Rethinking Quality Control: A New System for the Food Stamp Program (1987)

Pages
192

Size
5 x 8

ISBN
0309310857

Dennis P. Affholter and Fredrica D. Kramer, Editors; Panel on Quality Control of Family Assistance Programs; Committee on National Statistics; Commission on Behavioral and Social Sciences and Education; National Research Council

Rethinking Quality Control: A New System for the Food Stamp Program

Dennis P. Affholter and Fredrica D. Kramer, editors

Panel on Quality Control of Family Assistance Programs

John C. Neter, Chair

Committee on National Statistics
Commission on Behavioral and Social Sciences and Education
National Research Council

National Academy Press
Washington, D.C. 1987

This report was prepared for the Food and Nutrition Service, U.S. Department of Agriculture, under contract number 53-3198-6-29. Points of view or opinions stated in this document are those of the contractor and do not necessarily represent the official position or policies of the Food and Nutrition Service or the U.S. Department of Agriculture.

Printed in the United States of America

Available from:
Committee on National Statistics
National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

PANEL ON QUALITY CONTROL OF FAMILY ASSISTANCE PROGRAMS

JOHN NETER (Chair), College of Business Administration, University of Georgia
HENRY J. AARON, The Brookings Institution
MARY JO BANE, John F. Kennedy School of Government, Harvard University
LEON GILFORD, COBRO Corporation, Wheaton, Maryland
LEON GINSBERG, College of Social Work, University of South Carolina
DANIEL G. HORVITZ, Research Triangle Institute
GERALD J. LIEBERMAN, Department of Statistics and Operations Research, Stanford University
CARRIE MENKEL-MEADOW, Law School, University of California, Los Angeles
HELEN B. O'BANNON, University of Pennsylvania
S. JAMES PRESS, Department of Statistics, University of California, Riverside
SAMMIE LYNN PUETT, University of Tennessee
BERNARD J. STUMBRAS, Institute for Research on Poverty, University of Wisconsin
DAVID B. SWOAP, Franchetti & Swoap, San Francisco, California
JUDITH M. TANUR, Department of Sociology, State University of New York at Stony Brook

DENNIS P. AFFHOLTER, Study Director
FREDRICA D. KRAMER, Associate Study Director
GARY L. CAUSER, Research Associate
JUDITH A. UHLMANN, Research Assistant
ELAINE McGARRAUGH, Research Assistant/Production Editor
MICHAEL L. COHEN, Consultant
TORE DALENIUS, Consultant
RICHARD ROYALL, Consultant
WRAY SMITH, Consultant
PATRICE M. SNOWDEN, Administrative Secretary
DEBORAH REISCHMAN, Administrative Secretary

COMMITTEE ON NATIONAL STATISTICS

Contents

Preface

Controversy attends the quality control (QC) systems in the Food Stamp, Medicaid, and Aid to Families With Dependent Children Programs. That controversy, as well as the depths of feelings of all who have made their interests known to the panel, have made our task challenging and meaningful. It has been my pleasure to chair the panel appointed by the National Research Council to examine the quality control systems in question.

The panel would like to acknowledge all of the people who have given their assistance to, and provided support for, the panel's work. Many people offered to the panel their time, opinions, analyses, and insights into quality control. Individuals in public service and private firms, representing particular interests or their personal views, made contributions to our work. Many people throughout federal and state welfare agencies and various interest groups provided much help. Unfortunately, there are too many to be acknowledged here individually. The panel is most grateful to all of these people for their contributions.

The panel itself consists of individuals with varied experiences and views, some forged in public service and others in the disciplines of statistics and quality control. These views created intensive and divergent debates. In the end, the report that the panel has produced expresses a basic and fundamental consensus. David Swoap's disagreement and ultimate dissent on one fundamental issue—that of the proper scope of quality control in the

family assistance programs—helped to hone the panel's majority view on this basic issue.

The panel is greatly indebted to its staff for their effective work. We are particularly indebted to Dennis Affholter and Fredrica Kramer, the panel's study director and associate study director, respectively. They managed to organize and prepare a report that carefully reflects the panel's thinking and views, while making their own significant intellectual contributions. They pushed for needed clarifications when divergent views existed among the panel members, and they worked to accommodate those views within the context of the panel's consensus.

The panel also would like to express special appreciation to Gary Causer, research associate, for his efforts in describing modern thinking about quality, and Judy Uhlmann and Elaine McGarraugh, research assistants, for their substantial efforts in organizing the basic data for the analysis. Deborah Reischman and Patrice Snowden served as administrative secretaries to the panel. We could not have devoted our time and energy to the difficult matters at hand except for the smooth logistics that characterized all of our meetings.

We are also indebted to the panel consultants who assisted our work by developing working papers on a variety of subjects—Michael Cohen, Tore Dalenius, Richard Royall, and Wray Smith. Their papers were of great help to the panel.

The panel appreciates the assistance and support given it by staff officers of the National Research Council. David Goslin, executive director of the Commission on the Behavioral and Social Sciences and Education, provided the institutional backing to expedite every step of the process, from proposal development to the review and release of the panel's report. Eugenia Grohman, associate director for reports of the Commission, effectively managed editorial work on the report, and also provided the detailed follow-up required to meet the panel's tight schedule. Edwin Goldfield, director of the Committee on National Statistics, and Miron Straf, research director of the Committee, helped in establishing the panel and supporting its work.

Reviewers of this report from the Commission and from the Committee also are to be commended. Not only did their reviews substantially improve the panel's report, but they shouldered their review responsibilities under the same tight time schedule that the

Summary and Recommendations

SUMMARY

Background

In 1981 the Food and Nutrition Service of the U.S. Department of Agriculture began imposing financial penalties for states with excessive overpayment error rates in the implementation of the Food Stamp Program. Fourteen states were initially notified of sanction liabilities for fiscal 1981, totalling $29 million; for 1985, 45 states and 3 other jurisdictions operating the Food Stamp Program were notified of sanction liabilities that total $201 million. Aid to Families with Dependent Children and Medicaid—the two other major family assistance programs—also levy financial penalties for excessive overpayments. Observers estimate that by 1989 the total financial liabilities of the states for all three programs will be at least $3 billion, if the current systems for levying penalties remain in place.

Almost all of the states have appealed the sanctions through the administrative process, and several states are disputing Food Stamp liabilities for fiscal 1981 through the courts. In the face of accelerating confrontation and litigation, Congress in the Food

Security Act of 1985 instituted a temporary moratorium on the collection of sanctions, and it mandated studies of the system through which the sanctions are determined, the quality control system.

The Panel's Charge

Congress directed the Secretary of Agriculture to study the quality control (QC) system used in the Food Stamp Program and to contract with the National Academy of Sciences for an independent study. In separate legislation, Congress also directed the Secretary of Health and Human Services to study the QC systems used for the two other programs—Aid to Families With Dependent Children (AFDC) and Medicaid—and to contract with the Academy for an independent study of those QC systems. The Panel on Quality Control of Family Assistance Programs was set up under the Committee on National Statistics to carry out these studies.

The separate congressional mandates have resulted in separate reporting schedules. This first report of the panel focuses primarily on the QC system for the Food Stamp Program. The panel's second report will cover the quality control systems for the AFDC and Medicaid Programs. The panel notes that in all states the state agency that determines eligibility and benefits for the Food Stamp Program also determines AFDC eligibility and benefits; in many states, the same agency also determines eligibility for Medicaid. Consequently, the panel's second report, which will consider program operations for AFDC and Medicaid, may also have material that is relevant to operations of the Food Stamp Program.

The congressional mandates contain two charges: (1) to study "how best to operate" the quality control systems "in order to obtain information . . . to improve the quality of administration" of these three programs, and (2) to provide a reasonable basis upon which to withhold federal funding from states with "excessive levels of erroneous payments." The panel interprets its charges broadly: improving the quality of administration must be linked

to the achievement of program objectives; hence, maintaining financial accountability must also serve those ends.*

On the basis of its analysis, presented in this report, the panel recommends a new system to replace the current quality system for the Food Stamp Program.

Family Assistance Programs, Quality Control, and the Current Controversy

The Food and Nutrition Service (FNS) in the U.S. Department of Agriculture (USDA) is responsible for the Food Stamp Program, for which federal expenditures will be about $12 billion in fiscal 1987. The Office of Family Assistance (OFA) in the U.S. Department of Health and Human Services (HHS) is responsible for AFDC. Fiscal 1987 federal expenditures for AFDC will be $8-10 billion. The Health Care Financing Administration (HCFA) in HHS is responsible for Medicaid. In fiscal 1987 federal expenditures for Medicaid will be about $27 billion. Together these programs account for just over 4 percent of the federal budget, but about 40 percent of federal spending for benefit programs to assist poor people.

All three family assistance programs are administered by the states, either directly or through local agencies, often by the same agency. In all three programs, state and federal governments share the costs of administration. The federal government pays all of the benefit costs (food stamp issuances) in the Food Stamp Program, while states share in the benefit costs (payments) of both AFDC and Medicaid. States exercise a variety of policy prerogatives in all three programs, although the prerogatives are much more circumscribed under the Food Stamp Program. The exercise of these prerogatives creates intergovernmental conflict, since each jurisdiction has its own interests in policy and program administration.

Since the imposition of financial penalities in fiscal 1981, every state but one has been found liable for some penalties under the QC provisions of the programs under study. The imposition of these penalities has heightened the current controversy and led

*A panel member dissents from this interpretation and from several points in the panel's analysis and four of its recommendations; see Appendix A.

to the congressional mandates for the NAS studies. But that controversy also exposes the divergence of interests among those concerned with the administration of family assistance programs and the appropriate use of quality control systems in them.

To the federal government, the penalties represent one way of controlling program costs due to fraud and abuse and other so-called "misspent" funds. To some representatives of client interests the quality control systems raise concerns about denial of needed benefits. To the states, the systems represent a combination of these concerns, perhaps dominated by their financial liabilities. The panel finds that each of these perspectives represents policy objectives that "quality control" and management improvement systems can and should address.

Under federal law, state agencies operate quality control systems for all three programs. Since its inception in 1964, the quality control system for AFDC has focused consistently on overpayment errors, although underpayment (and certain nonpayment) errors are measured. A major change in 1970 established the QC system now used in AFDC; the AFDC QC system provided the model for the Food Stamp and Medicaid QC systems. The quality control systems in all three programs focus primarily on one type of payment inaccuracy, overpayments to eligible households or individuals and payments to ineligible households or individuals.

Since the 1960s, concerns have been raised in Congress about growing caseloads and costs, beginning with AFDC but extending to Food Stamps and Medicaid, and those concerns became linked with concerns about ineligibility and overpayments and about fraud and abuse. But caseload growth reflected the influence of many factors, including social, demographic, cultural, and economic factors and policy changes. Particularly in the Food Stamp Program, some caseload growth resulted from explicit policy choices that expanded the eligible population and also increased participation rates among eligible households.

The current QC systems do not measure the achievement of program objectives other than payment accuracy. For the Food Stamp Program, such other objectives include improved nutritional status, effective coverage of the eligible population, and the timely provision of benefits to eligible applicants.

Efforts to hold state agencies accountable to QC-based performance standards can be traced at least to 1973, although the

imposition of financial sanctions and disallowances on state agencies began only for fiscal 1981. Currently, the "official payment error rates" on which state-level performance standards are based are set at 5 percent for the Food Stamp Program and 3 percent for AFDC and Medicaid. Official fiscal 1984 estimates of the national payment error rates are 8.6 percent for Food Stamps, 6.0 percent for AFDC, and 2.7 percent for Medicaid.

The Food Stamp Program

Sampling and Reviews

The QC system for the Food Stamp Program has been characterized as a two-tiered system of state and federal government activities. State Food Stamp officials sample monthly from their active cases and "negative case actions" (terminations or denials of benefits), and state QC reviewers verify the accuracy of eligibility and benefit determinations in the sampled cases. Federal officials in regional offices subsample the completed QC reviews and re-review those case findings.

The state sample designs and sampling procedures must be approved by FNS and must meet prescribed standards for probability samples with minimum specified sample sizes, which vary with the size of a state's caseload. Federal law and regulations provide the standards for QC review findings. Before 1982, however, state plans that were approved by FNS provided standards for QC review findings. Federal-state disagreements on case findings are resolved through arbitration, first in regional FNS offices, then, if needed, in the national office.

Estimates of Error Rates

FNS estimates the official payment error rate, which includes the value of payments to ineligible recipients and the value of overpayments to eligible recipients. This payment error rate is based on the federal sample review data adjusted by data from the state sample reviews to improve accuracy. FNS also estimates an error rate for underpayments, again using this combined approach. In the state QC samples alone, FNS estimates an error rate for negative case actions—terminations or denials of benefits to eligible recipients.

However, the last two error rates, for underpayments and for negative case actions, are not included in the official payment error rate and therefore do not contribute to financial penalties.

For the period from fiscal 1980 through fiscal 1984, the reported annually national payment error rates show some decline: 9.51, 9.90, 9.55, 8.32, and 8.64 percent, respectively. The reported national underpayment error rates have hovered just under 2.5 percent for those 5 years. FNS reported negative case action error rates only through fiscal 1983, and these show a decline from about 4.0 percent in fiscal 1980 to 2.9 percent for fiscal 1983.

Sanctions and Rewards

If a state's official payment error rate is larger than the established performance standard, the state is notified of its potential liability for sanction. For each of the first three percentage points above the performance standard of 5 percent (5.00+ to 6.00 percent, 6.00+ to 7.00 percent, and 7.00+ to 8.00 percent), the state loses 5 percent of its federal funding for administration. For error rates above 8 percent, each additional percentage point costs the state 10 percent of its federal share for administrative costs. A state with an official error rate between 7 and 8 percent would incur a 15 percent loss in its federal administrative funding, but a state between 8 and 9 percent would incur a 25 percent loss. Because of this so-called step function, a small change in a state's error rate, say from 7.99 to 8.01 percent, can make a big difference in federal penalties.

States can also receive a performance-based bonus—60 percent instead of 50 percent of administrative costs from FNS—if the sum of a state's official payment error rate and its underpayment error rate meets or falls below the standard set for the payment error rate, *and* if the state's negative case action error rate also meets a standard set by the Secretary of Agriculture. Currently, the standard for the negative case action error rate is simply the national estimate for that error rate.

Waivers and Appeals

A state may request a waiver of its sanction liability for "good cause" as specified in the Food Stamp Act. Whether or not it seeks such a waiver, a state may appeal the liability established by FNS,

to a Grants Appeals Board that consists of USDA headquarters staff who do not work in FNS. Following the Board's decision, a state may file suit in a U.S. District Court.

Evaluation of QC System

Strengths and Weaknesses in the Food Stamp QC System

The panel finds that the Food Stamp QC system has certain strengths, which should not be lost in efforts to improve it. The current system produces and reports annual estimates of state payment error rates that can be tracked over time; such data were unavailable before the system was implemented. Producing those estimates is a complex job and for the most part it is done on schedule. Compared to some public programs with similar missions, the QC systems used in the family assistance programs are relatively sophisticated. The administrative resources required for the QC work represent a resource upon which to build. And the pervasive emphasis on doing good work throughout the Food Stamp Program, reinforced by the current QC system, ought not to be lost as changes are made.

The Food Stamp QC system primarily serves the federal government's interest in monitoring one kind of payment accuracy, overpayments. It offers little to state and local program managers in support of continued improvement in administration. The QC emphasis on overpayments may also lead to the neglect or detriment of important program objectives other than overpayment accuracy. Some critics of the QC system, for instance, claim that "case churning" (recipients or eligible applicants losing benefits for a short time, for procedural reasons) is a direct result of eligibility determination procedures designed to protect against the possibility of overpayment errors. The current QC system also fails to address the problem of state agency responses to sanction liabilities that may directly conflict with the achievement of other program objectives.

The panel finds that federal monitoring of states' performance, and the connection between measured performance and financial consequences to promote improved performance, can be improved. The current performance standard ignores state-to-state differences in caseload mix that may be beyond the control of state administrators and that may influence measured performance. The

panel also finds problems with the regression-adjusted estimator, now used to determine official payment error rates. In addition, QC sample designs cause the reliability of state-level estimates to vary from state to state. As a result, the risk of oversanction or undersanction due to sampling error varies from state to state. Furthermore, FNS violates sound statistical reporting practice by failing to routinely calculate and publish estimates of sampling error. Finally, the current sanction and reward structure provides inconsistent incentives for improving performance because of the step function for sanctions and the single bonus awarded at one performance threshold.

The Concept of Quality Improvement

Quality improvement experts take a broad view of quality and its management, and the panel finds that view relevant to the Food Stamp Program. Quality improvement involves a broad range of activities designed to assist in the achievement of program objectives. Systems for quality improvement include three kinds of activities: process control, process improvement, and service or program improvement. Process control is concerned with maintaining performance at the limits of capability of a process as designed. Process improvement examines possibilities for process redesign to improve performance capability. Service or program improvement considers questions of program redesign to better satisfy existing objectives or to serve different program objectives.

Quality improvement systems require both routine and special studies of performance, measured against a broad array of program and policy objectives. The panel finds that certain features of comprehensive quality improvement systems are relevant to family assistance programs: quality improvement is driven by the analysis of performance data that measure the achievement of program objectives; quality improvement requires timely feedback, so that the responsible officials have the needed access to performance information; and quality improvement requires action that is consciously designed to improve performance and is monitored and assessed to ensure its effectiveness. These features apply at all levels in a service operation, from the point of delivery to the central office, although the studies, measures, and tools most useful for one level may differ from those most useful at or for another.

The panel finds that the Food Stamp QC system lacks many of the elements of a comprehensive quality improvement system. For process control and for planning process improvements, state and local managers require more performance data, at lower levels of aggregation, than are now provided through the QC system. Furthermore, the federal government also has an interest in state and local improvements: it is only at the point of service delivery that the federal government's program objectives may be realized. The current QC system does not serve that interest well.

RECOMMENDATIONS

The panel recommends a number of changes that together will render the QC system a more effective tool for quality improvement. This section summarizes the panel's recommendations for a new quality improvement system. The panel's detailed recommendations follow this section, and a full discussion of the recommendations appears in Chapter 6.

Basic Structure for Quality Improvement

The so-called two-tiered "quality control" system hinders quality improvement and unnecessarily fosters conflicts among state and federal participants. Federal monitoring activities for payment accuracy should be separated from other state and federal quality improvement activities (recommendation 1).*

Improving quality requires that performance measures reflect the achievement of program objectives. If the measures do not represent all important objectives, the quality improvement system may promote some to the neglect of others. Performance measures to be used in quality improvement systems for the Food Stamp Program should reflect broadly the intent expressed in the Food Stamp Act (recommendation 2).*

Building State Quality Improvement Systems

States should develop their own quality improvement systems but with federal assistance and guidance. In order for that development to occur, Congress should mandate a 5-year maintenance

*A panel member dissents; see Appendix A.

of fiscal effort for state QC activities to protect current resources from competitive budget pressures. States should redirect the use of those resources to building more effective state and local quality improvement systems (recommendation 3). Federal and state agencies should work cooperatively, jointly, to expand their quality improvement activities, with the federal government taking a lead role for the provision of various forms of assistance (recommendation 6). Federal and state agencies also should use this 5-year period to evaluate the states' emerging quality improvement systems to determine how much to spend on those systems (recommendation 4) and to develop and refine other performance measures to which states could be held accountable (recommendation 5).

Federal Monitoring and Holding States Accountable

The panel proposes that Congress and FNS also adapt a broad quality improvement perspective to the task of monitoring state performance. Quality improvement systems should be designed to identify and attend to systemic problems in performance and to assign the responsibility for improving performance to the appropriate levels and managers within the Food Stamp Program's service delivery system, some of which may be national and some of which may be state specific.

Setting Performance Standards with Financial Consequences

The panel precedes recommendations 7 through 11 with a caveat. The panel cannot set precise performance standards or magnitudes of sanctions and rewards tied to measures of achievement against those standards: they are policy choices; the empirical evidence necessary to frame them precisely is not available; and they should change over time to reflect changes in capability.

The panel offers these recommendations as a whole, however, to make clear its conclusions about the direction and relative magnitude of these choices, based on assumptions about capability and in order to promote desired changes in the current system.

In the classic view, a quality control system for a program operating within attainable levels of performance sanctions and rewards only those performers at the extremes of performance. In

the face of demonstrated improvements in performance from fiscal 1980 through fiscal 1984, a QC system that finds three-fourths or more of its operating units to be performing poorly has one of two problems, or perhaps both: either the standards by which performance is judged are too stringent, or the QC system has failed to put in place the necessary quality improvement systems to achieve the performance levels indicated by the standards within the prescribed time period. In light of the evidence on improved state performance, a Food Stamp sanction and reward mechanism should identify only state agencies whose performance is unusually good or bad.

Within this framework, the panel sets forth the key elements for a new quality system that would, in their related effects, set direction and relative magnitudes of sanctions and rewards:

- performance standards should change over time to reflect changing performance capabilities;
- when states are operating at attainable levels of performance, thresholds for sanctions and rewards should be set so that only states at the extremes of good or bad performance are subject to rewards or sanctions;
- standards should reflect state-to-state differences in caseload composition;
- the system should include rewards as well as sanctions that are calculated against benefit costs;
- measures of performance to which financial consequences are attached should include criteria in addition to overpayments;
- federally approved state plans should provide the measurement standards against which to assess performance.

For this new system for setting sanctions and rewards, the panel recommends particular federal monitoring responsibilities. They include monitoring the consequences of sanctions within states that fall below the desired standard and special monitoring of and technical assistance for states with serious performance problems.

The panel recommends that FNS set performance standards to reflect state-to-state differences in caseload composition with respect to groups of recipients that may be differentially prone to error (recommendation 7). The panel also recommends that rewards as well as sanctions be established as a proportion of the

benefit dollars in error above or below established performance standards (recommendation 8).* With this change in determining the magnitude of sanctions, the panel notes that the proportion should be established such that a state's liability should not be substantially larger than that which would have been incurred under the current system.

Rewards for good performance or penalties for bad performance should be based on reliable measures of the achievement of desired objectives. The panel finds that only data on payment or issuance inaccuracy are now widely available as reliable measures of program performance. Measures of underpayment errors and erroneous nonpayments, as well as overpayments and payments to ineligible recipients, should be included (recommendation 10).* The panel recognizes that policy makers may differentially value different kinds of payment errors. Such preferences can be represented by attaching numerical weights that correspond to the relative values of the different components of error, but the panel believes that each of the four types of error is important and that the weights should reflect this.

The panel recommends that FNS resume approval of state plans in order to establish standards against which to hold states accountable. Only payment errors associated with violations of the state plans should be counted for the purposes of levying sanctions and awarding performance bonuses (recommendation 11). Federal reviewers should continue to collect information on other sources or causes of error, however, so that systemic problems of performance can be identified and corrected (recommendation 11).

