling error of its estimates, information about the estimates' precision has not been readily available and the adequacy of current sample sizes could not be readily assessed. Routinely calculating sampling error would conform t accepted statistical practice, help to determine the need for or desirability of changing sample sizes or sampling designs, and be less expensive and time-consuming than calculating the sampling error when a specific need or request for the data occurs. Service officials agreed tha the computer programs could be modified to automatically produce the sampling error estimates, and they said that they were planning to mak such a change.

Recommendations

We recommend that the Administrator, Food and Nutrition Service, amend Service policy and guidance to

- require that an acceptable expected minimum stratum sample size be established when approving a state's sampling plan;
- require the routine calculation of the sampling error of payment error-rate estimates; and
- in conjunction with actions taken in response to studies that Agricultur and the National Academy of Sciences are conducting on food stamp error rates and sanctions,
  1. eliminate the assumption that cases for which reviews have not been completed have the same error rate as completed cases and
  2. require that the standard deviation of the regressed payment error rate, instead of the standard deviation from the state sample, be used i calculating the adjusted, or official, error rate.

?ayment Error-Rate Target Goals and ;anction Amounts

The Food Stamp Act Amendments of 1982 established payment error-rate target goals of 9 percent of a state's total issuances for fiscal year 1983, 7 percent for fiscal year 1984, and 5 percent for fiscal year 1985 and beyond. However, the amendments also provided that any state with a payment error rate exceeding 9 percent during a 6-month base period—October 1980 through March 1981—could avoid a sanction for fiscal years 1983 and 1984 by meeting an individually determined target error rate. The target of such states for fiscal year 1983 was a reduction in a state's actual error rate equal to at least one third of the difference between its base-period rate and the 5-percent target for fiscal year 1985. For fiscal year 1984, the reduction was to be at least two thirds of this difference. For fiscal year 1985 and beyond, all states were to have payment error-rate targets of 5 percent.

For example, a state with a 14-percent error rate in the base period would have had to reduce that error rate by at least 3 percentage points in fiscal year 1983 (one third of the 9-percentage-point difference between 14 percent and 5 percent) and by at least 6 percentage points (two thirds of the original 9-percentage-point difference) in fiscal year 1984. The state would have to further reduce its error rate by another 3 percentage points in fiscal year 1985 to achieve the 5-percent error-rate target.

The amendments also required that the federally funded share of the state's administrative costs be reduced by 5 percent for each of the first 3 percentage points or fraction thereof by which the state's payment error rate exceeded the maximum rate for a fiscal year and by 10 percent for each additional percentage point or fraction thereof by which the maximum payment error rate for the fiscal year was exceeded. For states with 5-, 7-, and 9-percent error-rate targets for fiscal year 1984, for example, the sanction amounts would have been as follows:

Table III.1: Sanction Amounts for States With Error-Rate Targets of 5, 7, and 9 Percent

Payment error rate (percentage)	Sanction as a percentage of federally reimbursed administrative costs for state with error-rate target of		
	5 percent	7 percent	9 perce
5.00 or less	none	none	no
5.01 — 6.00	5%	none	no
6.01 — 7.00	10%	none	no
7.01 — 8.00	15%	5%	no
8.01 — 9.00	25%	10%	no
9.01 — 10.00	35%	15%	
10.01 — 11.00	45%	25%	
11.01 — 12.00	55%	35%	
12.01 or more[a]			

[a]The sanction increases by 10 percentage points for each additional percentage point or part of a percentage point in the payment error rate. No state, however, can be sanctioned more than the value of benefits issued in error above its target.

?otential Error When Assuming an Identical ?roportion of Cases Subject to Review in \ll Strata

When the assumption is made that the proportion of all cases subject to review is the same in each stratum, the potential exists for inaccurate estimates. A hypothetical state where cases are divided into two strata illustrates that potential bias, as shown in table IV.1:

ble IV.1: Hypothetical Case Data for ate With Two Strata

	Strata	
Cases	**1**	**2**
Eligible for payment	100	150
Subject to review	100	100
Subject to review that are in error (percentage)	5%	20%
Subject to review that are in error (number)	5	20

From the above information, we know that for the entire state, 25, or 12.5 percent, of the 200 cases subject to review are in error.