The panel wishes to ensure both that appropriate corrective action is taken by poorly performing states and that state agencies do not take measures in order to pay QC penalties that hinder program administration or whose effect will directly conflict with other program objectives. The panel therefore recommends that FNS carefully monitor the consequences of sanction to ensure that corrective action is taken. It further recommends that such monitoring assess the consequences of sanction payments to state agencies and take appropriate steps to minimize the dysfunctional effects created by the burdens of sanctions (recommendation 9).

*A panel member dissents; see Appendix A.

The panel also recommends the establishment of a special Secretary's discretionary fund, for the provision of aggressive technical assistance to and on-site monitoring of states with serious performance problems (recommendation 9). The panel believes that under the new monitoring system, fewer states would be subject to sanctions and therefore that such careful monitoring of the consequences of sanctions is feasible.

Sampling Error, Resolution of Differences, and Other Data Reporting

The panel recommends that FNS/USDA establish procedures for state reviews of findings of erroneous issuances by federal reviewers and that those procedures include steps for the resolution of state-federal differences (recommendation 12). The panel further recommends that FNS design samples for federal estimates of issuance inaccuracy, for which state agencies will be held accountable, to achieve equal precision in total issuance inaccuracy rates across states (recommendation 14). The panel also recommends that FNS routinely report estimates of the standard errors of all performance estimates that it publishes (recommendaton 13).

Resolving the Problems of Past QC Sanctions

The panel recognizes that its recommendations for the development of quality improvement systems cannot feasibly be applied retroactively for resolving the pending QC-based sanctions under the Food Stamp Program. The panel therefore makes separate recommendations on how FNS can clear up the backlog.

The recommendations on disputed sanctions are recommendations of expedience. FNS should use the lower bound of an interval estimate of payment error rates, for the years from fiscal 1981 through fiscal 1987, to determine whether a state should be sanctioned (recommendation 15). Such a measure would mitigate some of the potential unfairness built into current sanctioning procedures. However, FNS should use a point estimate for establishing the magnitude of a sanction once the sanction decision has been made in order to use the best estimate of the true error rate (recommendation 16). The panel has examined the regression-adjusted estimator currently used for estimating official error rates and it finds this estimator lacking for a variety of reasons. The panel is

exploring alternative estimators and will report its critique as well as proposed solution(s) in its second report.

DETAILED RECOMMENDATIONS

Recommendation 1: The FNS/USDA sampling and measurement activities for monitoring issuance accuracy in the Food Stamp Program should be made independent of state-level sampling and measurement activities for quality control, quality measurement, and quality improvement.*

Recommendation 2: Measures of quality should be set against broad program objectives, beginning with the language of intent in the Food Stamp Act. Federal performance measures should begin with payment accuracy but should be added to over time and as data from monitoring and evaluation activities permit greater breadth and specificity.*

Building State Quality Improvement Systems

Recommendation 3: Congress should mandate that state welfare agencies maintain current levels of effort in their quality control operations for 5 years and that state quality control resources be redirected toward the development of quality improvement systems.

Recommendation 4: Congress, FNS, and the states should each set in place evaluations to determine appropriate levels of support for quality improvement at the end of the 5-year maintenance-of-effort period.

Recommendation 5: Congress and FNS should mount evaluation efforts during the 5-year maintenance-of-effort period to develop objective standards for expanding the scope and specificity of quality measurements to which states can be held accountable.

Recommendation 6: The federal government should help states to develop quality improvement systems. It

*A panel member dissents; see Appendix A.

should provide technical expertise and financial and regulatory support and should disseminate information about best practices. FNS should support activities that include multistate, randomized experimentation with innovative practices and interagency cross-program efforts. State agencies should develop quality improvement strategies and share information with other states with similar needs or interests.

Improving the Federal Monitoring System

Recommendation 7: FNS should set performance standards to reflect state-to-state differences in caseload composition with respect to groups that are differentially prone to error. FNS should initially examine the possibility of using at least three groups as likely candidates for such corrections at this time: all households with reported earnings, all nonearning households with one or more elderly recipients, and all other households.

Recommendation 8: Sanctions and rewards should be calculated as a proportion of the benefit dollars in error above or below the performance standards. Because the base for calculating sanctions will be different from that under current procedures, the sanction proportion will need to be revised.*

Recommendation 9: FNS should monitor the consequences of sanctions to ensure that corrective action is undertaken. Congress and FNS should take steps to minimize potentially harmful effects on program administration and program objectives, for example, in the structuring of payment schedules for sanctions. In addition, as a non-routine response to major deficiency findings, USDA should establish a special Secretary's discretionary fund for aggressive technical assistance and on-site monitoring when a state Food Stamp agency has been determined to have serious performance problems.

Recommendation 10: The USDA should use estimates of issuance inaccuracy in the Food Stamp Program for the

*A panel member dissents; see Appendix A.

purpose of linking financial consequences to state agencies' performance, based on the absolute dollar value of each of four distinct types of errors: issuances to ineligible recipients, overissuances to eligible recipients, underissuances to eligible recipients, and issuances that should have been made to applicants or recipients who were wrongly denied or terminated. Issuance inaccuracy *rate* estimates should be made for each type of error by dividing the estimates of the total dollar value of errors for each component by the total value of issuances actually made. The performance estimate upon which to make sanction and reward decisions should be based on a weighted average of the four estimates of issuance inaccuracy rates. Although the panel believes that the determination of weights is a policy decision, it believes that each of the four types of error is important and that the weights should reflect this.*

Recommendation 11: FNS should approve state plans for the Food Stamp Program, and the plans should constitute the criteria against which states are held financially accountable. Federal reviewers should identify those sources of error that are associated with violations of the state plan. Federal reviewers should continue to report on the sources or types of errors that are now measured through QC systems.

Recommendation 12: Under the new system that the panel recommends, FNS/USDA should establish a process to accommodate state differences with federal findings of issuance inaccuracies and should establish procedures for the timely resolution of differences through the national office with respect to those findings.

Recommendation 13: FNS should routinely report estimates of sampling errors for all performance estimates based on sample data.

Recommendation 14: FNS should design samples for the purpose of estimating issuance inaccuracy rates to achieve equal precision across states.

---

*A panel member dissents; see Appendix A.

Resolving Past Sanction Liabilities

Recommendation 15: For the purpose of determining whether a state agency should be sanctioned under the current QC system for fiscal 1981 through fiscal 1987, FNS/USDA should use a lower bound for the yet-to-be-recommended estimator. The width of the interval between the point estimate and the lower bound should be the same for all states. The width should be based on the estimated sampling error for the state agency with the largest sampling error in its payment error rate estimate. If the lower bound is at or below the sanction threshold, then a state would not be sanctioned. If the lower bound is above the threshold, then a state would be sanctioned.

Recommendation 16: If the decision to sanction for years fiscal 1981 through fiscal 1987 has been made, the yet-to-be-recommended point estimate of payment error should be used to establish the magnitude of the sanction to be levied.

1
The Current Controversy

BACKGROUND

Since the imposition of financial penalties in fiscal 1981, every state but one has been found liable for some penalties under the quality control (QC) provisions of the three major family assistance programs: Aid to Families With Dependent Children (AFDC), Food Stamps, and Medicaid. Notifications of QC-based liabilities for fiscal 1985 for AFDC have been sent to 47 jurisdictions (43 states, the District of Columbia, Guam, Puerto Rico, and the Virgin Islands); 48 jurisdictions (45 states, the District of Columbia, Guam, and the Virgin Islands) have been notified of liabilities for fiscal 1985 Food Stamp penalties. At stake are hundreds of millions of dollars per year: by some estimates, the total will be at least $3 billion if the current system remains in place through 1989. For two of the largest state programs, in California and New York, liabilities already assigned through fiscal 1984 total $141 million and $123 million, respectively.

Several states have jointly sued the U.S. Department of Agriculture (USDA) over QC-based sanctions in the Food Stamp Program for fiscal 1981. Many states also have administrative appeals for fiscal 1981 pending within the U.S. Department of Health and

Human Services (HHS), concerning QC-based sanctions for both Medicaid and AFDC. If the state welfare agencies lose their appeals in the Department, more legal actions can be anticipated.

In the face of heightened tension and legal battles, the 99th Congress intervened. It placed a temporary moratorium on the collection of sanctions and disallowances based on the information reported from the quality control systems; it also directed the secretaries of USDA and HHS to conduct studies of the QC systems and to contract with the National Academy of Sciences (NAS) for "concurrent independent" study of those systems.

THE PANEL CHARGE

The congressional mandate is twofold, reflecting the goals of making QC systems serve state management purposes and of holding states accountable for payment accuracy. The Food Security Act of 1985, signed in December 1985, section 1538 (a)(1), pertaining to quality control of the Food Stamp Program, states:

> (B) The study shall—
>
> (i) examine how best to operate such system in order to obtain information that will allow the State agencies to improve the quality of administration; and (ii) provide reasonable data on the basis of which Federal funding may be withheld for State agencies with excessive levels of erroneous payments.

Similarly, the Consolidated Budget Reconciliation Act of 1985, signed in April 1986, section 12301 (a)(1), pertaining to the AFDC and Medicaid programs, states:

> The study shall examine how best to operate such systems in order to obtain information which will allow program managers to improve the quality of administration, and provide reasonable data on the basis of which Federal funding may be withheld for States with excessive levels of erroneous payments.

The panel interprets its charges broadly: improving the quality of administration must be linked to the achievement of pro-

gram objectives; hence, maintaining financial accountability must too serve those ends.*

THE ROLE OF QUALITY CONTROL

Despite the current controversy, few people oppose the idea of "quality control" in public programs. Quality control is of particular interest for programs that provide money or in-kind services. The Food Stamp, AFDC, and Medicaid programs represent 40 percent of such federal benefit spending, and moreover, that spending represents money or services defined as necessary to provide minimum cash, food, and medical care to people in need. Interest in quality control in these programs is heightened because they involve a web of federal, state, and local administrative relationships, which complicates their administration and generates a set of potentially conflicting interests. The current dispute accentuates the federal government's interest in payment accuracy, but both state and federal governments have obvious interests in developing quality control systems, as do welfare advocates. Those interests can be expressed in at least three ways.

First, public officials at all governmental levels are accountable for the expenditure of funds; governments that delegate spending authority seek to ensure that the authority is properly executed. The intergovernmental delegation of spending authority for many family assistance programs makes public accountability particularly complicated and diffuse. Second, measures of program performance can provide useful information for the policy process and assist in making policy corrections. Federal and state policy makers and others involved in the policy process all have an interest in identifying and understanding program achievements and limitations. Third, monitoring program operations through QC systems can promote administrative efficiency; efficiency is of greatest interest to those directly responsible for program administration at the state and local level, but it may also result in cost savings at the federal level.

*A panel member dissents from this interpretation and from several points in the panel's analysis and four of its recommendations; see Appendix A.

THE PANEL STUDY

This report deals primarily with quality control in the Food Stamp Program. It is the first report of the Panel on Quality Control of Family Assistance Programs. The panel has reviewed relevant literature, and conducted field work. It has amassed an extensive library of materials relevant to quality control, over 500 items currently. The panel has examined detailed QC data provided by the Food and Nutrition Service (FNS), the arm of USDA that administers the Food Stamp Program at the federal level. It has reviewed the quality control and quality improvement literature, and it has reviewed relevant legislative history to better understand the context in which the QC systems have developed.

The panel has reviewed the major QC evaluation research efforts that have been undertaken by the U.S. General Accounting Office, contract research firms, and others, from the formative days of QC to the present. And it has received dozens of detailed analyses from states, public interest organizations, and others with direct interest in the QC controversy and the panel's deliberations. The panel held six full panel meetings between July 1986 and May 1987; for two full days, it heard testimony directly from several of the interested groups, as well as from representatives of the interested federal agencies. It has interviewed former and current federal officials with knowledge of federal QC history. The panel's staff also interviewed officials engaged in quality control in other public programs and in private industry.

Members of the panel and staff visited local, state, and federal welfare programs to observe local intake and eligibility procedures, state QC review processes, and the federal re-review process. During these site visits, panel members and staff interviewed line workers, line supervisors, welfare administrators, QC review staff at the state and federal level, and those responsible for corrective action planning, as well as technical and policy staff at the state and federal level.

The visits covered a range of geographic regions and program types. Reviews were conducted in local or state programs in California, Georgia, Maryland, New Jersey, New York, North Carolina, South Carolina, Tennessee, and Wisconsin, and with officials from FNS, the Office of Family Assistance and the Health Care Financing Administration regional offices in Trenton, Philadelphia, Atlanta, and Sacramento. These lengthy visits served to introduce

the panel to the QC system at federal, state, and local levels. The panel also has undertaken an intensive analysis of the current statistical techniques used for calculating official federal error rates and for imposing sanctions.

The panel's work is not done, and under other circumstances it would have chosen to complete its work and release one report that covered all three QC systems. The panel recognizes, however, the force of congressionally set schedules. Those schedules result in contractual obligations to submit a report to USDA and to Congress on the Food Stamp Program several months before a similar report is to be submitted to HHS and to Congress on AFDC and Medicaid. The congressional schedule also required that the panel compress a study schedule of what might otherwise take 2 years or longer into half that time.

The staggered reporting schedule poses other difficulties for the panel. Although there are some differences between the QC systems in the three programs, they share much in common. Furthermore, at state and local levels, administration of these programs is not so easily or so clearly distinguished as are the federal programs themselves. This report suggests a basic strategy for quality improvement with specific recommendations explicitly for the Food Stamp Program—though many of them are likely to be applicable to the other programs as well. The panel will consider issues of program interrelationships between the Food Stamp, AFDC, and Medicaid Programs in its second report. The panel's recommendations are in Chapter 6. They are preceded by the panel's findings and analysis. Chapter 2 describes the nation's basic family assistance programs, the three programs under study, their trends over time, and the history of QC systems in those programs. Chapter 3 describes the current QC system for the Food Stamp Program, Chapter 4 addresses the basic critiques of the current system, and Chapter 5 offers lessons from quality control experts.

2
Quality Control in Family Assistance Programs

The AFDC, Food Stamp, and Medicaid Programs serve related purposes and overlapping populations, and their quality control systems are similar. The Food Stamp and Medicaid quality control systems are modeled on the one developed for the AFDC Program in the early 1970s. Each focuses on measuring and controlling the cost of ineligibility and abuse. This chapter describes the context in which the current QC systems developed and in which they now operate: how the three programs fit into the larger framework of federal family assistance programs, how the programs work, how they have changed over time, and finally, how those changes relate to the development of quality control.

FEDERAL BENEFIT PROGRAMS

The combined total of federal spending for Food Stamps, AFDC, and Medicaid was more than $43 billion in fiscal 1985; estimates for fiscal 1987 by the Office of Management and Budget and the Congressional Budget Office are $47 and $49 billion, respectively. In fiscal 1985, the average monthly caseload for AFDC was 10.9 million individuals, food stamps were issued to an average of nearly 20 million individuals monthly, and nearly 22 million people per

TABLE 1 Federal Spending for Human Resource Programs 1985

Program Category	Expenditures (in $ billions)
Social Security	188.6
Income Security[a] (includes AFDC and Food Stamps)	128.2
Medicare	65.8
Health[b] (includes Medicaid)	33.5
Education, Training, Employment, and Social Services[c]	29.3
Veterans Benefits and Services	26.4
Total	471.9

[a]Includes general retirement and disability insurance, federal employee retirement and disability, unemployment compensation, housing assistance, food and nutrition assistance, and other income security.
[b]Includes health care services, health research, education and training of health care work force, and consumer and occupational health and safety.
[c]Includes elementary, secondary, and vocational education; higher education; research and general educational aids; training and employment; other labor services; and social services.

Source: Data from Office of Management and Budget (1986c).

month received services from Medicaid. Because the recipient populations overlap, many of the individuals counted in each of the caseloads may also be counted in the other caseloads.

Despite these totals of cost and caseload, the three programs represent only a little more than 4 percent of the whole federal budget and only a small part of federal expenditures for benefit programs. They accounted for just over 9 percent of all human resources spending in fiscal 1985, while spending for old age, survivors, and disability insurance benefits (commonly known as Social Security) accounted for 40 percent (see Table 1).

The three family assistance programs, however, account for about 40 percent of spending for benefit programs specifically targeted for people in need, which include housing assistance, Supplemental Security Income (SSI), and many other, smaller programs.

A U.S. General Accounting Office (GAO) (1985) survey identified obligations for 150 "benefit programs" that accounted for almost one-half of the federal budget. Although the study surveyed expenditures for fiscal 1983, relationships among programs have remained much the same. Of the programs surveyed, 95 were "noncontributory" or "needs based": that is, recipients are not required to make any contributions to the benefit pool because they qualify on the basis of need. The remaining 55 were "insurance based," such as Social Security, Medicare, unemployment compensation, and civilian and military pensions. The 95 needs-based programs accounted for only 18 percent of the $409 billion total expenditures for "benefit programs," or about 8.6 percent of the total federal budget for that year. Figure 1 approximately replicates the GAO construction, using fiscal 1985 budget figures; they illustrate the large portion of total federal spending that benefit programs represent, and the relatively small portion that constitutes the three programs under study.

The five largest needs-based programs, however—Medicaid, Food Stamps, Supplemental Security Income (SSI), AFDC, and Section 8 housing assistance—accounted for 63 percent of the total expenditures for the 95 needs-based programs studied for fiscal 1983. Unlike the other three programs, neither SSI nor Section 8 rent subsidies rely on state administration (although some states do supplement SSI payments). The AFDC, Medicaid, and Food Stamp Programs, therefore, constitute the principal components of the intergovernmental system of needs-based assistance. Table 2 shows federal spending levels for these three programs.

AFDC, MEDICAID, FOOD STAMPS: BASIC CHARACTERISTICS, COMMON THREADS

The AFDC, Food Stamp, and Medicaid Programs have multiple and related objectives, and serve overlapping populations. All three operate in the 50 states, the District of Columbia, Guam, and the Virgin Islands. AFDC and Medicaid also operate in Puerto Rico, and Medicaid operates in American Samoa and the Northern Mariana Islands. Since 1982 Puerto Rico has had its own block grant program to deliver food stamps. All three programs have experienced substantial growth in total program costs in the last two decades, and finally, share some important, common problems.

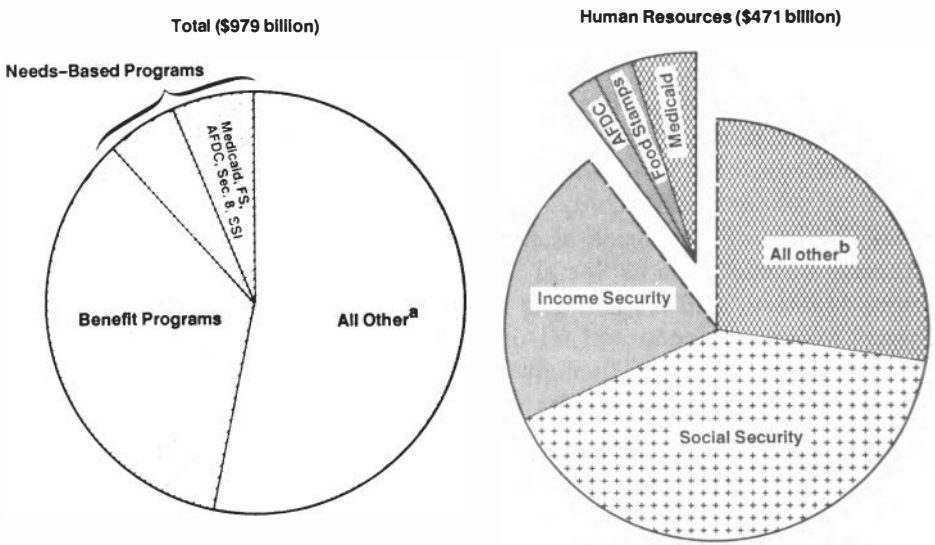


FIGURE 1 Total Federal and Human Resource Budgets, 1985

[a]Includes national defense, physical resources, and other functions.
[b]Includes health, education, training, employment, social services, and veterans benefits and services.

Source: Data from Office of Management and Budget (1986b, 1986c).

Program Populations

AFDC is the principal cash transfer program for needy children and their parents—mostly single mothers. At state option and with certain restrictions, coverage may be extended to families with a second, unemployed, parent—usually a father. Half of the states offer this "unemployed parent" program. Parents, guardians, and children who are over age 16 are entitled to AFDC benefits if they meet income and assets test and fulfill certain conditions. One such condition is the requirement to register for employment and training services unless exempted for specified

TABLE 2 Federal Spending for Family Assistance Programs (in $ billions)

Program	Fiscal 1985 Actual	Fiscal 1987 Estimated	
		OMB[a]	CBO[b]
Medicaid	22.8[c]	26.7	27.3
Food Stamps	11.7[d]	11.8	11.7
AFDC	8.0[c]	8.0	10.1

Note: The 1985 figures for Medicaid and Food Stamps include administration and benefit costs; the figure for AFDC includes benefits only.

[a]Data from Office of Management and Budget (1986a).
[b]Data from Congressional Budget Office, personal communication.
[c]Data from U.S. Department of Health and Human Services (1986).
[d]Data from Office of Management and Budget (1986b).

reasons, which include responsibilities for child or other dependent care, certain incapacities including age, work of at least 30 hours per week, and full-time school attendance for children over 16.

Medicaid is the principal vehicle for providing health care to the nation's poor. It pays health care providers to serve all AFDC recipients, for whom Medicaid eligibility is automatic, and most SSI recipients. In addition, others who meet Medicaid eligibility criteria—a population defined as medically needy, including many in long-term care facilities—may be covered at state option. Medicaid requires that certain health-care services be included in each state's program, permits specific other services to be included at state option, and prohibits the expenditure of federal Medicaid funds for all other health-care services.

The Food Stamp Program provides coupons (food stamps) redeemable for food or food products to people with low income. Since there are few conditions of eligibility other than need, the Food Stamp Program serves virtually anyone who can meet its income, assets, and work tests, including, for instance, working poor people, single individuals, the elderly poor (most of whom

also receive SSI), and AFDC recipients who meet the program's income and assets tests. While it does not serve individuals in long-term care facilities, the program does offer assistance under a variety of conditions to low-income individuals who live in group settings, in certain treatment facilities, and who receive prepared meals.

Program Structures

The federal government sets certain national standards for, and restrictions on, state and local administration of the AFDC, Medicaid, and Food Stamp Programs. Federal policy, starting with AFDC and later with Food Stamps and Medicaid, evolved from a tradition of local control. AFDC was added on to a set of traditionally local activities serving the poor, and the other programs added on to it. All three programs preserve shared federal, state, and local responsibilities, although state prerogatives and financial participation in Food Stamps are more restricted than they are in AFDC and Medicaid. State and local agencies responsible for the administration of AFDC also administer key functions of the other welfare programs. Therefore, the administration of AFDC at the state and local level dominates and largely determines the setting in which the Food Stamp Program is administered.

AFDC

AFDC is authorized under Title IV of the Social Security Act. State agencies administer the program or supervise local administration. There is no national benefit standard or standard of need: states set standards of need, and they establish benefit levels against their own standards; therefore, they may pay less than what they themselves have defined to be minimum standards of need.

The federal government's share of benefit costs is based on a formula that reflects state per capita income; hence, the federal government bears more of the cost in poorer states. The federal share of benefit costs can vary between 50 and 83 percent. For program administrative costs, most are split 50-50 between federal and state governments, and local governments may share some portion of the state costs where AFDC is locally administered under state supervision. Federal sharing of administrative costs is

higher for certain expenditures, including design and implementation of automated claims processing.

The federal government places many restrictions on who may receive AFDC benefits. The law mandates the maximum amount of resources a recipient household may have—$1,000—with certain "exclusions" and "disregards." The law also mandates various recipient safeguards, reporting requirements, and other conditions concerning eligibility, participant compliance, and state administrative responsibilities.

Although the principal benefit of AFDC is cash payments, the program may, and in some cases must, provide for a variety of other services toward the goals of achieving rehabilitation and self-sufficiency. These include, for example, family planning services, employment and training services, certain forms of emergency assistance, and social and supportive services.

Since the federal government does not administer AFDC directly, states and localities must carry out the needed referrals, administrative linkages to other programs, or direct service provision. Thus, often complicated linkages have been attempted to a variety of other social and rehabilitation services, including those administered by the state employment services, the local administrators of the Job Training Partnership Act, a variety of state and local social service agencies, and the state child support enforcement system. Since AFDC also confers categorical eligibility for health care through Medicaid, a separate set of relationships is necessarily created with the state's health care providers.

Medicaid

Medicaid is authorized under Title XIX of the Social Security Act. As with AFDC, state and local agencies operate the program: they certify recipients' eligibility and pay providers. State and local welfare agencies determine eligibility for most Medicaid recipients. All AFDC recipients and most SSI recipients are categorically eligible for Medicaid. In addition, states may extend coverage to individuals who by certain criteria are determined to be medically needy. Medicaid has its own set of intergovernmental relationships deriving from its population in institutions and long-term care facilities and its relationship with health care providers.

The federal government reimburses states for all covered health-care services by a formula that, as with AFDC, is inversely

related to state per capita income, and pays between 50 and 83 percent of benefit costs (except for family planning services, which are 90 percent reimbursed). Administrative costs are split 50-50, with some exceptions that include the development of automated claims processing systems and control of fraud and abuse.

The law requires that Medicaid provide inpatient and outpatient hospital services, laboratory and X-ray services, skilled nursing facility care for people over 21, some home health services, early periodic screening diagnosis and treatment (EPSDT) for people under 21, family planning services, and physicians' services. In addition, states have many prerogatives concerning services that will be covered under their Medicaid programs. Optional services include drugs, services in intermediate care facilities, eyeglasses, and inpatient psychiatric care for people between 21 and 65. States also may set limits on the amount of care in different categories. States must provide certain services if they provide coverage to the medically needy. The Tax Equity and Fiscal Responsibility Act of 1982 also permits states to impose "nominal" copayment charges, with exceptions for certain recipient groups and certain services.

Food Stamps

The Food Stamp Program, as currently constituted, is authorized by the Food Stamp Act of 1977 and subsequent amendments. Unlike AFDC and Medicaid, the Food Stamp Program is wholly federally designed and benefits are wholly federally funded. The program also requires fewer links with other programs than does AFDC. But the operations of the Food Stamp and AFDC Programs are intermingled at the state and local level in at least two ways. On a mechanical level, Food Stamps is administered by the same state agencies that administer AFDC. State agencies' procedures must provide for a single interview for food stamp recipients who may qualify for AFDC despite different income rules and different eligibility verification procedures. And while SSI recipients may apply for food stamps at Social Security offices, the state AFDC agency still certifies eligibility, determines benefits, and issues coupons. On a policy level they became intermingled because income determines Food Stamp benefits. Since AFDC grants count as income in such determinations, states may set their AFDC benefits to take into account the value of food stamps as an offset to the valuation of AFDC benefits.

As with AFDC and Medicaid, state agencies are reimbursed for 50 percent of their Food Stamp administrative costs. Also, as is true for AFDC, certain administrative activities, including investigations and prosecutions and certain automatic data processing activities, are reimbursed at a higher rate.

The Food Stamp Program's principal objectives are to alleviate hunger and malnutrition among low-income households. There are few services mandated other than the provision of food stamps to achieve these program objectives. Though the vehicle of delivery is normal channels of trade, there is some expectation that stimulating food purchasing power will benefit the nation's agriculture by promoting, in the words of the 1977 Act, ". . . the distribution in a beneficial manner of the nation's agricultural abundance . . . ." Little else in the enabling legislation emphasizes this link.

Program Changes Over Time

Although it is arguable whether the basic objectives of the programs have changed substantially over their respective lives, the programs have expanded coverage—either through incremental changes in administrative procedure or by explicit legislative changes. As a result of this expansion and other forces external to the program, they have experienced enormous growth in the last two decades, in both populations served and in program expenditures. Tables 3 through 5 and Figures 2 through 4 illustrate that growth for the Food Stamp, AFDC, and Medicaid Programs.

Expenditure and caseload growth have been a concern since the 1960s in the AFDC program. That concern was an important factor in the focus on ineligibility in the creation of AFDC's quality control system; the Food Stamp quality control system, in turn, grew out of that history. In fact, the major growth occurred at different times and for different reasons in each of the programs. While costs have risen for all three programs, the AFDC caseload has actually been relatively stable since the early 1970s; the Medicaid caseload has been stable since 1975 (although its cost has risen faster than the costs for AFDC and Food Stamps); the Food Stamp Program caseload started to decline only recently. And when expenditure growth is adjusted for inflation, growth in real costs is not nearly as substantial as it appears when expressed in

TABLE 3 Food Stamps Expenditures and Caseload, 1964-85

Year	Outlays[a] (in $ millions)	Caseload[b] (in thousands of people)
1964	30.5	366.8
1965	34.4	424.7
1966	69.5	864.3
1967	114.1	1,447.1
1968	184.7	2,210.0
1969	247.8	2,878.1
1970	576.8	4,340.0
1971	1,567.7	9,367.9
1972	1,909.2	11,109.1
1973	2,207.5	12,165.7
1974	2,844.8	12,861.5
1975	4,599.0	17,064.2
1976	5,632.0	18,548.7
1977	5,398.8	17,077.1
1978	5,498.8	16,000.8
1979	6,821.7	17,652.9
1980	9,117.1	21,071.0
1981	11,252.9	22,430.6
1982	11,014.1	22,133.4
1983	11,839.2	21,621.4
1984	11,651.0	20,867.5
1985	11,701.1	19,900.0

Note: Expenditure data include administrative costs.

[a]Data from Office of Management and Budget (1961-1984).
[b]Data from U.S. Congress, Senate (1985:167-172).

current dollars. But the perception of growth is related to current dollars.

There have been many explanations for the caseload and expenditure growth. Some growth was clearly due to explicit legislated changes. AFDC's caseload has changed dramatically both in size and composition since 1935, due to changes outside the program, as well as to two explicitly legislated expansions in coverage. The population receiving food stamps and the size of benefits have been expanded by explicit and repeated changes in the law. Some growth represented shifts from one funding source to another. For instance, prior to 1974 many individuals who later participated in

TABLE 4 AFDC Expenditures and Caseload, 1960-85

Year	Benefit Payments[a] (in $ millions)		Caseload (in thousands of people)
	Federal Share	Total	
1960	611.2	1,021.1	3,005
1961	655.9	1,119.0	3,354
1962	770.2	1,338.6	3,676
1963	826.8	1,425.9	3,876
1964	883.0	1,536.8	4,118
1965	956.5	1,724.9	4,329
1966	1,031.0	1,859.0	4,513
1967	1,170.5	2,065.2	5,014
1968	1,404.4	2,541.7	5,705
1969	1,717.2	3,189.1	6,706
1970	2,187.0	4,081.9	8,466
1971	3,008.3	5,477.4	10,241
1972	3,611.9	6,553.6	10,947
1973	3,856.1	7,003.2	10,949
1974	4,071.3	7,371.0	10,864
1975	4,625.5	8,412.1	11,346
1976	5,257.8	9,676.0	11,304
1977	5,625.6	10,388.0	11,050
1978	5,701.3	10,591.2	10,570
1979	5,825.4	10,779.1	10,312
1980	6,448.4	11,956.0	10,774
1981	6,928.0	12,844.7	11,079
1982	6,922.3	12,856.7	10,358
1983	7,332.2	13,605.8	10,761
1984	7,721.7	14,380.5	10,831
1985	8,027.2	14,943.2	10,855

Note: Expenditure data do not include administrative costs.

[a]Data from U.S. Department of Health and Human Services (1986).

the Food Stamp Program received food through the commodities surplus program. Medicaid combined many already covered services under the new construction of Title XIX, although Medicaid costs have been significantly influenced by inflation and provider costs. But program costs in each program have grown considerably.

TABLE 5 Medicaid Expenditures and Caseload, 1967-1985

Year	Expenditures[a] (in $ millions)		Caseload[b] (in thousands of people)
	Federal Share	Total Cost	
1967	1,208.8	2,368.0	n.a.
1968	1,836.7	3,685.7	n.a.
1969	2,275.5	4,165.8	n.a.
1970	2,574.2	4,731.0	n.a.
1971	3,358.8	6,174.0	n.a.
1972	4,138.2	7,642.4	17,606
1973	4,979.2	9,105.1	19,622
1974	5,833.2	10,170.9	21,462
1975	7,056.4	12,635.5	22,007
1976	10,529.7	18,581.8	22,815
1977	9,727.2	17,211.2	22,831
1978	10,763.9	19,136.9	21,965
1979	12,048.8	21,808.0	21,520
1980	14,138.9	25,227.7	21,605
1981	16,766.3	29,759.5	21,980
1982	17,749.9	31,797.5	21,603
1983	19,326.2	34,684.6	21,554
1984	20,698.0	37,442.6	21,557
1985	22,843.7	41,259.3	21,808

Note: Expenditure data include administrative and training costs and medical assistance payments.

[a]Data for 1967-69 from Office of Management and Budget (1961-84); data for 1970-85 from U.S. Department of Health and Human Services (1986).
[b]Data from U.S. Department of Health and Human Services (1986).

AFDC

The AFDC Program was begun in 1935 as Aid to Dependent Children (ADC), offering coverage only to children deprived of parental support "by reason of death, continued absence from the home, or physical or mental incapacity of a parent . . . ." Until the 1950s, only the children were financially supported. The first program change occurred when the mothers of those children became eligible, and the program became Aid to Families With Dependent Children.

The second change involved the working poor. In the original New Deal conception, ADC was intended for those presumed to be permanently out of the labor force. Workers were to be covered in the core social insurance system by unemployment insurance

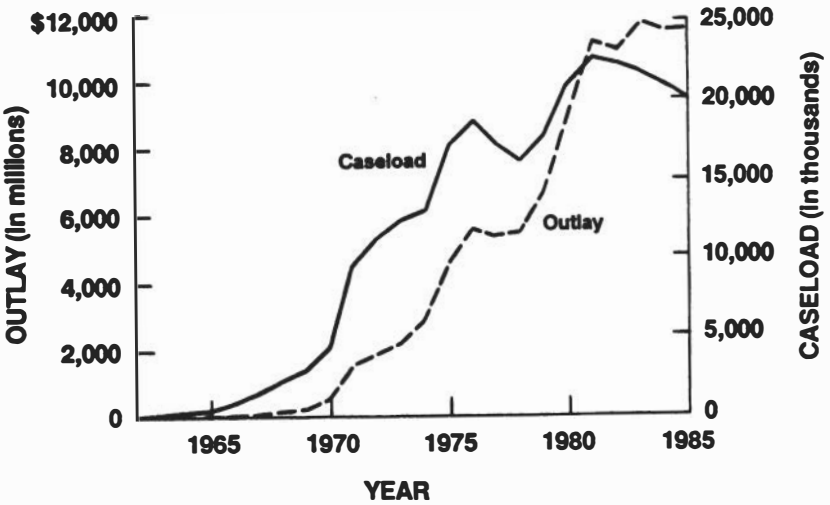


FIGURE 2 Food Stamps: Expenditures and Caseload

Source: Same as Table 3.

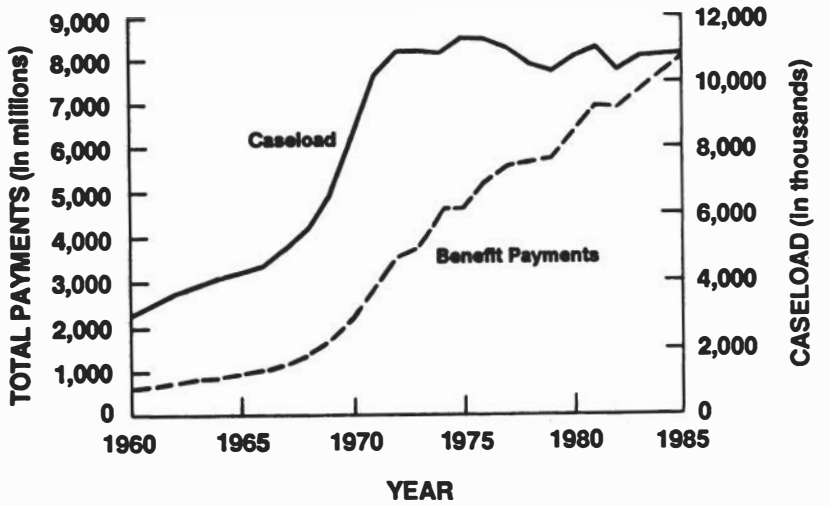


FIGURE 3 AFDC: Expenditures and Caseload

Source: Same as Table 4.

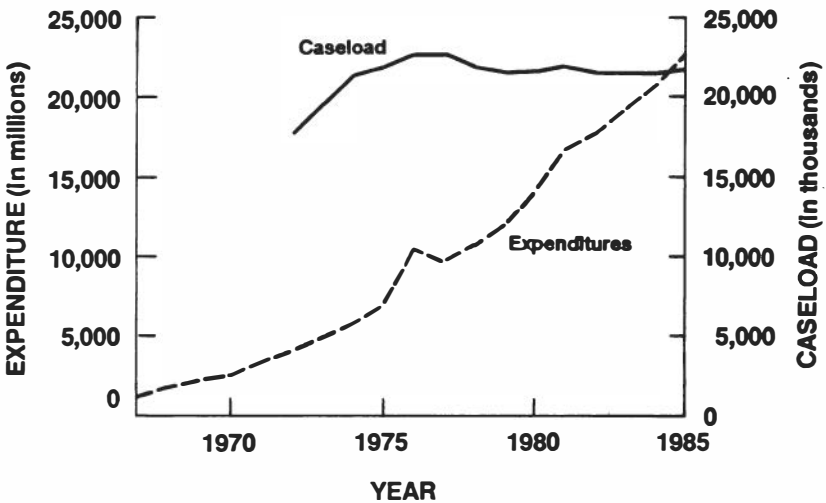


FIGURE 4 MEDICAID: Expenditures and Caseload

Source: Same as Table 5.

and old age assistance. But for a variety of reasons, many of the working poor were not covered by that system (Salamon, 1977), and when concern about poverty reached national attention again in the 1960s, the program was extended, at state option and with several restrictions, to two-parent households in which the father was unemployed.

Other shifts have challenged the assumptions of the original program design and have resulted in further program expansion. Originally, the program was designed on the assumption that mothers' principal responsibility was child rearing. As female labor force participation increased, beginning in World War II, AFDC mothers became increasingly indistinguishable from other women who worked; hence, assumptions about employability were challenged, and this issue continues to dominate the AFDC policy debate. At the same time, poor southern blacks, forced in large numbers out of employment by agricultural mechanization, also joined the caseload after the war (Piven and Cloward, 1971).

Today, the typical AFDC household is a single mother with one or two children. The majority of the national caseload is

nonblack, although blacks on AFDC tend to be concentrated in urban areas and are therefore an especially visible recipient group. Public attitudes toward welfare have continued to show substantial ambivalence, particularly around issues of work and program costs (Free and Cantril, 1967, cited in Wilensky, 1975:37; Gilford et al., 1983:43).

A dominant program goal in AFDC is self-sufficiency and the dominant strategy to achieve that goal is cash support, but AFDC has attempted to accompany cash benefits with the possibility of training, rehabilitation and other supportive services, with limited success. Until 1961 limited social services were provided by state or local welfare agencies. The Social Welfare Amendments of 1962 attempted to integrate cash and services and to emphasize rehabilitation and prevention rather than cash relief in the national program. But by 1969 new regulations and reorganization within the Department of Health, Education and Welfare (HEW) split cash support from services and minimized much of the casework function (Derthick, 1975). The Work Incentive Program (WIN), begun in 1967 to enforce the work requirement, also created an elaborate set of mandated services. But historically, few AFDC recipients have received the prescribed services and participated in WIN. Certain performance standards in the Job Training Partnership Act, created in 1982, attempt to increase employment and training services for AFDC recipients, and several other employment and training options created since 1981 attempt to increase the work effort of welfare recipients and reduce long-term dependency.

At the service delivery level, AFDC remains a cash assistance program with relatively little casework and only episodic ancillary support. However, the various attempts throughout the program's history to develop multiple strategies to promote work and self-sufficiency have enhanced state prerogatives. They have also resulted in intricate and complicated program rules, a complex web of client protections, and complicated and frequently cacophonous relationships between state and federal administering agencies.

Food Stamps

The Food Stamp Program also began in the New Deal, but it was ended in 1943 when the war effectively eliminated farm surpluses.

Begun again on a pilot basis in 1961, the program was made permanent in 1964.

The Food Stamp Program was developed to complement and compensate for gaps in coverage in the family assistance system (Berry, 1984:64; MacDonald, 1977:1), and its history through the first decade was one of repeated and explicit expansions of coverage through changes in federal policy. When the program was made permanent in 1964, it was serving 370,000 persons a month at an annual cost of $30.5 million (Congressional Research Service, 1979). From then until 1974, repeated legislated changes expanded the population eligibile for food stamps. That expansion took two forms: increased geographic coverage, reaching the entire nation in 1974, and changes in eligibility standards.

In 1969, as a direct result of congressional and public concern over low participation, separate benefit schedules then operating for northern/western and southern states were merged. This merger reduced the amount participants had to pay to purchase food coupons and, in many instances, increased the value of monthly allotments and therefore program costs (Congressional Research Service, 1979:4). In 1970 amendments set uniform national income and eligibility standards, which in some areas of the country increased the eligibile population. The 1970 amendments also created annual indexing of allotments (Congressional Research Service, 1979), and in 1973, semi-annual indexing was begun. Indexing was made annual again in 1980.

Some growth in the Food Stamp Program reflects efforts to increase participation among those who are eligible. An "outreach" amendment enacted in 1971 required states to undertake "effective action . . . to inform low-income households concerning the availability and benefits of the Food Stamp program." In April 1975, the U.S. Department of Agriculture (USDA) issued new regulations in order to comply with a court decision forcing more effective outreach. The department established one full-time outreach coordinator in each state who was responsible for developing and monitoring annual state outreach plans, including efforts directed toward elderly and disabled people, migrants, the rural poor, and ethnic groups.

To increase state participation, the federal matching rate for administrative costs was increased from approximately 30 to 50 percent, making state participation more attractive—though other

requirements effectively mandated state participation. Coverage also was expanded to new groups, including alcohol and drug addicts in treatment programs and others receiving prepared meals in groups and in home settings. One analyst attributes the 1973 expansions in part to the failure of the Family Assistance Plan (FAP)—President Nixon's welfare reform proposal (MacDonald, 1977). By 1976 not only had the entire country been covered, but state agencies also had outreach efforts under way.

The Food Stamp Act of 1977 repealed the Food Stamp Act of 1964 and completely rewrote the law to its present form. Although several measures were instituted to restrict eligibility and benefits, participation was once again increased by the elimination of any purchase requirement. Until that time, participants were required to purchase their coupons, at a discounted price, rather then simply to receive monthly allotments outright. The changes in the program were major enough to require about 2 years for the transition to be completed. (During the transition, QC operations were suspended; they were resumed in fiscal 1980.) The elimination of the purchase requirement also was implemented in fiscal 1980. The caseload rose by about three and one-half million persons from fiscal 1979 to fiscal 1980 (see Table 3, above).

Since 1981 there have been several further efforts to contract the program. Among these, the Omnibus Budget Reconciliation Act (OBRA) of 1981 forbid any federal reimbursement to state agencies for outreach activities, lowered the so-called "net" income limit for eligibility, and established in addition an income eligibility limit on gross income after allowable exclusions but before deductions.

Food stamps may be used to purchase a variety of products in addition to food associated with enhancing basic nutrition: including, for instance, any food for personal consumption, seeds and plants, and, for certain eligible households in Alaska, equipment specifically used for procuring food through hunting and fishing. It also includes prepared meals for elderly and handicapped people, and for certain individuals in treatment facilities and group settings.

Because the Food Stamp Program explicitly covers working poor people, it has escaped some of the conflict about employability that has dominated the AFDC policy debate. The Food Stamp Program has had a work requirement since 1971 and allows

workfare. Otherwise it mandates little—and engages in little at the delivery level—in the way of direct services. It does create administrative relationships with food retailers, who turn in the coupons to banks for reimbursement.

HISTORY OF QUALITY CONTROL

The QC system for the Food Stamp Program is built directly on the one for AFDC, which was developed in the early 1970s around specific concern with ineligibility, fraud, and abuse. The history of quality control in the family assistance programs reveals that despite the concern and conflict around poverty, and attempts at comprehensive welfare reform, quality control has never focused on anything but ineligibility and overpayments.

1962-1970: Early Quality Control Focus on Ineligibility and Payment Inaccuracy in AFDC

The period of biggest increase in caseload and cost for the AFDC program provided the backdrop for the creation of the quality control system for that program. Quality control systems for the Food Stamp and Medicaid Programs evolved from that developed for AFDC (U.S. Department of Agriculture, 1986b; U.S. Department of Health and Human Services, 1980). Although program-specific differences may be found, the three QC systems share a common orientation and approach. All three focus on recipient eligibility status and benefit payment accuracy. None focuses on the effects of the receipt of benefits or on precisely how those benefits are used with respect to the achievement of program objectives.