Next, suppose that we know for each stratum the number of cases eligible for payment—100 in stratum 1 and 150 in stratum 2—but not the number of cases subject to review. In terms of the number of eligible cases, stratum 1 contains 100 of the 250 total cases, or 40 percent, while stratum 2 contains the remainder, 60 percent. If we assume that the same proportion of eligible cases in each stratum is subject to review, then stratum 1 will contain 40 percent of the cases subject to review while stratum 2 will contain the remaining 60 percent. Using the percent of cases subject to review and taking a weighted average leads to an estimate of 14 percent.

$$\left(\frac{100}{250} \times 5\%\right) + \left(\frac{150}{250} \times 20\%\right) = 14\%$$

In this example, the assumption that each stratum contained the same proportion of cases subject to review led to an estimate of a 14-percent error rate when, as shown above, the error rate was actually 12.5 percent. This example is illustrative only. We do not intend to suggest that the differences in actual error rates would be as large as shown in this example.

Formulas for Adjusting Error Rates for Incompletely Reviewed Cases

The Service's fiscal year 1984 formula for the adjusted error rate (A), which was the same for both stratified and nonstratified samples, was as follows:

$$A = r(R) + (1-r)(R+2QCSD)$$

The definitions the Service used for the symbols in the formula, along with the definitions we suggested, are given in table V.1 for nonstratified samples and table V.2 for stratified samples.

Table V.1: Definitions Used in the Adjusted Error-Rate Formula for Nonstratified Samples

Symbol	Current Service method	Suggested method
R	Regressed error rate	Same as current method
r	Proportion of required sampled cases subject to review that were completed by the state	Same as current method
QCSD	Standard deviation of the state-estimated error rate based on the full quality control sample[a]	Standard deviation of the regressed error rate[b]

[a]See app. VI for the formula.

[b]See p. 18 for a reference to proposed method.

Table V.2: Definitions Used in the Adjusted Error-Rate Formula for Stratified Samples

Symbol	Current Service method	Suggested method
R	Regressed error rate assuming that the same proportion of cases are subject to review in each stratum and that incompletely reviewed cases have the same error rate as completed cases	Regressed error rate for completed subject-to-review cases
r	Unweighted proportion of cases subject to review that were completed	Weighted proportion of cases subject t review that were completed
QCSD	Standard deviation of the state-estimated error rate based on the full quality control sample[a]	Standard deviation of the regressed error rate[b]

[a]See app. VI for the formula.

[b]See p. 18 for a reference to proposed method.

Formulas for Computing the Standard Deviation From the Full State Sample

imple Random Sample ormula Used by ervice in Fiscal Year 984

The standard deviation is computed as the square root of:

$$\frac{s^2}{n\bar{y}^2}$$

where:

$$s^2 = \frac{\Sigma \quad (x_i \; - \; py_i)^2}{n-1}$$

x_i = amount of payment in error to sample case i

y_i = total allotment issued to sample case i

n = total number of completed sample cases

$$p = \frac{\Sigma x_i}{\Sigma y_i}$$

$\bar{y} = \frac{\Sigma y_i}{n}$ = average allotment per household

iscal Year 1984 ervice Formula for tratified Samples

The standard deviation is computed as the square root of:

$$\frac{1}{N^2} \Sigma \frac{N_h^2 s_h^2}{n_h \bar{y}_h^2}$$

and where:

N = total number of cases

N_h = total number of cases in stratum h

s_h^2 =
n_h = } defined as shown on preceding page for each stratum h
$\bar{y}_h$ =

Suggested Formula for Stratified Samples

According to statistical texts,[1] the standard deviation would be computed as the square root of:

$$\frac{1}{N^2\bar{Y}^2} \sum \frac{N_h^2 s_h^2}{n_h}$$

and where:

$\bar{Y}$ = true average allotment per household.

Generally, the true average allotment per household ($\bar{Y}$) is not known; however, an estimate ($\bar{Y}_e$) is available from the state sample. This estimate is calculated as:

$$\bar{Y}_e = \frac{1}{N} \sum N_h \bar{y}_h$$

The formula the Service uses for stratified samples is not algebraically equivalent to the results obtained by substituting the estimated average allotment per household into the textbook formula. We believe the latter approach would yield a more statistically correct estimate.

[1]Cochran, Sampling Techniques, and Hansen, Hurwitz, and Madow, Sample Survey Methods and Theory.

Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are $2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300

Address Correction Requested

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100