Quality control in AFDC can be traced to pilot programs as early as 1952, but the first in a series of sustained efforts that led to the current QC system began in 1962. Senator Robert Byrd (D-W.Va.), then chair of the Senate Subcommittee for the District of Columbia, ordered a special study of ineligibility among District of Columbia recipients of AFDC. The study reported that more than half of the District's recipients were ineligible. The study was criticized on methodological grounds—including sampling procedures and the validity of the measurements—but it had the effect of raising concerns in Congress about the implications for other large urban areas (U.S. Department of Health, Education and Welfare, 1974; U.S. Congress, House, 1977:204.)

The full Committee ordered a national study of the problem, and HEW and the U.S. General Accounting Office undertook a national eligibility review in 1963 involving, among other things, full field reviews. This study showed substantially lower rates of ineligibility in AFDC caseloads than did the District of Columbia study—estimates ranged from 2 to 15 percent—but rates high enough to be of concern.

In 1964 HEW wrote regulations to establish "quality control" systems nationwide, prompted by the concern over ineligibility rates, according to one federal official present at the time, and by the sense that the responsible federal agency should never again be unable to know what the error rate is (Bowes, 1987). The focus of the new system was on only one aspect of management improvement, the correctness of caseworkers' actions. The sample design used caseworker actions as the sample units, which does not facilitate an accurate assessment of overall error rates, since cases on the active caseload but between redeterminations are not considered.

By 1968 HEW was convinced that the 1964 version of QC was not working. States were slow to establish and make use of QC units within their agencies, and the QC systems were not controlling caseload growth and costs (U.S. Congress, House, 1977): AFDC caseloads had increased by 90 percent between 1960 and 1968, and costs had more than doubled (see Table 4, above).

In 1968, other studies, particularly one of New York City's AFDC caseload, reported levels of ineligibility much higher than had been suspected (General Accounting Office and New York State, 1968, cited in U.S. Congress, House and Senate, 1972). Although the implementation of a QC system had permitted HEW to track error rates based on caseworker actions, the agency was once again embarrassed because its sample inevitably produced a lower error rate than one that sampled all active cases (Bowes, 1987), as did the study in New York City.

At this same time, HEW also began considering a simplified form of eligibility determination, one that would rely primarily upon the declaration of applicants and recipients concerning the elements of their eligibility status. The so-called "declaration" system followed the split between social services and income maintenance. The split meant that many methods of verification

that state agencies had had available, like home visits, were effectively gone. It also carried with it a view, in some eyes, that "income maintenance was perfunctory; any clerk could do it" (Bowes, 1987). With this change in the way eligibility was determined and the new eligibility studies, another fear was coupled to that of high error rates: with fewer verification requirements, greater potential would exist for abuse of the program by persons ineligible for benefits but claiming them nonetheless (U.S. Congress, House, 1977). HEW began a comprehensive reassessment of the QC system, which led to a major revision in 1970. Meanwhile, AFDC caseloads continued to rise, and costs rose even faster.

Three concerns came together, therefore, in the 1970 revision of QC systems for AFDC: (1) the apparent failure of the 1964 system to accurately estimate and report ineligibility rates and its failure to control the costs associated with payments to ineligible recipients; (2) the failure of state agencies to fully implement and make effective use of QC systems; and (3) with simplified eligibility procedures, the fear of greater abuse of the program, and, consequently, of even larger amounts of misspent funds (or spending uncontrolled).

1970-1976: From QC Measures to Standards, with Consequences

The 1970 QC system differed from that of 1964 in at least three important ways (see U.S. Congress, House, 1977). First, QC reviews were to be sampled from active caseloads, not from caseworkers' actions. (Not until 1976 was sampling from terminations and denials, and a review of their correctness added back to the QC system.) Second, more thorough reviews were to be done of recipients' actual eligibility status and the benefits awarded, rather than simply relying on a review of material in the case file to determine the accuracy of caseworkers' actions. While some work in addition to so-called desk reviews had been done in QC eligibility determinations in the past (Bowes, 1987), the field reviews to be done after 1970 required more extensive verification. Third, national performance standards were set, at 3 percent on the case ineligibility measure, and 5 percent each on over- and underpayments in the state caseloads. The consequence of any state's failing to meet one or more of these so-called "tolerance levels" (plus a factor

for sampling error in the design of the state's QC sample) was a requirement that the state agency plan and implement corrective action.

The 1970 QC revision required substantial changes to the system set up after 1964. To carry out the federal agency activities that were to be required, Congress approved federal QC staff positions in 1971 and 1972 (U.S. Department of Health, Education and Welfare, 1974:17). But many state agencies did not move quickly to implement the new system. One source (U.S. Department of Health, Education and Welfare, 1974:17) cites insufficient incentives for state compliance. Another source reports that during this time much attention and federal interest were diverted to welfare reform initiatives, and, if those reforms were enacted, they would have rendered much of the QC system obsolete (U.S. Congress, House, 1977). President Nixon's welfare reform proposal, the Family Assistance Plan (FAP), went down to defeat in the Congress in 1969 and 1970, but reform still occupied part of the White House agenda.

In the early 1970s, HEW contracted with a consulting firm, Westat, for help in improving the AFDC quality control system. Two early reports from Westat (1971, 1973) show a considerable effort to bring to AFDC then-current thinking about "quality control." Westat recommended, for instance, a sampling and review scheme that closely resembles the present QC systems. Westat developed a regression-adjusted estimator for obtaining estimates from two-phase samples; that estimator is now used in all three family assistance programs. Westat also suggested a shift from case accuracy to payment accuracy measures, and all three systems now measure the dollar value of erroneous payments.

Many of the Westat recommendations, however, were not implemented. Those not implemented included the use of much larger samples, which would have offered more detailed information with which to identify and plan corrective action strategies; the use of the lower confidence bounds around state-level error-rate estimates to trigger corrective action planning; the use of national performance "goals" with state-specific performance "standards" based on performance history; and routine cost accounting of QC and corrective action, for cost-benefit analyses. HEW's interest appeared increasingly to be in estimating payment accuracy at the state level.

In a 1972 study of welfare administration by the Joint Economic Committee of the U.S. Congress (U.S. Congress, House and Senate, 1972), a link between payment error and fraud was made explicit: a full 7 pages of the 44-page report was a section entitled "Error and Fraud Uncontrolled." In the committee hearings, then Under Secretary of HEW John Veneman reported that an April 1971 study of case records showed a 5 percent ineligibility rate, with over- or underpayment errors in about 25 percent of the cases reviewed. He cited the complexity and confusion of the national system, not fraud or abuse as the primary cause. Although the committee primarily agreed, its report concluded that ". . . the QC system . . . appears to be relatively ineffective in enabling welfare administrators to prevent error and fraud" (p. 41).

HEW made a major revision to the QC system for AFDC early in 1973: financial consequences were tied to performance as measured against QC-based error rates. For state agencies with error rates exceeding the federal "tolerance levels," HEW would disallow the federal portion of costs represented in that excess. Planning and implementing corrective action would be required of all states, regardless of performance; continued improvement was the obligatory goal. A federal official involved in QC at the time has reiterated that HEW's interest was in recovering erroneous overpayments (Schutzman, 1987).

HEW made some other revisions to the QC system over the next 2 years (U.S. Congress, House, 1977:208-209). These included the introduction of the regression-adjusted estimator for a federal re-review process and a distinction between QC-based point estimates and estimates of the lower bound of confidence intervals constructed around the point estimates. The point estimates would be used for "discussion" purposes—presumably analysis and corrective actions—while the lower bounds would be used for the enforcement of financial disallowances.

1976-1982: Financial Penalties Overturned, Then Restored

In 1976 the U.S. District Court for the District of Columbia (*Maryland v. Mathews*) (Fed. Supp. 415:1206-1214) overturned a critical section of the regulations governing the system, implementing

some of the revisions, those on which tolerance levels had been established. The court ruled that the tolerance levels—set in regulations, not in law—were arbitrary and capricious standards against which to hold state agencies accountable. Hence, HEW could not establish disallowances on the basis of performance measured against such standards. In response to the 1976 court decision federal program administrators moved in other ways to hold state agencies accountable for "excessive" levels of payment error. By 1979, a joint federal-state task force developed a scheme for setting and resetting standards for AFDC on the basis of national rates of error reduction, which is one approach to reflecting performance history and system capability. The Food Stamp Program moved to adopt a similar approach shortly thereafter.

Food Stamp and Medicaid officials readily acknowledge their historical and procedural debt to the AFDC QC efforts (U.S. Department of Agriculture, 1986b; U.S. Department of Health and Human Services, 1980). Their QC systems also rely on an intensive review of samples of cases to make eligibility determinations and an assessment of payment accuracy (in Food Stamps, coupon issuance value) for eligible recipients. State-sampled reviews are subsampled and re-reviewed by federal officials. Although neither program is only a cash assistance program like AFDC, neither program includes in its QC system other possible sources of errors, such as those of food retailers, who accept food stamps and present them for redemption, and those of health care providers, who bill state agencies for the services that they provide to Medicaid recipients. (Both programs do have other systems that examine and attempt to monitor such errors.) All three family assistance programs use regression-adjusted official estimates of payment error rates, and all three now make financial sanction or disallowance decisions based on the point estimates produced.

In an effort to remedy the loss of disallowances as a result of *Maryland v. Mathews*, Representative Robert Michel (R-Ill.) introduced an amendment that legislated overpayment performance standards for Medicaid and AFDC, to become effective in fiscal 1981. State officials contended that this amendment was never enacted since the appropriation bill to which it was tied died when funds were appropriated in a continuing resolution, but federal agency officials contended that the continuing resolution in effect

incorporated the Michel Amendment, and they therefore implemented its provisions.

The standards set for overpayment errors by the Michel Amendment were more stringent than those that would have resulted from the federal-state task force's approach, and an amendment to the Tax Equity and Fiscal Responsibility Act (TEFRA) of 1982 revised these standards and those for Food Stamps, making them still more stringent. For fiscal 1983, AFDC and Medicaid standards were 4 percent; TEFRA lowered the AFDC standards to 3 percent beginning in fiscal 1984, and the Medicaid standards to 3 percent beginning on April 1, 1983. Food Stamp standards were to reach 5 percent by fiscal 1985, with a graduated reduction for the 2 preceding years.

1982-Present: QC Systems Retain Focus on Ineligible Recipients and Overpayments

The quality control systems for the AFDC, Medicaid, and Food Stamp Programs underwent further changes after 1982, but not in their major aspects. They still use a two-phase sample approach, with state reviews and federal re-reviews. They use a regression-adjusted estimator to arrive at an official payment error rate (described in Chapter 3). Although underpayments are measured and an estimated error rate is reported, that error rate is not included in the performance standards. Negative case actions—terminations and denials—are also sampled, reviewed, and the error rates for them are reported, but these reviews get less intensive scrutiny (particularly at the federal re-review stage), and they, too, are not incorporated in performance standards.

The focus on the costs of ineligibility has been consistent throughout QC history. The early studies that inspired the implementation of QC in family assistance programs pointed to a problem of ineligible recipients. Each QC redesign has focused on eligibility issues and reducing the possibilities for abuse. In 1970 the system moved to thorough, full field reviews on a sample of the active caseload. Over the next few years the Westat recommendations for estimating state-level overpayments were implemented, but other recommendations applicable to state quality improvement were not. Each major change has also followed closely the

substantial growth of AFDC caseloads from 1960 through the mid-1970s, with even greater increases in program costs.

By the time caseload growth had become a source of general concern, the major studies related to QC concerns had already sounded an alarm about ineligibility. It is not surprising that concerns about caseload and cost increases might have become linked with concerns about eligibility, and implicitly about fraud and abuse.

Food Stamp Error Rates: 1980—Present

Since fiscal 1980 the Food Stamp QC system has reported state-level estimates of underissuance error rates, payment error rates (which include overissuances to eligible recipients, as well as issuances to ineligible recipients), and "error concentrations." Error concentrations show the relative magnitude of payment errors attributed to particular types or sources of error. Only one such error code per case may be coded in a case review, so the error concentrations attribute the entire value of the error for any case to a single source, even though multiple sources may be present.

Table 6 shows that Food Stamp payment errors for the most part are related to income or to deductions from income. Together these two sources or types of error consistently account for more than two-thirds of erroneous payments.

The national overpayment error rate has declined between fiscal 1980 and 1984, but the decline has not been consistent: the rates were 9.51, 9.90, 9.55, 8.32, and 8.64 for the 5 years, respectively. The national trend masks a variety of different behaviors among the states, although the state-level payment error rates over time also show improvements in performance over those 5 years (U.S. Department of Agriculture, 1986a:Table 5). At the same time, underpayment error rates show little change over those 5 years. Table 7 summarizes the two sets of state error-rate distributions for fiscal 1980-1984. Figures 5 through 9 show greater detail in the distributions of the overpayment error rates for those years.

TABLE 6 Percent Distribution of Benefit Dollars Overissued, by Source

Month/Year	Income	Deductions	Resources	Other	Nonfinancial	Computation
10/79-3/80	46.1	24.6	20.4	n.r.	8.9	0
4/80-9/80	41.0	29.2	22.6	n.r.	7.3	0
10/80-3/81	38.9	28.6	24.9	n.r.	7.6	0
4/81-9/81	37.1	29.8	26.4	n.r.	6.7	0
10/81-3/82	39.9	33.1	17.9	n.r.	6.7	2.4
4/82-9/82	38.3	33.5	19.4	2.8	6.1	n.r.
10/82-9/83	31.4	38.1	12.3	3.4	14.8	n.r.
10/83-9/84	34.6	39.2	6.6	3.2	16.5	n.r.

Note: n.r. indicates not reported.

Source: Data from U.S. Department of Agriculture (1979-82; 1983-86).

TABLE 7 Number of States in Different Error Rate Ranges, Food Stamp Program, by Fiscal Year

Error Rate Percentage	Fiscal Year 1980	1981	1982	1983	1984
Underpayment					
≤1.70	10	5	8	8	11
1.71-2.99	34	38	35	33	30
≥3.00	7	8	8	10	10
Overpayment					
≤7.00	5	3	5	13	13
7.01-10.99	39	31	36	31	35
≥11.00	7	17	10	7	3

Source: Data from U.S. Department of Agriculture (1986a:8-9, Tables 4 and 5), includes District of Columbia.

FIGURE 5 State-Level Food Stamp Overpayment Error Rates: Fiscal 1980

4.08, 4.49	NV,HI
6.62, 6.68, 6.97 7.14, 7.40, 7.43, 7.48, 7.51, 7.65, 7.69 8.00, 8.10, 8.29, 8.45, 8.61, 8.62, 8.70, 8.83, 8.91	
9.13, 9.18, 9.19, 9.20, 9.35, 9.48 9.61, 9.74, 9.84, 9.95, 9.97	
10.26, 10.28, 10.29, 10.36, 10.39, 10.40 10.41, 10.42, 10.48, 10.48, 10.86, 10.87	
11.88	
12.16	
13.16, 13.50	NM,RI
14.60, 14.86	MD,DC
15.78	NY

Note: Units are percentages to two decimal places. The first entry, for example, shows an error rate of 4.08 percent for Nevada; the last entry is 17.78 percent for New York.

Source: Data from U.S. Department of Agriculture (1986a:8, Table 4), includes District of Columbia.

FIGURE 6 State-Level Food Stamp Overpayment Error Rates: Fiscal 1981

```
 3.39                                                                NV

 5.16                                                                ND
 6.97
 7.11, 7.36, 7.45, 7.50, 7.65, 7.74, 7.75, 7.89
 8.08, 8.09, 8.26, 8.49, 8.50, 8.52, 8.99

 9.03, 9.09, 9.11, 9.17, 9.21, 9.28
 9.30, 9.31, 9.40, 9.49, 9.56, 9.80

10.10, 10.24, 10.45, 10.50
11.02, 11.12, 11.13, 11.31, 11.37
12.25, 12.37, 12.53, 12.85, 12.87
13.12, 13.34, 13.48, 13.63, 13.78                    DC,NM,MT,NY,CT
14.22                                                           MD

23.23                                                           AK
```

Note: See Figure 5.

Source: Data from U.S. Department of Agriculture (1986a:8, Table 4), includes District of Columbia.

FIGURE 7 State-Level Food Stamp Overpayment Error Rates: Fiscal 1982

```
 1.48                                                                NV

 5.72, 5.96                                                       AL,HI
 6.40, 6.89
 7.15, 7.40, 7.41, 7.56

 8.02, 8.20, 8.32, 8.34, 8.37, 8.49, 8.56
 8.61, 8.68, 8.72, 8.90, 8.93, 8.99

 9.03, 9.11, 9.25, 9.57, 9.64, 9.69
 9.69, 9.70, 9.71, 9.79, 9.82

10.04, 10.25, 10.26, 10.36, 10.51, 10.64, 10.67, 10.87
11.10, 11.40, 11.42, 11.98
12.73, 12.85
13.38                                                           MA

15.07                                                           CO
16.29                                                           NH

20.80                                                           AK
```

Note: See Figure 5.

Source: Data from U.S. Department of Agriculture (1986a:8, Table 4), includes District of Columbia.

FIGURE 8 State-Level Food Stamp Overpayment Error Rates: Fiscal 1983

2.17	NV
4.28, 4.94, 4.98	HI,DE,ND
5.52, 5.52	WV,MT
6.46, 6.72, 6.73, 6.83, 6.90, 6.90, 6.98	
7.12, 7.22, 7.23, 7.48, 7.57 7.70, 7.84, 7.86, 7.92, 7.95	
8.20, 8.27, 8.33, 8.37, 8.48 8.51, 8.77, 8.79, 8.88, 8.90	
9.09, 9.45, 9.79, 9.88, 9.98, 9.99 10.03, 10.08, 10.08, 10.09, 10.37 11.43 12.63, 12.80	
13.29, 13.60, 13.86	UT,MA,AK
16.71	VT

Note: See Figure 5.

Source: Data from U.S. Department of Agriculture (1986a:8, Table 4), includes District of Columbia.

FIGURE 9 State-Level Food Stamp Overpayment Error Rates: Fiscal 1984

2.54	NV
3.59, 3.69	ND,HI
5.83	MO
6.09, 6.27, 6.40, 6.46, 6.65, 6.74, 6.85, 6.88, 6.95 7.08, 7.11, 7.22, 7.47, 7.35, 7.61, 7.63, 7.67 8.18, 8.31, 8.51, 8.64, 8.77, 8.79, 8.80, 8.98	
9.00, 9.08, 9.18, 9.23, 9.24, 9.29, 9.38 9.57, 9.60, 9.66, 9.71, 9.77, 9.86, 9.97	
10.09, 10.14, 10.16, 10.41, 10.66 11.43, 11.83	
13.35	AL

Note: See Figure 5.

Source: Data from U.S. Department of Agriculture (1986a:8, Table 4), includes District of Columbia.

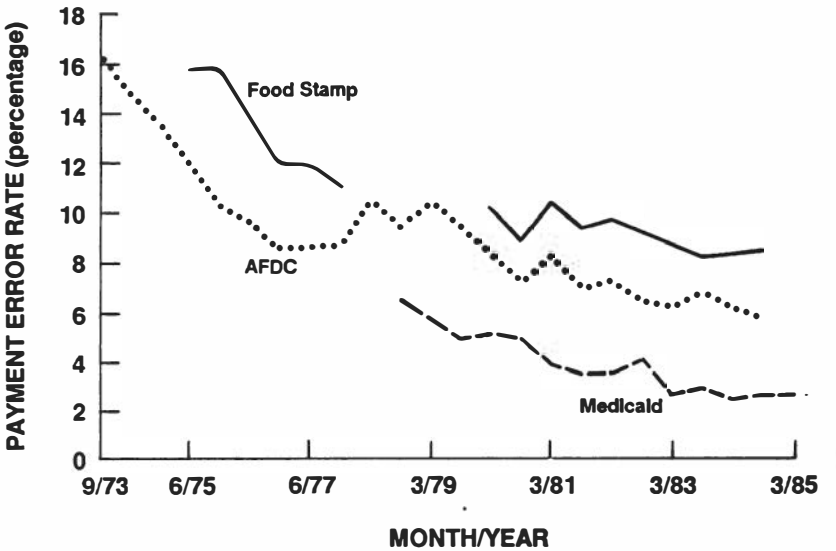


FIGURE 10 Trends in Overpayment Error Rates: AFDC, Food Stamp, and Medicaid Programs

Note: Food stamp quality control operations were suspended after the passage of the Food Stamp Act of 1977 due to the major changes in the program; it was resumed in 1980.

Source: Rust (1986).

CONCLUSION

The history of QC beginning with AFDC suggests that QC resulted in large part from concern about overpayments, in the face of rising program costs. As shown in Figure 10, error rates for all three family assistance programs have been reduced since QC systems began. But QC systems cannot have been expected to reverse caseload growth or program costs because the major elements of that growth have not been ineligibility or overpayments. In fact, costs during the period described grew faster than caseloads. Caseload growth is frequently the result of policy changes and intended expansion in coverage. Growth in costs reflect a variety of factors, including explicit policy choices and inflation in both the Food Stamp and Medicaid Programs, since food stamp allotments are indexed to increase with inflation, and

Medicaid costs reflect the increase in provider fees. QC systems now in place do not measure the effects of such choices. Nor do they measure the achievement of program objectives other than payment accuracy, or the relationship between the achievement of those objectives and costs.

REFERENCES

American Public Welfare Association
n.d. Chronology of Federal Quality Control in AFDC, Medicaid and Food Stamps 1952 Through December 1982. American Public Welfare Association, Washington, D.C.

Berry, J.M.
1984 *Feeding Hungry People: Rulemaking in the Food Stamp Program.* New Brunswick, N.J.: Rutgers University Press.

Bowes, J.
1987 Personal interview, March 6. Special Assistant to Acting Director, Office of Family Assistance, formerly Director of Quality Control, AFDC, U.S. Department of Health and Human Services, Washington, D.C.

Congressional Research Service
1979 *A Concise History of the Food Stamp Program.* Washington, D.C.: Library of Congress.

Derthick, M.
1975 *Uncontrollable Spending for Social Services Grants.* Washington, D.C.: The Brookings Institution.

General Accounting Office
1985 *Federal Benefit Programs: A Profile.* GAO/HRD 86-14. Washington, D.C.: U.S. Government Printing Office.

Gilford, D.M., D.P. Affholter, and L. Ingram, eds.
1983 *Family Assistance and Poverty: An Assessment of Statistical Needs.* Committee on National Statistics. Washington, D.C.: National Academy Press.

MacDonald, M.
1977 *Food, Stamps, and Income Maintenance.* New York: Academic Press.

Office of Management and Budget
1961-84 *Appendix to the Budget.* Washington, D.C.: U.S. Government Printing Office.

Office of Management and Budget
1986a *Appendix to the Budget 1987.* Washington, D.C.: U.S. Government Printing Office.

Office of Management and Budget
1986b *Budget of the United States Government, Fiscal Year 1987.* Washington, D.C.: U.S. Government Printing Office.

Office of Management and Budget
1986c *Historical Tables: Budget of the United States Government, Fiscal Year 1987.* Washington, D.C.: U.S. Government Printing Office.

Piven, F.F., and R.A. Cloward
1971 *Regulating the Poor: The Functions of Public Welfare.* New York: Vintage Books.

Rust, D.A.
1986 Presentation made before the Panel on Quality Control of Family Assistance Programs, July 11, 1986. National Academy of Sciences, Washington, D.C.

Salamon, L.M.
1977 *Toward Income Opportunity: Current Thinking on Welfare Reform.* Durham, N.C.: The Institute of Policy Sciences and Public Affairs of Duke University and the Ford Foundation.

Schutzman, F.
1987 Personal Communication, February 13. Director, Bureau of Quality Control, Health Care Financing Administration, U.S. Department of Health and Human Services, Washington, D.C.

U.S. Congress, House
1977 *Administration of the AFDC Program.* Committee on Government Operations. 95th Congress, 1st Session. Washington, D.C.: U.S. Government Printing Office.

U.S. Congress, Senate
1985 *The Food Stamp Program: History, Description, Issues, and Options.* Committee on Agriculture, Nutrition, and Forestry. S. Prt. 99-32. Washington, D.C.: U.S. Government Printing Office.

U.S. Congress, House and Senate
1972 Issues in welfare administration: an administrative nightmare. *Studies in Public Welfare.* Paper No. 5 (Part 1). Joint Economic Committee, 92nd Congress. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Agriculture
1979-82 *Semiannual Summary Report of Food Stamp Quality Control Reviews 1975-82.* Food and Nutrition Service. Washington, D.C.: U.S. Department of Agriculture.

U.S. Department of Agriculture
1983-86 *Food Stamp Quality Control Annual Report, Fiscal Years 1983-86.* Food and Nutrition Service. Washington, D.C.: U.S. Department of Agriculture.

U.S. Department of Agriculture
1986a Food Stamp Quality Control Executive Overview, Fiscal Year 1984. Food and Nutrition Service, U.S. Department of Agriculture, Washington, D.C.
1986b *History of Quality Control in the Food Stamp Program.* Food and Nutrition Service. Washington, D.C.: U.S. Department of Agriculture.

U.S. Department of Health, Education and Welfare
1974 *Training for Federal Quality Control Review.* Office of Management. SRS. Washington, D.C.: U.S. Department of Health, Education and Welfare.

U.S. Department of Health and Human Services
1980 *Background Paper on Medicaid Quality Control for Moynihan Hearing.* HCFA. Washington, D.C.: U.S. Department of Health and Human Services.

U.S. Department of Health and Human Services
1986 Unpublished data from Office of Family Assistance and HCFA, U.S. Department of Health and Human Services, Washington, D.C.

U.S. District Court
1976 State of Maryland vs. Mathews, *Federal Supplement* 415: 1206-1214.

Westat Research, Inc.
1971 *Evaluation of Quality Control System in Public Assistance: Report on Phase II.* Rockville, Md.: Westat Research, Inc.

Westat Research, Inc.
1973 Evaluation of the Tolerance Levels Established for the Quality Control System in Public Assistance. Unpublished report submitted to the U.S. Department of Health, Education and Welfare. Westat Research, Inc., Rockville, Md.

Wilensky, H.L.
1975 *The Welfare State and Equality: Structural and Ideological Roots of Public Expenditures.* Berkeley, Calif.: University of California Press.

3
The Food Stamp Program and Its QC System

The Food Stamp QC system focuses on the accuracy of states' eligibility and benefit determinations. This chapter describes the operations of the Food Stamp Program, as established in law and regulation, that might relate to quality control operations. They include: eligibility criteria and benefit standards established in law, with certain requirements of states in certifying and recertifying eligibility and determining benefits; national research, demonstration, and evaluation requirements that are related to program quality; and the operations of the Food Stamp QC system itself. The Food Stamp Program sets out requirements in many other areas, such as those that govern allowable and unallowable uses of food stamps; the printing, distribution, issuance, and redemption of food stamps; and the registration of retail outlets where stamps may be accepted. Although these provisions could clearly affect the quality of the program, they have never been tied to the QC system.

The law itself and the rules and regulations governing eligibility and the process of determining eligibility are complex. It is reasonable to question whether and to what extent the complexities themselves are important sources of error. The current QC

system is not designed to undertake such analyses or to provide information that support such judgments.

BASIC PROGRAM REQUIREMENTS

Eligibility and Benefits

Eligibility for food stamps is based primarily on income and assets, and the benefits are based on the amount of money required to provide a minimum but nutritionally sound diet. The case unit in the program is a household, which can consist of 1 or more persons. A food stamp household may not include all persons who live in that household—it is even possible for some individuals within a household to be barred from participation. There is a requirement for work registration and participation for some in an employment and training component of the program, but unlike other family assistance programs, there are few restrictions to participation other than need. State agencies must undertake certain verifications, provide for a variety of protections of applicants' and participants' rights, and provide for appropriate employment and training services for certain participants.

Income and Assets Tests

Eligibility for food stamps requires, in most cases, two income tests and an assets test. Assets for a household usually cannot exceed $2,000; ($3,000 for households with an elderly member). Income definitions specify deductions and exclusions (detailed below). Households with an elderly or disabled person are not eligible if their income after specified exclusions and deductions exceeds the poverty line. For all other households, income after exclusion but before deductions may not exceed 130 percent of the poverty line, ***and*** income after both exclusions and deductions may not exceed 100 percent of the poverty line.

Income. Determining countable income can be complicated. Income includes wages and salary, and gross income from self-employment; training allowances; assistance payments such as SSI, AFDC, and general assistance; unemployment compensation; annuities, pensions, and retirement benefits; strike benefits; foster

care payments; support payments or alimony; and dividends, interest, and royalties. Most income that accrues for periods shorter than a year, other than piecework or self-employment, is averaged over the entire 12-month period; some income, such as money for education-related living expenses, is averaged over the period for which it is received. Income that is not subject to monthly reporting may be estimated prospectively; other income is reported retrospectively.

The allowable exclusions from income include: noncash payments and infrequent income up to $30 per quarter; some deferred education and other loans; money for third-party beneficiaries not in the household; income of a student under 18; the costs of producing self-employment income and exact expense reimbursements; energy assistance payments or allowances; and certain nonrecurring lump-sum payments; and income that other federal law excludes for food stamp eligibility determination.

Deductions from income include a standard deduction of $99 per month per household for the continental United States (more for other areas), adjusted annually for changes in the Consumer Price Index (CPI) for urban consumers, plus 20 percent of all earned income to compensate for mandatory salary, education, and other work expenses. A dependent care deduction is allowed for actual expenses of up to $160 per household per month, to allow for employment, training, or education. An excess shelter expense is also allowed of up to $149 per month for the continental United States (annually adjusted by the CPI for urban consumers), for shelter costs above 50 percent of monthly household income after all other deductions. The excess shelter deduction may include a standard utility allowance. Households containing an elderly or disabled person are also entitled to an excess medical expense deduction for allowable medical expenses exceeding $35 per month and an unlimited excess shelter deduction.

Assets. Liquid and nonliquid assets may not exceed $2,000 for a household or $3,000 for a household of two or more in which one person is age 60 or older. Assets include savings and retirement accounts, the value of licensed vehicles in excess of $4,500 at fair market value (except if used to produce income or necessary to transport a physically disabled household member), and the full value of any other vehicles and mobile homes for vacation purposes.

Transferring assets in order to attempt to qualify for food stamps disqualifies an applicant for up to 1 year.

Eligibility

Eligibility is established for periods of 6 to 12 months for individuals who are required to submit monthly reports of income, and for 1 to 12 months for all other households, depending upon the predictability of the household income. For people categorically eligible for food stamps by virtue of their receipt of AFDC, SSI, or aid to the aged, blind, or disabled under appropriate titles of the Social Security Act, the period of eligibility must coincide with that for assistance in selected other programs, but may not be longer than 12 months.

Individuals must be disqualified for making false or misleading statements or otherwise violating federal or state laws or regulations governing the program. Disqualifications are for 6 months for the first offense, 1 year for the second, and permanently for the third offense.

Although the program's history is one of changes made to expand coverage, more recent changes restrict coverage. Eligibility does not extend to strikers unless the households were already eligible prior to the strike. Students are subject to special restrictions, and many are not eligible. Eligibility requires periodic reporting of income, depending on income and earnings status, and the provision of a social security number for each household member, for use in a variety of reporting and verification procedures.

In most households with earners aged 16 to 60 (except those of migrant farmworkers) those persons are required to register annually for work. An individual may be disqualified permanently if he or she refuses to accept a job without good cause or for 90 days if he or she voluntarily quits a job without good cause. If the individual is the household head, the entire household is disqualified. Exemptions to the work registration requirement include: compliance with AFDC work registration requirements; dependent care responsibilities, including those for a child under age 6; provisions for some students enrolled at least half time and others who comply with special work requirements for students; regular participation in an alcohol treatment program; actual employment

for a minimum of 30 hours per week; physical or mental inability to work; and receipt, or pending receipt, of unemployment compensation.

Benefits

Food Stamp allotments are determined on the basis of the USDA Thrifty Food Plan—a particular diet required to feed a family of two adults (man and woman) and two children—adjusted for household size. The cost of the Thrifty Food Plan is adjusted for areas outside the continental United States and updated annually. It is assumed that a poor household will spend 30 percent of its countable income (that is, total income less exclusions and deductions) on the purchase of food. The Food Stamp benefit, or allotment, is the difference between the cost of the Thrifty Food Plan for an eligible household and 30 percent of its countable income. The minimum allotment for eligible households of 1 or 2 persons is $10. After eligibility is determined, a household's allotment for the first month is prorated from the date of application for that month.

Requirements of State Agencies

The Food Stamp Program is administered by the state agency responsible for administering federally aided public assistance programs, according to an approved state plan. State plans must include bilingual personnel and materials as appropriate, but since 1981, states are prohibited from conducting outreach activities with federal funds. The law also mandates that states "to the extent practicable, verify the income and liquid resources of the household prior to issuance of coupons . . . ."

Other rules aim to provide various safeguards for people who are eligible. State agencies are required to provide an application on the same day as an initial request and to accept the application on that day; to comply with standards permitting telephone and mail contacts for certain individuals; and to determine eligibility within 30 calendar days of application. The law establishes standards for timely notification of recertification, for safeguarding privacy, and for expedited coupon issuance (within 5 days) under certain circumstances. The law requires that the state agency provide fair hearings; promptly restore coupons wrongfully denied;

have a plan for providing benefits to disaster victims; and prominently display USDA posters and make available to participants USDA pamphlets that provide information about good nutrition and other related programs.

HHS and USDA are required to develop a system of single interviews to determine eligibility for AFDC and food stamps; to permit SSI recipients to apply for food stamps at Social Security offices and be certified on the basis of information in the Social Security Administration's files; and to develop procedures that may permit, at state option, a single application form for food stamps and AFDC or other public assistance programs, as well as permit, at state option, a single verification based on the public assistance application.

Furthermore, the secretaries of USDA and HHS are required to consult in issuing regulations so that, to the extent feasible, the definition, valuation, and calculation of income and assets for persons who receive benefits under both the AFDC and Food Stamp Program will be comparable. (About 40 percent of those who receive food stamps also receive AFDC benefits.) That comparability is not always achieved. The bases for valuation of an automobile, for instance, are specified differently in the laws governing the two programs, and they cannot be made comparable through regulations.

Other rules are aimed at facilitating proper and accurate eligibility and benefit determinations. For verification of eligibility, states must request information available from the state employment service, and in other cases, from the Social Security Administration or Internal Revenue Service. States are required to set up a system to protect against multiple receipt of benefits in more than one jurisdiction. States must also develop and submit for approval a plan for the use of an automatic data processing and information retrieval system. States are encouraged to have systems that include elements for the determination of eligibility, calculation of benefits, issuance of benefits, computer matches for income and asset verifications, reconciliation procedures, generation of statistics for program reporting, and coordination with related federal and state programs. Finally, states also must operate a quality control system (discussed below).

National Research, Demonstration, and Evaluation Responsibilities

Section 17 of the Food Stamp Act of 1977 as amended specifies a number of activities, described below, aimed at assessing program effectiveness, including several that are required. The activities may not relate directly to current QC activities. Also, some research efforts listed below such as the eligibility simplification and the evaluation of the effects of the Omnibus Budget Reconciliation Act (OBRA) of 1981, have been completed. The important point to be noted is that these activities are relevant to improving the quality of the Food Stamp Program. The law specifies such activities as pilot programs and experimental studies to test mechanisms for improved program efficiency and improved delivery of benefits, including the use of cash, checks, or vouchers in place of coupons. Also included is a requirement for up to 14 pilot projects to test eligibility simplification for recipients of food stamps and AFDC, Medicaid, or SSI. The law permits the USDA secretary to conduct up to ten pilot studies for implementing workfare, five statewide and five at the substate level.

The law also mandates two sets of activities to assess the effects of the program and specific legislative changes on specific program objectives. It mandates that the USDA secretary:

> . . . develop and implement measures for evaluating, on an annual or more frequent basis, the effectiveness of the food stamp program in achieving its stated objectives, including . . the program's impact upon the nutritional and economic status of participating households, the program's impact upon all sectors of the agricultural economy, including farmers and ranchers, as well as retail food stores, and the program's relative fairness to households of different income levels, different age composition, different size, and different regions of residence.

The law further mandates that the secretary:

> . . . implement pilot programs to test various means of measuring on a continuing basis the nutritional status of low income people, with special emphasis on people who are eligible for food stamps, in order to develop minimum common criteria and methods for systematic nutrition monitoring that could be applied on a nationwide basis.

The law also requires that the secretary conduct studies of the effects of reductions in benefits as a result of the Omnibus Budget Reconciliation Act of 1981, the Food Stamp Amendments of 1981,

the Food Stamp Commodity Distribution Amendments of 1981, the Food Stamp Amendments of 1982 and "any other laws enacted by the Ninety-seventh Congress which affect the Food Stamp Program." The study efforts must include evaluations of the effects of retrospective accounting and repaying procedures on benefit and administrative costs and error rates, and the degree to which eligible households are denied food stamp benefits for failure to complete periodic reports. Food stamp research and evaluation work is carried out by FNS; the budget for those activities for fiscal 1985 was $8.9 million.

THE QUALITY CONTROL SYSTEM

The quality control system establishes the means by which the federal government monitors state agencies' performance with respect to benefit accuracy, develops estimates of that performance, and imposes sanctions or rewards for that performance. This section describes the sample design for QC, how the reviews are done and the performance measures are developed, how performance estimates are derived from QC samples, and the sanctions and rewards that follow from estimated performance.

As with most of the costs of state administration of the Food Stamp Program, the federal government pays 50 percent of the costs of state QC operations. To promote certain activities that are intended to improve program administration, the federal government pays a larger share of specific administrative costs; including 75 percent for the costs of designing and installing automatic data processing systems and for investigations and prosecutions of fraud. States are permitted to keep 50 percent of the value of payments that are recouped from ineligible recipients and of overpayments to eligible recipients, if the household is found to have intentionally violated program rules. For other recoupments, the state may keep 25 percent, as long as the overpayments were not due to state agency error. Although such incentive funding is not tied directly to the QC system, it is intended to improve the administration of the program.

Performance as estimated under the QC system can affect the basic 50 percent federal share for most administrative costs. Poor performance results in a diminished federal share (detailed below);

good performance may result in an additional 10 percent enhanced funding for basic administrative costs.

The QC Sample Design

Every month, state officials sample separately from two universes of cases: active cases, households that received food stamps, and negative cases, households that were denied food stamps or were terminated from the program. The state agency must submit a QC sampling plan for FNS approval. Federal Food Stamp officials at the regional level subsample from the state's QC sample.

State Samples

States may use simple random sampling, stratified random sampling, systematic random sampling, cluster sampling, or whatever probability sample design best meets a state's needs. If a state agency chooses to adopt a sample design other than simple random sampling, the design must be fully described and documented, submitted for approval as part of the state plan, conform to probability sampling principles, and provide for estimates of payment error rates with at least the precision that would be obtained by simple random samples of the size that result from the use of FNS formulas for sample size calculation for simple random samples.

FNS Handbook 311 (p. 12) states that:

> . . . the determination of any sample size is based on several factors, both statistical and administrative . . . A minimum sample size is established to provide sufficient data to support the analytic demands placed on the quality control sample. An upper limit on the sample size is set to balance the costs of additional data collection with the potential gains in precision.

In part, then, the prescribed sample sizes for simple random QC samples seem to reflect consideration of the precision of estimates of payment error rates to be derived from the sample.

Above the minimum sample size requirements for the smallest states, however, size calculations are primarily established by the state's Food Stamp caseload. For active cases:

(a) if a state's average monthly caseload is under 10,000, then the minimum QC sample size is 300 cases per year;

(b) if a state's average monthly caseload is over 60,000, then the standard minimum QC sample is 2,400 cases per year and the optional minimum is 1,200 per year; and

(c) for all states with average monthly caseloads from 10,000 to 60,000, the standard and optional minimum samples sizes for QC are derived by formula, n_s and n_o, respectively,

$$n_s = 300 + 0.042\ (N - 10{,}000) \text{ and}$$
$$n_o = 300 + 0.018\ (N - 10{,}000),$$

where N is the average monthly caseload.

The optional minimum sample size may be chosen by state agencies if they agree not to dispute later sanctions, which are tied to estimated performance, on the basis of the precision of the estimate derived from the sample. If a state agency is unwilling to forego its ability to contest a sanction based on sampling error in the estimate, then it must use the standard minimum sample sizes or an alternative design that will provide for at least the precision afforded by the standard minimum random sample.

For negative case actions, the sample sizes also vary by caseload size, totalling 150-800 case actions.

Federal Subsamples

The subsamples for federal re-reviews are drawn from both state samples, active cases and negative actions, and are a function of the state sample sizes. For active cases, federal subsamples range from 150-400 cases per year; from one-half (in the smallest caseload states) to one-third (in the largest caseload states) of the optional minimum state sample sizes. For negative actions, they range from 75-160 per year, but in fact negative action QC samples are rarely subsampled for federal re-review. If a state's sample design is stratified, the federal subsample will follow that stratification. (FNS Handbook 315 contains descriptions of subsample designs.)

By following the criteria for setting state QC sample sizes, the subsample design seems to reflect primarily considerations of caseload rather than the precision of estimates to be derived from QC samples and subsamples. The established minimum sample sizes for small states, however, reflect some consideration of the precision of estimates.

Quality Control Reviews

The purpose of drawing samples of active cases and negative actions is to subject them to intensive review, to assess the accuracy of eligibility determinations and benefit calculations. The data collected during QC case reviews make possible a variety of analyses of sources and types of errors and can be used to help with corrective action planning later. The review process runs on a schedule, as does the re-review of subsampled cases by federal officials in regional offices. FNS maintains a system for the resolution of differences in findings between federal and state QC reviewers for the subsampled cases.

Eligibility and Benefit Reviews

FNS Handbook 310 sets out minimum standards of evidence for state reviews of each element of eligibility, including all sources of income and possible assets, but states may use stricter standards. Consequently, the standards vary across states. FNS describes primary and secondary standards of evidence, but what constitutes acceptable documentation may also vary for elements of eligibility that are to be verified. For the most part, documents from a government or public agency constitute primary evidence, and collateral information from any of a number of sources constitutes secondary evidence. When primary evidence is not available and secondary evidence is used, at least two pieces of secondary evidence are required.

For some elements of eligibility, especially income, the depth of the investigation depends on the nature of an applicant's case. Affirmative statements by an applicant of wages earned, for instance, may be documented with pay stubs, and self-employment income requires receipts and other documentation. But denials of income by an applicant require that the QC reviewer establish a work history, contact previous employers to ascertain the recipient's employment status with those employers for the month in question, and also query either the state employment security agency's or the Social Security Administration's records of reported wages or both. The QC reviewers are also instructed to look for other indications of possible employment, for instance (FNS Handbook 310:108):

> **Information obtained while verifying other elements of eligibility, with individuals such as relatives, school officials, landlord, etc., may show the participant is frequently absent from the home. This may be an indication of . . . % employment. Additional situations which the reviewer shall investigate are: indications that the head of the household is not at home during regular working hours and leaves at the same time every day; difficulty in finding the head of the household at home; seasonal employment at its peak in the area where the participant lives; shelter costs higher than reported income; and indications of a relatively high standard of living.**

Standards of evidence are described for all elements of eligibility and benefit calculations, including all the sources of income and deductions and assets detailed above. Indications of unreported income or assets must also be verified. For instance, a computer match of bank records may turn up an account that the participant had not reported. Household composition also must be verified.

Federal and state reviewers use the same standards of evidence, but federal re-reviews are not done independently of the state reviews. Federal re-reviewers work from the completed QC case files submitted by states: they may simply accept the evidence reported; verify the evidence reported; or gather additional evidence, based on their own judgment about the usefulness or probable correctness of prior verification.

Sources and Types of Errors

State and federal QC reviewers not only identify errors, but also calculate the value of such errors (except for those that lead to erroneous negative case actions). The calculations cover payments to ineligible recipients, overpayments to eligible recipients, and underpayments to eligible recipients. In addition, certain case characteristics are recorded, such as household composition and income by source. State and federal reviewers distinguish "client-caused" errors from "agency-caused" errors, assign a judgment about willful misrepresentation by participants as opposed to other explanations for client-caused errors, and characterize major sources of error such as income, deductions, resources, and numerical mistakes in calculating benefits. Finally, the source and type of the primary error also is coded on data entry forms.

Review Schedules and Reports

State agencies must report monthly on the list of cases and actions sampled for review. They must also report on the disposition of all sampled cases and actions, including those dropped from the sample for permissable reasons and not subject to review. The findings for 90 percent of the sample are due in the FNS regional office within 75 days of the end of the sample month. All case findings must be received by FNS within 95 days of the end of the sample month, including an explanation for all cases for which reviews were not completed and a schedule for their disposition. The reporting mechanism, a monthly status report, also must include a description of the sampling process. Finally, on a yearly basis, the edited results of all state QC reviews for the year are to be sent to FNS no later than 95 days after the end of the annual review period.

Resolution of Disagreements Between Federal and State QC Reviewers

Federal and state QC reviewers sometimes disagree on the findings for particular cases. If their differences cannot be resolved by mutual discussion, states may submit their differences for arbitration, first within the FNS regional office and then, if they are not satisfied, to the central FNS office. Federal and state reviewers agree on the vast majority of QC reviews, both for the many cases for which no errors are found and also for many in which some kind of error was found. Neither FNS nor others who addressed the panel could tell it what proportion of cases in the federal subsamples are submitted for arbitration in regional offices. Furthermore, none could tell the panel what proportion of these cases from regional office arbitration are appealed to the FNS national office. Of those case disagreements submitted by the states for further arbitration to the national office, however, about 20 percent of the regional FNS office findings are overturned or altered (Center on Budget and Policy Priorities, 1986).

Error Rate Estimates

Several error rates are estimated annually from the QC sample and reviews. Some are based on the full state sample, others on

the federal subsample, and an official payment error rate then combines different estimates. This combined estimate is the basis for the "official payment error rate," the key for the establishment of state-level liabilities for sanction: the performance standard for state agencies is set for the official payment error rate, and states are sanctioned for failing to meet the standard.

Estimates in the State Sample

Four case error rates and four payment (or issuance) error rates are estimated from the state sample. Case error rates represent the proportion of all cases reviewed that are in error; payment error rates are the proportion of the value of all food stamps issued that were issued in error to all payments. The case error rates are for: (1) ineligible cases among the active case sample; (2) overpayments in the active case sample; (3) underpayments in the active case sample; and (4) cases in the negative sample that were wrongly denied or terminated.

The first three of the payment error rates correspond to the first three case error rates, those among the active case sample. The fourth payment error rate is an error rate for total overissuances: the sum of payment error rates for ineligible recipients of food stamps and for overissuances to eligible recipients, that is, the sum of the payment error rates for (1) and (2), above.

Estimates in the Federal Subsample

The federal re-review is not particularly concerned with case error rates. The federal re-reviewers do calculate and report the four payment errors rates. In addition, the federal re-reviewers calculate two other estimates based on state findings (not federal findings) for the subsample: the underpayment (underissuance) error rate (i.e., (3) above), and the summed overpayment error rate (issuances to ineligible recipients plus overissuance to eligible recipients). The state-based estimates in the subsample generally do not agree precisely with those for the larger sample, nor do they necessarily agree with those for the federal re-review of the subsample.

Official Error Rates

Three separate estimates of error rates—the separate federal estimate based on the subsample, the state-based estimate in the federal subsample, and the state estimate from the full state sample—are combined to arrive at official error rates. A regression-adjusted estimator is used to arrive at two official error rates: one for underpayments (item (3) above), the other for overpayments (items (1) and (2) above). The latter is the "official payment error rate."

The use of a regression-adjusted estimator rests on statistical theory for deriving an estimate from a two-phased sampling design such as that used in the QC systems (Cochran, 1977:Chapter 12). In the larger sample, one measure is taken (say x_i); in the subsample, a different measure is taken (say y_j), usually a more precise and more expensive measurement. To estimate a population parameter such as the payment error rate (call it ER) in these QC systems, information from both sets of measurements is used. In the larger sample, $\bar{x}_1$ provides one estimate of ER; in the small sample, both $\bar{y}$ and $\bar{x}_2$ provide other estimates. (The subscripts on $\bar{x}$ distinguish the estimate based on the larger or full sample from the estimates based on the subsample). Sampling theorists use the two $\bar{x}$ estimates to adjust or calibrate the $\bar{y}$ estimate—that is, to estimate what $\bar{y}$ would have been if the y_j measures had been taken in the full state sample.

One class of estimators simply adds to $\bar{y}$ a weighted difference between the two $\bar{x}$ estimates, that is,

$$\widehat{ER} = \bar{y} + c(\bar{x}_1 - \bar{x}_2) .$$

One way to determine a value for c in the equation above is to replace it with an estimated regression coefficient, $\hat{b}$, from the regression in the subsample of the y_js on the x_js. The estimator then becomes

$$\widehat{ER} = \bar{y} + \hat{b}(\bar{x}_1 - \bar{x}_2) .$$

This is the regression-adjusted estimator used in the QC systems for the family assistance programs, where the y_j measures are from the federal QC re-reviews and the x_is are the measures from the state QC reviews. In essence, the regression-adjusted estimator is designed to calibrate the federal error rate estimate by the difference between the two corresponding state-based numbers. The statistical theory that underlies this official estimate of payment

error rates uses the presumed greater accuracy of the federal findings as the standard and incorporates the greater precision of the larger state QC sample to arrive at an estimate that is "better" than either a state-based or federal-based estimate alone.

If there is no pattern of agreement between state and federal reviewers, then the expected value of $\hat{b}$ is 0. The official estimate will be that based on the federal reviews alone, and it will have only the precision afforded by the smaller, federal subsample. When state and federal reviewers agree completely, two important consequences ensue. First, $\hat{b}$ equals 1.0. Second, the federal and state estimates in the federal subsample are equal, that is, $\bar{y} = \bar{x}_2$. Then the official estimate of the payment error rate, $\widehat{ER}$, becomes the state estimate in the full state sample, $\bar{x}_1$.

Performance Standards

Official payment error rates are estimated separately for underpayments and for overpayments. However, only the latter, the "official payment error rate," is related to sanctions that are tied to estimates of performance. Sanctions are set as functions of differential federal sharing in the costs of administering the program.

For fiscal 1981 and 1982, standards were set in a formula that reflected national average rates of error reduction. Beginning with fiscal 1983, however, states were to reduce their payment error rates over 3 years, reaching 5 percent by fiscal 1985. For most states, the error rates were set at 9 percent for fiscal 1983, 7 percent for fiscal 1984, and 5 percent for fiscal 1985 (and subsequent years). For states with very high error rates (over 11 percent) for fiscal 1982, the declining targets were reductions over each of the next 2 years of one-third of the difference between the fiscal 1982 performance and the 5 percent target for fiscal 1985.

A number of consequences follow from performance estimates that derive from QC samples. State agencies must engage in corrective action planning, they may be subject to financial sanctions, and they may be eligible for rewards, for "enhanced" federal sharing in administrative costs for exceptionally good performance.

Corrective Actions

In response to findings of deficiencies in QC reviews, and if their estimated error rates do not meet the established standard, a

state agency must develop a corrective action plan and submit it for FNS approval. Some deficiencies that may trigger corrective action planning may arise outside of the QC process—for instance, from FNS or USDA audits or investigations. Failure to complete QC reviews on time for 95 percent of the QC sample also will trigger corrective action plans, as will negative case action error rates higher than 1 percent.

Corrective action plans must address the problems identified. The primary factors for determining the most serious deficiencies to be addressed in the plan are the magnitude of the deficiency, the geographic extent of the deficiency, and the likelihood of success of planned corrective actions. Corrective action plans must include documentation of the problems to be corrected, their magnitude and geographic extent, the source by which they were discovered, and the anticipated actions that will eliminate or reduce the problems.

Corrective action plans also must be developed for appropriate project or management levels within the state agency, tailored to specific units for which the plan is operative. Such substate plans also must meet the minimum requirements for the state corrective action plan. The state must monitor and evaluate the corrective action program at all levels at which it works within the state and must describe the monitoring and evaluation program in its corrective action plan. The intent here is to ensure that the planned corrective actions are implemented properly and that the anticipated results are achieved within the specified schedules.

State corrective action plans are open-ended and remain in effect until all deficiencies are reduced or eliminated. As the deficiences are corrected or improved according to the plan, FNS is to be notified in writing, with documentation to support removing the deficiency from the plan. Such removal is subject to FNS review and validation. As a practical matter, corrective action plans frequently become annual activities because they are triggered by annual error rate findings.

Sanctions

When a state's official payment error rate does not meet the standard, the federal share of administrative costs is reduced. This is the sanction in the Food Stamp Program. Federal funds for administration are reduced by 5 percent for each percentage point

or fraction above the error rate standard, for the first 3 percentage points above the standard, and by 10 percent for each percentage point or fraction above the 3 percentage point difference. That is, official payment error rates above the target rate by 1, 2, and 3 percentage points would result in sanctions of 5, 10, and 15 percent of the federal share of administrative costs, respectively, and official payment error rates above the target by 4, 5, and 6 percentage points would result in sanctions of 25, 35, and 45 percent of the federal share, respectively. The sanction against administrative dollars, however, cannot exceed a total that is equal to the value of food stamps issued in that year in excess of the allowable error rate.

A provision of the Food Security Act of 1985 requires that states must be notified of their potential liability within 9 months of the end of the fiscal year. There are a variety of conditions under which states may request a "good cause" waiver from total or partial payment of sanctions, if such a request is made within 30 days of the notification of potential liability. The conditions include natural disasters or civil disorders, strikes by workers involved in eligibility determination, significant caseload growth, changes in the Food Stamp or other programs adversely affecting the Food Stamp Program, misapplication of federal policy by or because of FNS representation, or other circumstances beyond the states' control. There is no time requirement for FNS to respond to requests for waivers based on "good cause" claims.

A state may also appeal the federal sanction finding by requesting a hearing, in addition to a review of the record, within 10 days of notification and submitting within 30 days a written statement attesting to issues, position, pertinent facts and related items. The hearing must be scheduled within 60 days, and both parties may submit further information within 10 days of its conclusion. A determination must be made within 30 days. Within 30 days of final determination, the state may file for judicial review in the appropriate U.S. District Court. Figure 11, shows the time line for the appeals process.

Rewards

The reward for good performance—which also is tied to federal sharing of administrative costs—uses the underpayment error rate and the negative case action error rate in addition to the official

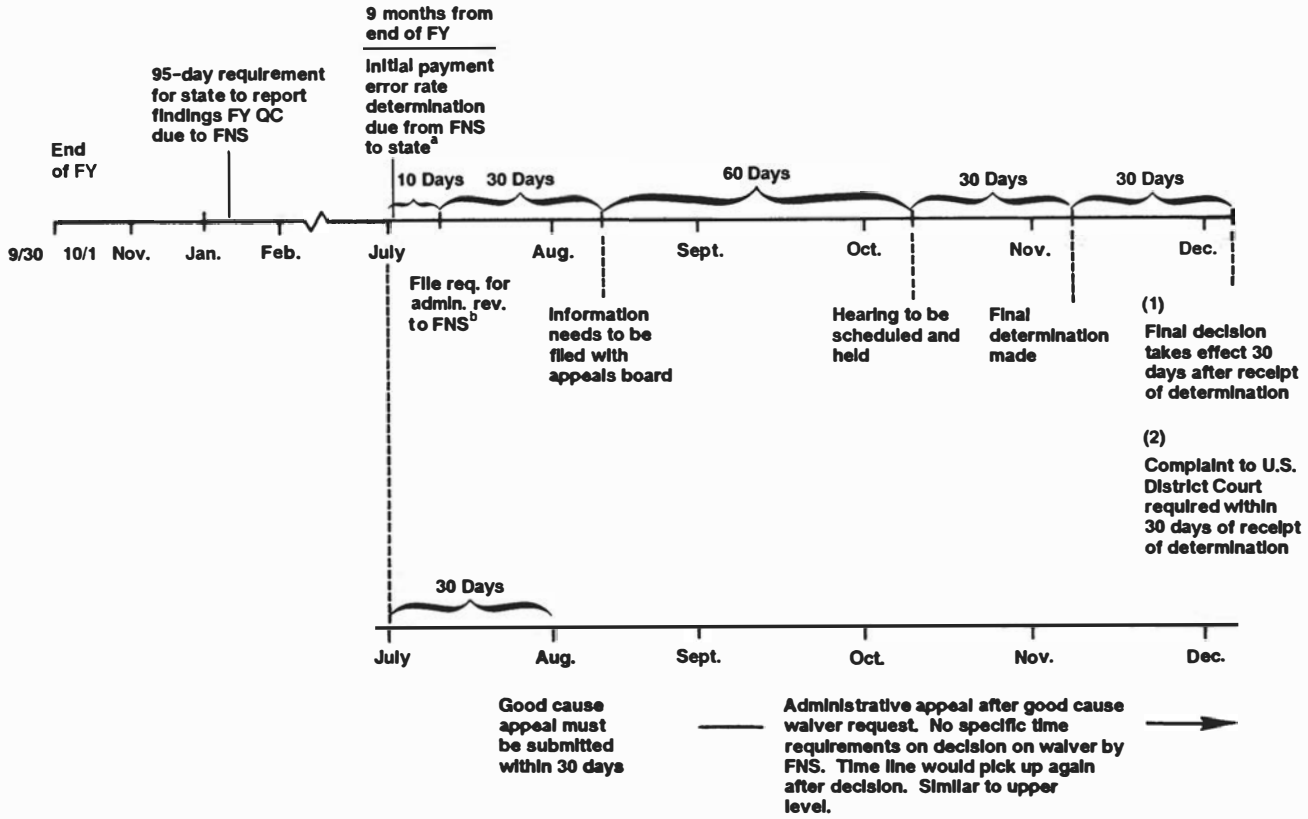


FIGURE 11 Schedule for the Determination of QC Sanctions

[a]Error rate takes affect if no administrative appeal or request for good cause waiver.
[b]Process and time line for administrative appeal if no good cause waiver request.

payment error rate. First, in order to be eligible for the incentive funding, the sum of a state's official payment error rate and its underpayment error rate must be no larger than the 5 percent performance standard. Then the negative case action error rate is examined: a state gets incentive funding only if its negative case action error rate is not larger than the national average negative case action error rate.

When a state meets these tests, it is eligible for enhanced federal shares in the basic costs of administering the program; 60 percent instead of the usual 50 percent. Very few states have ever qualified for this incentive payment, although one state, Nevada, has received it several times.

REFERENCES

Center on Budget and Policy Priorities
1986 Issues Related to Quality Control and Error Rate Sanctions System, Particularly in the Food Stamp Program. Center on Budget and Policy Priorities, Washington, D.C.

Cochran, W. G.
1977 *Sampling Techniques*, 3rd ed. New York: John Wiley and Sons, Inc.

4
Critiques of the Food Stamp QC System

"Quality control" is a much studied subject. It has many applications—in manufacturing, in service industries, and in some public programs. Quality control of the family assistance programs has generated intense feelings among many who are involved in or affected by either the program, the consequences of the QC measures, or both. These programs also present a unique set of problems associated with maintaining or improving quality. The panel has therefore drawn upon a variety of sources in its assessment of the current Food Stamp Program QC system.

The panel gathered a large body of analyses of the three QC systems under study. That body includes rigorous studies of program effects as well as a range of other analyses and observations that have been presented to the panel in writing, orally, and in connection with the panel's site visits. (The reference list at the end of the chapter includes items specifically cited and other resources used.) The issues that the panel considers salient in those critiques are summarized in this chapter. Because the QC

systems for family assistance programs share common roots and approaches, the panel reports some analyses that do not speak specifically to the Food Stamp QC system. In large measure, this work is critical of the QC systems. Since this work typically is done in reaction to perceived problems with the QC systems, this result is not a surprising one. In general, the critics call the QC systems into question on grounds of "fairness," although the meanings and standards of fairness vary. Systems that levy penalties tend to generate concerns about fairness. Some criticisms speak directly to an issue of "fairness" in the operations of or outcomes of the current system, others are only implicit. Imbedded in each is some sense of unfairness to one or more of those interests to which the programs are accountable: to taxpayers; to local, state, and federal officials; to recipients; and to the target or eligible populations.

For purposes of analysis and presentation, the panel presents its discussion of critiques of QC systems under four major headings: statistical issues, measures of quality, the use of performance standards with sanctions/disallowances, and procedural issues. The groupings are not necessarily those used by others in their assessments of QC systems.

STATISTICAL ISSUES

Statistical critiques fall under two headings, sample design and estimation. Questions focus on the validity, accuracy, propriety, and cost-efficiency of applications of statistical theory and methods. At issue is whether, or to what extent, current QC sample designs and estimates fairly represent state agency performance and balance state and federal financial risks.

Sample Design

Critics have taken on questions of sample sizes (too large, too small), the varying precision of sample-based estimates by state (state QC sample sizes vary as a function of caseloads), the dropping of cases from QC samples under certain conditions, and the cost-efficiency of the current "two-tiered" system of QC sampling and review.

Sample Sizes

Some critics argue that the current QC sample sizes are too large. In all states except those with the smallest samples, current samples provide greater precision of estimated payment error rates than is needed for federal monitoring purposes. The HHS Inspector General, for example, concludes that in all states but those with the smallest QC samples, AFDC QC sample sizes could be substantially reduced for federal monitoring purposes (U.S. Department of Health and Human Services, 1986). These critics also argue that current QC samples do not yield sufficiently useful additional information for various analysis purposes to justify the larger samples. For example, data do not permit the construction of so-called error-prone profiles: profiles that permit the identification of cases that are likely to be prone to error, and can be targeted for special handling in eligibility determination (see, especially, American Public Welfare Association, 1986; U.S. Department of Health and Human Services, 1986). Error prone profiling attempts to sort cases into groups and identify those groups that are more prone to error in their eligibility and benefit determinations.

Other critics, however, argue that the current QC sample sizes are too small: they cannot produce sufficiently reliable substate estimates of performance, such as for local offices or teams of workers within offices. Early advice given to HEW on the AFDC QC system suggested a sampling and review system for state and local management purposes, with larger samples and different kinds of reviews than are now used in the family assistance QC systems (Westat, 1971). Representatives of state interests have argued the need for useful management information, which includes reasonably precise substate estimates (see American Public Welfare Association, 1986; Sacramento County, 1986; National Council of State Human Service Administrators, 1986; Florida, State of, 1986; National Governors' Association, 1986).

Federal officials note that states have the option of supplementing QC samples and review activities and that the administrative costs of such supplementation will be matched at the 50 percent rate by the federal agencies. A number of states reportedly do use such supplemental samples and reviews. And New York City samples and reviews intensively, by federal QC review

standards and measures, the caseloads of small groups of workers (panel site visit, August 1986).

Precision of Estimates

QC sample sizes now vary according to state agencies' caseloads. The precision (or sampling error) of state-level estimates of payment error rates, therefore, varies by state as well. The differences in precision become important if and as interval estimates (or confidence intervals around point estimates) are used for administrative actions (see U.S. General Accounting Office, 1985b). Varying precision in the estimates, by state, also raises questions of equity with respect to the consequences or risks that arise from the uses of the estimates (U.S. General Accounting Office, 1986d).

The federal formulas for determining state QC sample sizes take into consideration the relative costs to states of undertaking QC reviews, which are labor intensive. Samples designed to provide estimates of equal precision across states would require sample sizes that are much more nearly equal. Such designs would result in greater administrative costs relative to caseloads among states with smaller caseloads.

Dropped Cases

Federal rules permit some cases to be dropped from QC samples under specified circumstances, usually if a person-to-person interview cannot be scheduled. (One example is that of the death of the recipient, for instance.) One study by the U.S. General Accounting Office (1986a) claims that failure to pursue and use some dropped cases leads to underestimates of payment error rates. The GAO examined 360 dropped cases properly dropped under current regulations and determined that, in principle, 242 might have been further reviewed. Of those, the GAO completed reviews for 95 cases. On the basis of the findings for those 95 cases, the GAO concluded that dropped cases are about twice as prone to error as those for which QC reviews now are completed. The GAO therefore recommended that the federal agencies revise their rules for dropped cases, to restrict the criteria by which cases can be dropped from QC samples.

Two-Phase Sample Design

Finally, many critics have questioned the two-phase sample design now used. Federal regional officials subsample the states' QC samples and re-review state reviewers' findings and determinations. This two-tiered system adds substantial time to the entire review process and, some argue, yields little or no added information (see U.S. Department of Health and Human Services, 1986; U.S. General Accounting Office, 1986d). Also, this system at times puts state QC reviewers in conflicting roles: responsible for finding errors of their own state agency that have major cost consequences for that agency derived directly from the federal interest. Although both state and federal governments have an interest in payment accuracy, it is the federal rules and procedures that dominate QC findings of error.

The U.S. General Accounting Office (1986d) reviewed the pros and cons of the two-tiered sample design as against a one-tiered alternative, federally drawn and reviewed, but made no recommendations. Several groups have supported a one-tiered alternative QC sample design for federal monitoring purposes, which might cost less, would turn results around more quickly, and would not lose any useful information (see National Council of State Human Service Administrators, 1987; Sacramento County, 1986; Illinois, State of, 1986). Such an alternative may make state-federal relationships more difficult, however, and could break an important state-level connection between QC reviews and planning effective corrective actions (see, especially, U.S. General Accounting Office, 1986d). The FNS currently has under way a trial demonstration of a one-tiered sample and review design, but that design is not intended to evaluate the possibilities of reduced sample sizes.

Estimation

Critics have attacked the regression-adjusted estimator that provides official payment error rates; the use of a point estimate rather than an interval estimate of error rates; the estimation of (or, failure to estimate) standard errors of payment error rate estimates; and the errors imputed to uncompleted QC review cases.

Regression-Adjusted Estimator

Perhaps the most frequently made and probably the most hotly debated criticism of current QC systems concerns the use of a regression-adjusted estimator from the two-phase sample. All three QC systems use the regression-adjusted estimator in order to derive official payment error rates for states. The preliminary estimate from federal re-review findings among the federal subsample for each state is adjusted with a weighted difference between two state-review-based estimates—one for the state sample and one for the federal subsample in the state. The purpose of the regression is to derive a "weight" for the difference between the state estimates from the full sample and the subsample. The weight is provided by the estimated coefficient for the regression of federal case level findings of error on state findings of error in each state's federal subsample.

Blischke (1985) and DMA Corporation (1985) make "model-based" criticisms of the regression-adjusted estimator. They argue that the usual meaning of an estimated regression coefficient cannot be sustained in this application because the model's assumptions are not borne out by the data. Fairley (1986) and Desmatics (1986) make similar arguments. In response, however, it can be argued that the regression-adjusted estimator arises from sampling theory for estimation from a two-phased sample design (Cochran, 1977) much like the two-tiered QC review system. The justification rests on sample "design-based" statistical theory that requires samples that are only "large enough" in order for the estimator to be valid.

Westat Research, Inc. (1971) initially recommended to HHS the use of the regression-adjusted estimator; it later reported (Westat Research, Inc., 1986) results of simulations of AFDC payment errors in order to demonstrate the characteristics of the regression-adjusted estimator. Westat found that the regression-adjusted estimator behaves well enough, within the sample sizes found in QC applications. That is, Westat's results suggest that QC sample sizes are large enough to justify reliance on large sample statistical theory. The panel has not formed a judgment on the study.

Westat also compared the performance of the regression-adjusted estimator to that of a related, but simpler alternative, a "difference" estimator. The difference estimator also adjusts the

federal estimate from the federal subsample with a weighted difference of the two state estimates. Westat investigated weights of 1.0, 0.9, and 0.8 in its simulations and reported that its simulations show essentially no difference in the statistical behavior of the regression and difference estimators for weights of 1.0 and 0.9.

Use of a Point Estimate

Any statistical estimate that is based on sample data contains some error based on sampling variability alone. An estimate that consists of a single numerical value, such as an average, is called a point estimate. Statisticians often construct confidence intervals or interval estimates that cover some range of values around the point estimate. Such intervals permit a probability-like statement of confidence that the interval covers the value of the population parameter for which the sample data provide an estimate, e.g., a 95 percent confidence level.

State welfare agencies and their representatives argue that the lower bound at 95 percent confidence, or the lower limit of an interval estimate, should be used to establish the amount of sanction instead of the point estimate of state-level payment error rates. They argue that sampling error can produce estimates that overstate actual payment error rates, and states should not have to assume the risk of oversanction; therefore, they argue for using the lower bound of estimated payment error (National Council of State Human Service Administrators, 1987; Sacramento County, 1986; Covington & Burling, 1986; Desmatics, 1986; Fairley, 1986; Florida, State of, 1986). An earlier report of the National Research Council also recommended the use of the lower confidence bound for the purposes of sanctions or disallowances (Gilford et al., 1983:63-64). In its early AFDC QC reports, even Westat Research, Inc. (1971, 1973) implied that a confidence bound should be considered for taking administrative action based on estimates of performance, although Westat's discussion pertained only to corrective actions and not to financial penalties. The U.S. General Accounting Office (1986d), however, notes that the use of a lower confidence bound would shift to the federal government more of the financial burden for erroneous payments relative to the federal-state shares when the point estimate is used. Since the precision of the payment error rates vary by state, the use of a lower confidence bound could alter the relative rankings of

states by estimated payment error rates (U.S. General Accounting Office, 1985b).

Estimates of Standard Errors

Problems arise with respect to the estimation of the standard errors of official estimates of state payment error rates. The U.S. General Accounting Office (1986b) remarked upon FNS' failure to make and report estimates of standard errors and recommended that FNS do so in the future. Westat's (1986) simulations demonstrate shortcomings in the usual estimates of standard error for AFDC estimates of error rates. The usual method for estimating standard error assumes a particular symmetry, but the regression-adjusted estimator has a skewed distribution. Westat recommended to HHS that it use another approach for estimating standard errors, called the jack-knife approach. The recommended jack-knife method partitions a data set into a number of subsets and then re-estimates the log-transformed payment error rate a number of times, each time leaving out one subset of the data. The set of estimates then provide information from which to estimate the standard error. Note that if confidence intervals are to be constructed, estimates of standard errors are required.

Uncompleted Reviews

Regulations set a schedule within which state agencies must complete QC reviews. Often, some reviews will not be completed within the prescribed time limit. Incomplete QC case reviews usually amount to fewer than 5 percent of the state samples. Federal agencies impute findings of error to uncompleted reviews, assuming that cases with uncompleted reviews are more prone to error. The overpayment imputed is at a value of two standard deviations above the average overpayment error in the sample of reviews completed by the state agencies (using the full state sample). This adjustment, according to critics, has not been supported empirically, and may overstate the errors actually made (see, e.g., Desmatics, 1986). The U.S. General Accounting Office (1986a, 1986b), however, suggests that the method currently used may actually understate errors. The imputation now is made for uncompleted cases before deriving the regression-adjusted, official payment error rate; GAO suggests making the imputation after deriving a regression-adjusted estimate for completed cases only.

MEASURES OF QUALITY

The particular measures of quality used in the QC systems for family assistance programs have come under attack: whatever the particular criticism, the measures taken can or do distort decision making—particularly when sanctions and disallowances are tied to those measures. To the extent that the measures ignore important program objectives or emphasize something at odds with basic program goals, actual quality as opposed to measured quality may deteriorate.

Other Performance Criteria

A number of critics argue that other important performance criteria are ignored in the QC systems, particularly those that pertain to other program objectives (American Public Welfare Association, 1986; Sacramento County, 1986; National Council of State Human Service Administrators, 1987; Florida, State of, 1986; National Governors' Association, 1986). Such things include the timeliness with which applications and redeterminations or recertifications are done, costs of administering the programs, training and job placements (with subsequent savings in benefit payments), child support collections, recoupment of misspent funds, and the effective provision of other kinds of services in addition to payments.

In another area of performance, some critics charge that the cost-effectiveness or cost-efficiency of QC activities, including corrective actions, has never been assessed—and that the focus on payment accuracy standards neglects or overlooks this important dimension of quality (see, e.g., Illinois, State of, 1986; National Governors' Association, 1986). In its early studies, Westat (1971, 1973) advised HEW to collect, track, and report the costs of and savings from QC and corrective actions.

These critics and others (Brodkin, 1986; Center on Social Welfare Policy and Law, 1986a, 1986b; Brodkin and Lipsky, 1983) also argue that the emphasis in the current QC systems on overpayments adversely affects some recipients or potential recipients. The QC system gives no weight to underpayments and incorrect negative actions. Brodkin (1986) and Brodkin and Lipsky (1983) argue that the QC system has encouraged eligibility workers to err on the side of "stringency" in making eligibility (re)determinations

and that some otherwise eligible poor people are not served, or not served well.

Case Churning

The Center on Social Welfare Policy and Law (1986a, 1986b) also stresses a related point, that the QC emphasis results in substantial "churning" of the caseloads. Recipients denied (or terminated) for procedural reasons—those documentation requirements for which noncompliance would create QC errors—may in fact be eligible and may be found eligible (or eligible again) within months. Because QC emphasizes procedural corrections, it ignores other program objectives, such as giving timely needed assistance. The Center on Social Welfare Policy and Law cites a rapid increase in applications and procedural denials and terminations as circumstantial evidence for this claim; it also cites studies in New York City, Michigan, and by Mathematica Policy Research, Inc. that show that many recipients who are denied or terminated for procedural reasons do indeed reapply within a few months and are found eligible. In the New York City study, 64 percent of the AFDC cases studied were reopened within 2 months and another 12 percent within 6 months.

Payment errors may indeed be avoided through procedural denials and terminations. But, critics charge, the denial of needed assistance to poor people for those months when they are not receiving benefits is a potential cost that is not captured through QC measures, and may be promoted through the incentives created by the QC systems. They argue that the risk of procedural denial is particularly high because the documentation required to meet QC standards is frequently difficult for this population to provide.

Negative Case Actions

The USDA states that underpayment error rates have remained roughly constant, nationally, while overpayment error rates have improved (Bode, 1986). The Center on Budget and Policy Priorities (1986) asserts, however, that the effect of improper negative case actions is more important than underpayments, and believes that informal denials may have increased: that is, potential Food Stamp applicants may be discouraged from making a formal application when they make an inquiry about program participation.

The Urban Institute (1985) concluded that as a result of legislative changes in 1981-82, participation in the Food Stamp Program decreased more than expected as a result of those changes, by about 600,000 participants in 1983.

The U.S. General Accounting Office (1986c) studied a sample of negative action cases reviewed for QC in Illinois: it determined that 23 percent of the cases had been improperly denied or terminated; the state QC reviewers found only a 3 percent error rate for negative actions. Moreover, GAO found that, contrary to regulations, benefits were not restored in the specific cases found by QC reviewers to have been improperly denied or terminated.

The panel notes that in its site visits to local, state, and regional offices around the country, it was reported that state QC reviews do not examine negative actions as intensively as they do the active caseload and that federal re-reviewers almost never re-review negative action cases. Also, in the QC systems, the negative case action errors that are found are never evaluated with respect to the value of the payments not made or food stamp coupons not issued.

Misleading Measurements

Technical Errors

AFDC and Food Stamp (but not Medicaid) QC systems record and report certain procedural or "technical" errors. Technical errors include the absence of evidence of Social Security numbers for individual recipients in assistance cases, the failure to make or document registration in work or training programs when such registration is required, and the failure to assign child support collection rights for AFDC children to the state. The FNS records and reports such errors, but excludes them from estimates of Food Stamp payment errors; under AFDC, in contrast, such errors are recorded, reported, and included in official estimates of AFDC payment error rates (see U.S. General Accounting Office, 1986d).

Many critics have argued, state welfare directors and state interest representatives in particular, that such "errors" do not represent misspent funds, since if corrected they would have no effect on the benefit costs (see, for example, American Public Welfare Association, 1986; Sacramento County, 1986; Wyoming, State of, 1986; California, State of, 1986; Center on Budget and Policy

Priorities, 1986). In response, HHS officials argue that these procedural requirements are established in law and therefore must be counted into AFDC payment error rates (see U.S. General Accounting Office, 1986d). Panel members have examined examples of when such errors may include total or partial mispayments, as well as the more obvious examples of when mispayments would not occur once the procedural errors were corrected.

Double-Counted Errors

The Center on Budget and Policy Priorities (1986) documents a class of cases for which erroneous payments are double counted for Food Stamp and AFDC QC purposes. When unreported income is discovered in a QC review for a case that received both AFDC and Food Stamp assistance, the AFDC grant is in error by a magnitude determined by the amount of income not reported. These misspent federal dollars are included in the AFDC payment error rate. In making a determination of error in the Food Stamp issuance, QC reviewers count the full amount of the AFDC grant received and add the amount of unreported income again in determining what the Food Stamp issuance should have been. By including the AFDC grant error in computing the Food Stamp error, the true total of misspent federal dollars is overstated. That is, if the AFDC benefit had been corrected, the Food Stamp overissuance would not have been as large as is now counted in QC determinations.

Omissions from QC Measures

Other critics note two troublesome omissions from QC measures. For example, AFDC QC does not count an overpayment of less than $5.00 as an error, and in all three programs only a single cause of error may be coded. Many experts believe that for management improvement purposes, all errors should be recorded and multiple causes of error (or all causes of multiple errors in a single case) should be reported (see, e.g., Center on Budget and Policy Priorities, 1986; U.S. General Accounting Office, 1980).

PERFORMANCE STANDARDS AND SANCTIONS

Critics have challenged the basis upon which performance standards have been set, variation in error rates in the three programs, the distorting effects of consequences now tied to measured QC performance, the reasonableness of holding state agencies accountable for clients' errors, and making the Food Stamp Program sanction a function of administrative costs.

Basis for Performance Standards

All three programs now use national performance standards, uniformly applied to all state agencies: 3 percent payment error rates for AFDC and Medicaid, 5 percent for Food Stamps. Various critics argue that state-specific circumstances that are beyond the control of welfare administrators can affect payment error rates. Such factors include caseload sizes, unemployment rates and general economic conditions in the states, demographic characteristics, programmatic complexity, policy choices made by governors and state legislatures, incidence of cases with earned income, and different indications of urbanization—all of which can affect caseload mixes and error rates. (In a related vein, these critics also note that the incentives created by QC-based disallowances under AFDC discourage states from operating the "unemployed parent" (UP) part of AFDC, which may be done at state option. AFDC-UP cases have a greater incidence of earnings, and therefore are more error prone (see American Public Welfare Association, 1986; Sacramento County, 1986; National Council of State Human Service Administrators, 1986, 1987; Florida, State of, 1986; National Governors' Association, 1986; Wyoming, State of 1986; California, State of, 1986).

From this perspective, either national standards should be adjusted for certain kinds of state-specific characteristics or standards should be set on a state-by-state basis (see, especially, California, State of, 1986; National Council of State Human Service Administrators, 1986, 1987; Florida, State of, 1986; National Governors' Association, 1986; Wyoming, State of, 1986.) GAO (1986d) notes, however, that the effects of such external influences on error rates may not be easily distinguishable from the effects of poor management.

From a different perspective but making a similar point, other critics have suggested that the performance capabilities of state welfare agencies necessarily vary and that the agencies should be financially accountable only to levels of performance for which they are capable. Others have argued specifically that if sanctions are to be used, then standards should be set at levels that a significant number of states could realistically reach with dedicated effort: the National Council of State Human Service Administrators (1986); Desmatics (1986); Center on Budget and Policy Priorities (1986); National Governors' Association (1986). Explicitly, they argue that standards that penalize a large number of state agencies are necessarily too stringent.

Westat Research, Inc. (1971, 1973), drawing on a common practice in quality control applications elsewhere, recommended that AFDC performance standards be state-specific, based on each state's record of performance. A system structured with appropriate incentives for improvement, and with a long-range national target or goal, could use such state-specific standards. Westat suggested that the national goal should represent a target achievable by a large number of states with sustained and diligent effort. Such an approach, if taken, would obviate any need to model and adjust national standards for state-specific differences.

Implicit in such criticisms are two principles. First, performance standards should reflect state characteristics, whether by adjustment of national standards for certain state-to-state differences or by setting state-specific standards that reflect each state agency's performance capability. Second, any standards that are set should reflect realistic thresholds; standards based on varied state capabilites create targets that can be reached with dedicated efforts.

Variation in National Payment Error Rates in the Three Programs

The Center on Budget and Policy Priorities (1986) examined differences in the national average payment error rates achieved for AFDC, Medicaid, and Food Stamps, in order to address the question of whether the Food Stamp Program was more poorly managed than the other two programs. The question arises because of higher national payment error rates for Food Stamps, and because

states do not share in the benefit costs for this program. The Center concluded that a combination of differences in caseloads and differences in definitions of errors that are measured fully explains the differences observed.

The Center's analysis found that the Food Stamp caseload has several characteristics that make it more error-prone than the AFDC caseload: it has three times the proportion of cases with earnings; it has about six times the fraction with unearned income, an indirect indicator of assets and another source of errors; it turns over more quickly; and it contains more two-parent families. The Center argues that Food Stamp error rates should be expected to be higher than those for AFDC, and by at least as much as the differences that do exist.

The Center on Budget and Policy Priorities (1986) also argues that the Medicaid error rates should be expected to be the lowest, which they are, because Medicaid does not count technical errors and has two other features that would lead to a lower expected error rate. First, it uses a "spend-down" provision for errors related to excess resources. If a recipient is formally ineligible for participation because of excess resources, an entire AFDC grant or Food Stamps issuance is deemed to be in error, regardless of the value of the resources in question. Under Medicaid, however, if the resources are less than the benefits paid, the amount of the error is the value of the resources—the amount that the recipient would have had to "spend down" in order to be eligible for the remaining benefit payments. Second, Medicaid payment error rates accumulate only payments made on behalf of ineligible recipients. There is no counterpart to AFDC overpayments or Food Stamps overissuances to eligible recipients. Under Medicaid, overutilization of care by recipients and excess billing by providers do not enter into QC-based payment error rates.

Distorting Effects of Consequences from Limited Measures of Performance

The QC systems for AFDC, Medicaid, and Food Stamps focus on and emphasize overpayments, including payments to or on behalf of ineligible recipients. As noted above, some critics charge that this narrow emphasis encourages performance to reduce these

kinds of payment errors, which—particularly with substantial consequences tied to measured performance—leads to the neglect of performance, or even dysfunctional program responses, in other important areas. The argument that QC induces "case churning" (described above) is a prime example. Examples of program objectives to which states might be held accountable include timeliness, nutritional benefits, improved health, child support collections, overpayment recoupments, placement in jobs or training, and self-sufficiency (see Brodkin, 1986; National Council of State Human Service Administrators, 1987; Florida, State of, 1986; and National Governors' Association, 1986).

Agency and Client Causes of Error

State officials argue that they should not be held accountable for clients' errors (see American Public Welfare Association, 1986; Sacramento County, 1986; Coalition of California Welfare Rights Organizations, 1986; National Governors' Association, 1986; New Jersey, State of, 1986; California, State of, 1986). This is a complex issue. Agency actions or systems may influence client errors (U.S. General Accounting Office, 1986d) through, for example, training of workers, particularly in interviewing techniques; differences in workers' skill or educational levels; clear forms and communications with recipients and applicants; and the use of computerized matches of bank records and wage reports. However, some recipients, for whatever reasons, do misreport such things as income, assets, shelter costs, and persons living in a household. It is not impossible, for instance, for a recipient to own but not know of a resource such as an insurance policy or a bank account. Moreover, such lack of knowledge might mean that the recipient did not have effective access to it, in which case it is not effectively available, and countable, income for determining need of public assistance.

Food Stamp Sanctions Against Administrative Costs

For AFDC and Medicaid, the penalty for poor performance, as measured under their QC systems, is a disallowance of the federal funds expended for benefits in excess of the those allowed within the established performance standard. The penalty to the states is a dollar-for-dollar reduction of the federal share of the

federal funds misspent. Under the Food Stamp Program, however, since there is no state share of benefit costs, the sanctions are against the federal share of administrative costs. The U.S. General Accounting Office (1986d) notes the inconsistency and lists pros and cons to establishing sanctions against benefit costs or against administrative costs. Another GAO official has argued for the use of Food Stamp sanctions tied dollar-for-dollar to the federal benefit costs associated with overpayment issuance errors (see Crowley 1985, 1986; U.S. General Accounting Office, 1985a). The Center on Budget and Policy Priorities (1986) argues that the intent in sanctioning administrative costs for Food Stamps is to penalize the program administrators, that is, the states, and not the beneficiaries.

A state's potential dollar liability for sanctions under the Food Stamp Program is a function of its total administrative costs for the program. But those costs are not easily identified. Since the same state (or local) agency administers both programs, the agencies have cost-allocation plans for allocating between Food Stamps and AFDC (and usually other programs as well) the proper shares of the total costs of administration. The Center on Budget and Policy Priorities (1986) argues that an OMB ruling on such cost allocations has had the effect of shifting a greater proportion of states' charges for administration to Food Stamp, and from AFDC, administration. This change alone, without any tie to measured performance, raises the potential liabilities that state agencies may incur under Food Stamps. The National Council of State Human Service Administrators (1987) argues for sanctions tied to benefit costs and not linked to administrative costs.

Another report cites the peculiar incentives created by the "step function" in the Food Stamp QC sanctions (Desmatics, 1986). Penalties change only at each whole percentage of estimated payment error. The financial consequences of a small shift in the estimated error rate can be very large, if that shift crosses a threshold that marks a step in the sanction formula, for example, from 8.01 to 7.99 percent. Yet there are no consequences to a large shift within a single step, for example, changes up or down anywhere between 8.9 to 8.1 percent.

PROCEDURAL ISSUES

Critics argue that certain QC procedures do not support uniform application or implementation of Food Stamp policy, create undue management hardships, and threaten due process protections in hearings and appeals.

Reviews of State Plans

Within state welfare agencies, state plans constitute the basic operating guide. Usually by reference, these plans include the manuals and guidelines for workers in local offices. In AFDC and Medicaid, QC re-reviewers at federal regional offices use both state plans and federal law and regulations to make their determinations for QC. In Food Stamps, federal re-reviewers use only federal policy documents (laws, regulations, and other guides from FNS), and since 1982, FNS officials have not reviewed and approved state plans. Since the states still use the state plans as the controlling documents in making their QC determinations, and the federal government uses their document, there is an inevitable source of conflict.

Discrepancies between state plans and federal guidance, in all three programs, have historically given rise to disputes between state and federal officials concerning interpretations of eligibility and benefit criteria. State officials have argued that federal agencies should review and approve state plans, so that in reviewing cases for QC determinations, all reviewers—federal and state—work from the same set of criteria (see National Council of State Human Service Administrators, 1987).

Administrative Grace Periods for Implementation of Program Changes

Currently, federal AFDC practice permits a time lag for implementation of program changes. Only after this administrative "grace period" are state agencies held accountable for those changes in the QC process. This lag is intended to give the state agencies time to make needed changes in their state plans and manuals; to print and disseminate those changes; to train workers, supervisors, and program directors at all levels in the implementation of such changes; and to work out the procedural bugs that often

accompany such changes. For Food Stamps, FNS does not permit such lags before counting errors from such changes in QC reviews. One argument for not permitting such time lags is that major changes are issued with a built-in lead time, but states argue that such lead time rarely if ever is sufficient for the administrative job required (see National Council of State Human Service Administrators, 1987; Sacramento County, 1986; Center on Budget and Policy Priorities, 1986).

Region-to-Region Differences

State officials claim that region-to-region differences in interpretation of federal policy and guidelines results in differential treatment of states in the QC process and that such differences result in QC findings of different stringency (Florida, State of, 1986; National Governors' Association, 1986). The U.S. General Accounting Office (1986b) and the Inspector General of the U.S. Department of Agriculture (1986a, 1986b, 1986c, 1986d, 1986e) documented regional variations in the calculation of Food Stamp error rates by state. These differences were of two sorts: one simply involved errors in arithmetic, in the computation of error rates from the aggregated reports of QC case findings; the other involved the inappropriate weighting of strata in some states with stratified QC samples. Either kind of error can affect a state's liability for sanction as a result of the official payment error rate estimate. The Inspector General recommended greater use of second party reviews of regional office work; the GAO noted that computerization of such calculations, with appropriate testing of the programming, would eliminate such regional differences.

Nonindependence of Federal Re-Reviews

Federal reviewers use the material in the state case review files to determine what further action, if any, is required. Therefore, it is not strictly an independent review. State officials contend that such nonindependence leads to higher estimates of error, because any reviewer presented with data in which findings—in this case errors—have already been determined will be predisposed to find more (see California, State of, 1986; U.S. General Accounting Office, 1986b). Westat Research, Inc. (1971) recommended that independent federal re-reviews be done, at least periodically

on a small scale, to properly assess state and federal reviewers' performance.

Proposed Revisions to FNS Arbitration and Appeals Procedures

Early in 1986, FNS proposed to "streamline" its arbitration and appeals procedures for the resolution of state-federal QC differences, ending the arbitration process with the findings of the regional office director. Currently, disputes may be pursued to the national FNS office, and about 20 percent of such disputes by states are found in states' favor. FNS also proposed to alter the appeals process; an appeal of a sanction liability would no longer be heard by a Grants Appeals Board consisting of USDA staff outside of FNS; instead, a single person named by the Secretary of Agriculture—who could be an FNS employee or official—would hear such appeals.

FNS reportedly has received numerous objections to the proposed revisions. The House subcommittee with jurisdiction over the Food Stamp Program held hearings on the proposed revisions. All who testified spoke against the proposed changes. State officials noted that the effect would be to deprive them of two processes that historically have found enough fault with FNS determinations to make a difference in payment error rate estimates and sanction liabilities. Others argued that the arbitration and appeals process now in place affords due process protection by the use of appropriately independent persons to settle disputes between interested parties and that the FNS proposals would eliminate that protection (see Center on Budget and Policy Priorities, 1986).

REFERENCES

This list includes sources cited in this chapter and other materials the panel used in reviewing the Food Stamp Program. Those references not cited are noted by an asterisk(*).

*American Public Welfare Association
n.d. Chronology of Federal Quality Control in AFDC, Medicaid, and Food Stamps, 1952 through December 1982. American Public Welfare Association, Washington, D.C.

American Public Welfare Association
1986 *Briefing Book on Quality Control and Fiscal Sanctions in the AFDC, Medicaid, and Food Stamp Programs.* Washington, D.C.: American Public Welfare Association.

*Arizona, State of
1984 Training Guide and Record of Progress for Eligibility Reviewers. Program Evaluation Section, Family Assistance Administration, Arizona Department of Economic Security.

Blischke, W.R.
1985 On the Use of Regression Analysis in Estimation of Error Rates in AFDC Payments. Unpublished manuscript. Van Nuys, Calif.

Bode, J.
1986 Remarks to the Panel on Quality Control Programs, National Research Council, July 11, 1986. Food and Nutrition Service, U.S. Department of Agriculture, Washington, D.C.

Brodkin, E.Z.
1986 *The False Promise of Administrative Reform: Implementing Quality Control in Welfare.* Philadelphia: Temple University Press.

Brodkin, E., and M. Lipsky
1983 Quality control in AFDC as an administrative strategy. *Social Service Review* (March):1-34.

California, State of
1986 Issues Raised for Consideration in the Study of the Quality Control System. Unpublished document. Department of Social Services, Sacramento, Calif.

*California, State of
1987 Letter from Linda S. McMahon regarding transmittal of issues which various jurisdictions within California raised as worthy of examination by the National Research Council. Department of Social Services, Health and Welfare Agency, Sacramento, California.

*Center on Budget and Policy Priorities
1985 A Comparison of Food Stamp, Medicaid, and AFDC Error Rates: Is the Food Stamp Program Poorly Managed? Center on Budget and Policy Priorities, Washington, D.C.

*Center on Budget and Policy Priorities
1985 Does the Food Stamp Error Rate Overstate the Loss to the Federal Government Through Errors? Center on Budget and Policy Priorities, Washington, D.C.

Center on Budget and Policy Priorities
1986 Issues Related to Quality Control and Error Rate Sanction Systems Particularly in the Food Stamp Program. Unpublished manuscript. Center on Budget and Policy Priorities, Washington, D.C.

*Center on Social Welfare Policy and Law
1986 Letter from T.J. Casey and M.R. Mannix regarding background material on AFDC QC and/or on denials of AFDC benefits. Center on Social Welfare Policy and Law, New York.

*Center on Social Welfare Policy and Law
1986 *Quality Control: A Disaster for the Poor.* Initial Statement of Houston Welfare Rights Organisation and Michigan Welfare Rights Organisation to the Panel on Quality Control for Family Assistance Programs, National Research Council. Center on Social Welfare Policy and Law, New York.

*Center on Social Welfare Policy and Law
1986a Quality Control and the "Churning Crisis," Part I. Unpublished manuscript. Center on Social Welfare Policy and Law, New York.

Center on Social Welfare Policy and Law
1986b Quality Control and the "Churning Crisis," Part II. Unpublished manuscript. Center on Social Welfare Policy and Law, New York.

Coalition of California Welfare Rights Organisations, Inc.
1986 Letter from K. Aslanian to express interest in the quality control issue. Coalition of California Welfare Rights Organisations, Inc.

Cochran, W.G.
1977 *Sampling Techniques*, 3rd ed. New York: John Wiley and Sons, Inc.

*Connecticut, State of
1986 AFDC Quality Control Narrative Analysis: April - September 1985. Department of Income Maintenance, Hartford, Connecticut.

*Connecticut, State of
1986 Letter from Patricia Day, Director, Management Planning and Evaluation to Charles Vann regarding AFDC quality control. Department of Income Maintenance, Hartford, Connecticut.

*Covington & Burling
1986 *Appellants Appeal File*, Vol. IV. Affidavit of Dr. William Fairley and supporting materials. In regard to fiscal year 1981 Medicaid quality control disallowances, before the Departmental Grant Appeals Board, U.S. Department of Health and Human Services. Washington, D.C.: Covington & Burling.

Covington & Burling
1986 *Appellants' Opening Brief on Common Issues.* In regard to fiscal year 1981 Medicaid quality control disallowances. Washington, D.C.: Covington & Burling.

Crowley, B.P.
1985 Statement Before the House Committee on Agriculture on the General Accounting Office's Overview of the Food Stamp Program, April 17, 1985. U.S. General Accounting Office, Washington, D.C.

Crowley, B.P.
1986 Statement Before the House Committee on Agriculture on the USDA and Other Proposals to Reduce Food Stamp Program Expenditures, April 6, 1985. U.S. General Accounting Office, Washington, D.C.

*Delaware, State of
1986 Delaware Error Reduction Project: Management Assessment Report. Division of Economic Services, Department of Health and Social Services, New Castle, Delaware.

Desmatics, Inc.
1986 *A Statistical Evaluation and Review of the Food Stamp Quality Control Sanction Imposed on Pennsylvania for FY84.* C.A. Mauro and D.E. Smith, authors. Technical Report No. 127-3. State College, Pa.: Desmatics, Inc.

DMA Corporation
1985 Statistical Evaluation of Federal Regression Methodology. Unpublished supplement report. DMA Corporation, Palo Alto, Calif.

*East Orange, City of
1986 How Does Quality Control Limit the Creativity of a Local Welfare Administrator from Better Serving His/Her Constituents. Report to the White House Domestic Policy Council Advisory Commission on Intergovernmental Relations on Welfare Reform. W. Bruce Potter, Director, East Orange Department of Public Welfare, East Orange, New Jersey.

Fairley, W.B.
1986 Affidavit of William B. Fairley in the Matter of Connecticut Department of Income Maintenance. State Food Stamp Appeals Board, U.S. Department of Agriculture, Washington, D.C.

*Florida, State of
n.d. Quality Control Training Handbook. Tallahassee, Florida.

*Florida, State of
n.d. Third Party Review and Training the New QC Analyst. Tallahassee, Florida.

*Florida, State of
1986 Letter from William J. Page, on quality control regarding independent study of the QC system utilised to assess the performance of the Food Stamp Program. Department of Health and Rehabilitative Services, Tallahassee, Florida.

Florida, State of
1986 The Quality Control System. Unpublished manuscript. Department of Health and Rehabilitative Services, Tallahassee, Florida.

*Georgia, State of
1986 Error Rate Study: AFDC, Medicaid and Food Stamp Program. Department of Human Resources, Atlanta, Georgia.

Gilford, D.M., D.P. Affholter, and L. Ingram, eds.
1983 *Family Assistance and Poverty: An Assessment of Statistical Needs.* Committee on National Statistics. Washington, D.C.: National Academy Press.

Illinois, State of
1986 Letter to John Neter from G.L. Coler, September 10, 1986. Illinois Department of Public Aid, Springfield, Illinois.

*Michigan, State of
1983 The Impact of Monthly Reporting in Michigan, Vol. 1. Planning and Evaluation Division, Office of Planning, Budget, and Evaluation, Michigan Department of Social Services.

*National Council of State Human Service Administrators
1986 Management and Quality Control Committee Meeting, Washington, D.C., June 3.

*National Council of State Human Service Administrators
1986 Resolution Concerning FNS Regulations to Restructure the Administrative Review and Arbitration Processes. National Council of State Human Service Administrators, Washington, D.C.

National Council of State Human Service Administrators
1986 The States' Views of the AFDC, Medicaid and Food Stamp Quality Control Systems. Talk presented to the Panel on Quality Control of Family Assistance Programs, National Academy of Sciences, July 11, 1986. National Council of State Human Service Administrators, Washington, D.C.

National Council of State Human Service Administrators
1987 Report of the APWA Task Force on an Alternative System for Performance Measurement. Unpublished document. National Council of State Human Service Administrators, Washington, D.C.

National Governors' Association
1986 Remarks by A. Smith Before the Panel on Quality Control of Family Assistance Programs, National Academy of Sciences, July 11, 1986. National Governors' Association, Washington, D.C.

*National State QC Directors' Association
1986 Programs Administered by the Ocean County Board of Social Services. National State QC Directors' Association, Washington, D.C.

National State QC Directors' Association
1987 Letter from John Petraborg to John Neter regarding transmittal of Report of the APWA Task Force on An Alternative System for Performance Measurement. National State QC Directors' Association, Washington, D.C.

New Jersey, State of
1986 Letter from D. Altman, Department of Human Services, Trenton, New Jersey.

*New York, State of
1984 Administrative Closings of New York City Public Assistance Cases. Department of Social Services, Albany, New York.

*New York, State of
1985 A Report on the Expansion of the Enhanced Quality Assurance System. Leonard Chapla, Director, Office of Quality Assurance, New York.

*Ohio, State of
1986 Correspondence and back-up materials on good faith waiver request. Coding error and problems with regression formula in Medicaid cases for Ohio. Department of Human Services, Columbus, Ohio.

Sacramento County
1986 Quality Control Issues Requiring Study by the Panel on Quality Control of Family Assistance Programs, National Research Council. Sacramento County Department of Social Welfare, Sacramento, Calif.

*Sacramento County
1986 Toward Excellence in Human Service Programs. An issue paper on the Aid for Families with Dependent Children Quality Control System. County Welfare Directors' Association, Sacramento, California.

*South Carolina, State of
n.d. Medicaid Quality Control Training Guide. Medicaid Quality Control Unit, Quality Control Division, Department of Social Services, Columbia, South Carolina.

*South Carolina, State of
1986 Local office procedures. Department of Social Services, Charleston, South Carolina.

*Texas, State of
1986 Narrative Section, Good Cause Waiver Request. Office of Programs, Texas Department of Human Services, Austin, Texas.

U.S. Department of Agriculture
1986a Audit Report: Food and Nutrition Service, Northeast Regional Office. U.S. Department of Agriculture, Washington, D.C.

U.S. Department of Agriculture
1986b Audit Report: Food and Nutrition Service, Southeast Regional Office. U.S. Department of Agriculture, Washington, D.C.

U.S. Department of Agriculture
1986c Audit of the Food Stamp Program Quality Control Error Rate Reduction System in California. U.S. Department of Agriculture, Washington, D.C.

U.S. Department of Agriculture
1986d Audit of the Food Stamp Program Quality Control Error Rate Reduction System in Michigan and Minnesota. U.S. Department of Agriculture, Washington, D.C.

U.S. Department of Agriculture
1986e Nationwide Audit of Quality Control Error Rate Reduction System. U.S. Department of Agriculture, Washington, D.C.

U.S. Department of Health and Human Services
1986 An Alternative and Less Costly Method to Measure Error Under the Quality Control System for the Aid to Families with Dependent Children (AFDC) Program. Unpublished manuscript. Office of Inspector General, U.S. Department of Health and Human Services, Washington, D.C.

U.S. General Accounting Office
1980 *Better Management Information Can be Obtained from the Quality Control System Used in the Aid to Families with Dependent Children.* GAO/HRD-80-80. Washington, D.C.: U.S. General Accounting Office.

U.S. General Accounting Office
1985a *Quality Control Error Rates for the Food Stamp Program.* GAO/RCED-85-98. Washington, D.C.: U.S. General Accounting Office.

U.S. General Accounting Office
1985b *Quality Control Error Rates for the Food Stamp Program.* GAO/RCED-85-98. Washington, D.C.: U.S. General Accounting Office.

U.S. General Accounting Office
1986a *Food Stamp Program: Refinements Needed to Improve Accuracy of Quality Control Error Rates.* GAO/RCED-86-195. Washington, D.C.: U.S. General Accounting Office.

U.S. General Accounting Office
1986b *Food Stamp Program: Statistical Validity of Agriculture's Payment Error-Rate Estimates.* GAO/RCED-87-4. Washington, D.C.: U.S. General Accounting Office.

U.S. General Accounting Office
1986c *Food Stamp Program: Restoration of Improperly Denied or Terminated Benefits.* GAO/RCED-87-51. Washington, D.C.: U.S. General Accounting Office.

U.S. General Accounting Office
1986d *Managing Welfare: Issues and Alternatives for Reforming Quality Control Systems.* GAO/HRD-86-117BR. Washington, D.C.: U.S. General Accounting Office.

Urban Institute
1985 *The Effects of Legislative Changes in 1981 and 1982 on the Food Stamp Program.* Washington, D.C.: Urban Institute.

Westat Research, Inc.
1971 *Evaluation of Quality Control System in Public Assistance: Report on Phase II.* Rockville, Md.: Westat Research, Inc.

Westat Research, Inc.
1973 Evaluation of the Tolerance Levels Established for the Quality Control System in Public Assistance. Unpublished report. Westat Research, Inc., Rockville, Md.

Westat Research, Inc.
1986 A Statistical Evaluation of AFDC Quality Control. Draft. Westat Research, Inc., Rockville, Md.

Wyoming, State of
1986 Letter from J.E. Robinson to H. Hansen regarding quality control error rates. Department of Health and Social Services, Cheyenne, Wyoming.

5
Lessons from Quality Improvement Experts

For more than 50 years, experts have systematically applied statistical methods to problems of controlling product and service quality. The contemporary field of "quality control," however, extends beyond the application of specific statistical methods, to a larger view of quality and its management. In this chapter the panel draws on lessons from the experts in order to assess the family assistance QC systems. This consideration of the tools and perspectives of the experts is selective, mindful of the special environments in which family assistance programs operate. The panel may have more to say about practical applications in its second report.

The larger view of quality experts is reflected in the diversity of terms that describe this field today—terms like total quality control, company-wide quality control, quality assurance, quality improvement, and zero defects—which creates problems for any discussion about "quality control." For the sake of coherent presentation, the panel adopted the following usage: quality control refers specifically to the systems known by that name in the Food Stamp, AFDC, and Medicaid Programs; statistical quality control refers to specific statistical applications for process control (which are described below); quality improvement refers to systems that

take on a larger view, more than simply process control. All other uses of those and similar terms are set off in quotes and used when necessary in order to make a particular connection or citation to work identified as such in the quality field.

The purpose of this chapter is to introduce readers to possibilities for the uses of quality improvement systems and tools in family assistance programs. It is not to provide a tutorial on quality improvement, nor a set of concrete, specific Food Stamp applications. A Food Stamp example is noted, to make connections between the world of assistance programs and the perspectives and tools of quality improvement. But successful applications in any system will depend substantially on the circumstances: such things as the environment in which a program or process operates, the degree of isolation or interdependence to other programs or processes, the resources and amount of flexibility available to various participants in the system, the objectives to be achieved by the product or services, and the knowledge available about the purposes and problems of the system.

QUALITY IMPROVEMENT: A PROCESS ORIENTATION

Quality improvement systems focus on processes. Since errors or imperfections may enter into a process at any step—from initial design to ultimate inspection and use of the finished product or rendered service—quality improvement systems are intended to identify, reduce, eliminate, and ultimately prevent, such errors.

Quality improvement systems involve three elements: process control, process improvement, and product or service improvement. Process control focuses on the operational steps in a process, and aims at "statistical control" of that process, that is, ensuring that production operates at the full capability of the process. Process improvement focuses on making the process work better, improving the capability of the process, which often requires changes in the process itself, and process redesign. Product or service improvement focuses on the effectiveness of the product or service, for the purposes of or uses intended for the product or service. Improved products or services often require improved design or redesign of the product or service itself.

The use of statistical tools for quality control and "quality assurance" has a longer history in manufacturing than in service

industries. "Quality control" texts, for instance, often contain many more manufacturing examples than service examples (see, e.g., Grant and Leavenworth, 1980). Providers of services often claim that services, and especially public services, represent more complex systems than do those for manufacturing. They have much more variation and complexity; they are subject to external disturbances; and they have to meet a wider variety of potentially conflicting objectives, and are accountable to a wider set of actors and interests. Hence, they argue, statistical quality control is not widely applicable in many service delivery environments.

In response, the quality improvement perspective notes that all organized activity, whether for manufacturing or service delivery, consists of processes. Any large collection of processes serves multiple purposes, and different purposes can often work against each other. At this level, all organized processes are complex, although the kind of complexity varies from specific process to process. The larger and more complex the overall process, the greater that complexity and the greater the need for weighing trade-offs among different and sometimes competing objectives.

The manufacture of a car, for instance, involves an elaborate set of processes that could be the focus for quality improvement systems: parts need to be produced and assembled to specifications, but such manufacturer's objectives as market share and short- and long-term profitability also require a focus on process and product improvement. Paying on credit with credit cards, a classic service operation, requires some system of quality control to ensure accurate billing to the appropriate users in the process, and may also require process improvement and service improvement systems in order to assure sufficient volume of credit use and cash flow to keep the company in business.

Similarly, the delivery of Food Stamps, from initial inquiry through eligibility determination to final issuance of an appropriate allotment of coupons, is a complex process. Some form of process control should be used to ensure that the correct allotment goes to the right individual in timely fashion. A quality improvement system would also incorporate process improvement and service improvement efforts, however, to promote improved efficiency and effectiveness in the achievement of this and other desired or required program objectives.

Large-scale processing systems, whether for products or services, can be organized or subdivided into smaller, interrelated sets of processes which can be by function, for instance, or by specificity of processing detail within organizational units. Process control, process improvement, and product or service improvement apply at all of these subprocess levels as well as at the largest aggregation of processes.

Process control, process improvement, and product or service improvement can be pursued at any point in the system—for instance, in the manufacture of car doors, or more finely yet, in the manufacture of car door handles. The measures of and studies for quality improvement of door handles, for instance, will be limited or shaped by the constraints of the design of the car door and the car body, but may in turn suggest design changes to the car door or car body.

Similarly, the perspectives and tools of quality improvement can be applied to the Food Stamp Program. Like the door handle, the process of eligibility determination will be limited or shaped by the constraints or requirements of federal policy. But that process may in turn suggest design changes to federal policy, either for eligibility in its narrowest sense or in the service of other program objectives that could be addressed at the point of eligibility determination.

PROCESS CONTROL THROUGH CONTROL CHARTS

Traditional statistical quality control focuses on what statisticians call process control, at a very discrete, detailed level at which characteristics of the items (or transactions) produced can be measured against specifications. Statistical quality control was born from an important insight, attributed to Walter Shewhart in 1925; in the words of Grant and Leavenworth (1980:1):

> Some stable "system of chance causes" is inherent in any particular scheme of production and inspection. Variation within this stable pattern is inevitable. The reasons for variation outside this stable pattern may be discovered and corrected.

Causes of inevitable variation within the stable pattern cannot be eliminated or corrected without changes in the system or process. Such causes of variation in performance are known as "common causes"—that is, they are common to the process

as designed, not attributed to special or idiosyncratic problems. Causes that are not inherent in the process, or not a part of the inevitable variation within a stable pattern, are known as "assignable" causes. Corrective action may be taken to eliminate assignable causes, once identified, without changing the system or process. "A process operating in the absence of any assignable causes of erratic fluctuations is said to be in *statistical control*" (Bowker and Lieberman, 1959:378).

It follows, then, that when any processing system is working up to its capability, measured performance will still show some variation. A processing system's capability will, even under the best of circumstances, produce dispersion in performance around an average level. That is, not every item or transaction will be perfect.

However, events or problems may arise that cause the quality to decline significantly from the performance capability of the system. Machines or machine parts may break, or settings may slip drastically in an automated process; a new worker in a local welfare office may introduce errors of magnitude and frequency sufficient to alter the service quality for the whole office. To recognize the difference between the assignable causes of variation in performance and common causes, certain measurements must be taken, baselines determined, and ongoing measurement systems established.

One mechanism for such a measurement system is a control chart, which can be used in the manufacture of door handles as well as the processing of Food Stamp coupon allotments. In its simplest form, a control chart shows the quality measurement of the output—items or transactions—from a single process. It does so by sampling the items or transactions from that process over a period of time and physically charting the quality measurement of each sample, usually with an easily read graph.

A control chart shows the sample measures and average performance of the process as well as established limits, called control limits. When sample estimates fall outside the control limits, the control chart indicates that the process may be out of statistical control, a clear signal to look for problems.

Control limits often are set at "3 sigma," plus or minus three times the standard deviation of performance measured (Bowker and Lieberman, 1959:384-385), which is usually estimated from

samples of performance measures. (In many control charts, the range between highest and lowest measures taken in each sample may be used, rather than the standard deviation.) When a system is in control, measurements that fall outside of these control limits can be expected only 27 times in every 10,000 measurements taken, roughly 3 in 1,000. A single measurement outside of the control limits, therefore, is likely but not certain to indicate the introduction into the process of a problem with a specific or assignable cause.

Control limits may be established other than by the "3 sigma" rule. For instance, if the cost of hunting for a specific cause of a problem is extremely expensive or disruptive to the process, wider control limits may be set. On the other hand (Wadsworth et al., 1986:223):

> if the cost, or effort, or time to check a process to conclude that no change has occurred is low, then a larger risk may be warranted, by using a lower multiple of sigma (say, 2.5 or 2.0) for control limits.

Many control charts require only simple arithmetic for their construction and use. Most can be started after a small number of samples and used effectively with very small samples. Bowker and Lieberman (1959:383) report that 25 subgroups (or samples) typically should be used to start a control chart, and that subgroups (or sample sizes) of 4 or 5 seem to work quite well in manufacturing applications. Wadsworth et al. (1986:223) echo that assessment, but also note specific considerations that can alter the sizes of the subgroups needed: if it is important to detect small changes in the process, or, in effect, to tighten your control limits, then larger subgroups are needed; if the cost of inspecting or testing the samples is high, then smaller subgroups may be appropriate, even as small as 1. Montgomery (1985) devotes an entire chapter in his quality control text to the economic design of control charts.

The key to the productive use of control charts, however, is the basis for grouping items or transactions for sampling and measurement for the control chart. The groups should offer the maximum opportunity for differences or variations in performance between groups and minimum variations within groups (Grant and Leavenworth, 1980:119). The rest of this section presents an example of how a control chart could be used in the Food Stamp Program, focusing on a measurable outcome.

A control chart could be established for monitoring the timeliness of disposing of applications by eligibility workers. One variable entered on the chart would be the average number of days to disposition. If four or five dispositions per day can be sampled from a worker's output, then a control chart could be established for the worker within about 5 weeks. The chart could be revised over time, depending on the measured performance. If a worker's processing output is too small or too irregular for sampling daily, then a higher level of the process, such as a work group or local office, or a longer period of time, may be a more appropriate focus for a control chart. If daily samples are taken for a work group, then again about 5 weeks would be required to establish the chart. If a worker's output can be sampled only weekly, then about 6 months would be required to establish the chart.

For a control chart, two measures need to be taken from each set of the sampled dispositions: the average number of days to disposition and the range between the greatest and least number of days within each sample. The average number of days to disposition can then be calculated across samples, as well as the average range. From standard charts, tables, or texts, upper and lower control limits—UCL and LCL—can be calculated (see, e.g., Wadsworth et al., 1986). From that point on, the performance of the worker (or work group or local office) can be tracked and brought under statistical control. (See Figure 12 for an illustration of just such a control chart.)

Data from control charts can be summarized periodically (say, weekly) for use by the next higher management level. Assuming that each control chart shows a process in control, discrepancies between these charts—across workers, work units, or local offices—would suggest possibilities for process improvement. With or without comparative differences, if managers deem the process to take too long or to be unduly variable, either process or service redesign might be undertaken. This in turn might require additional study.

As an adjunct to control charts, supervisors may also wish to track backlogs, or flow into and out of the backlog, of pending applications. Even though the process is apparently in control, a rising backlog could signal forthcoming trouble. A falling backlog could indicate an opportunity for the reassignment of one or more workers to other functions. And a sudden increase in applications,

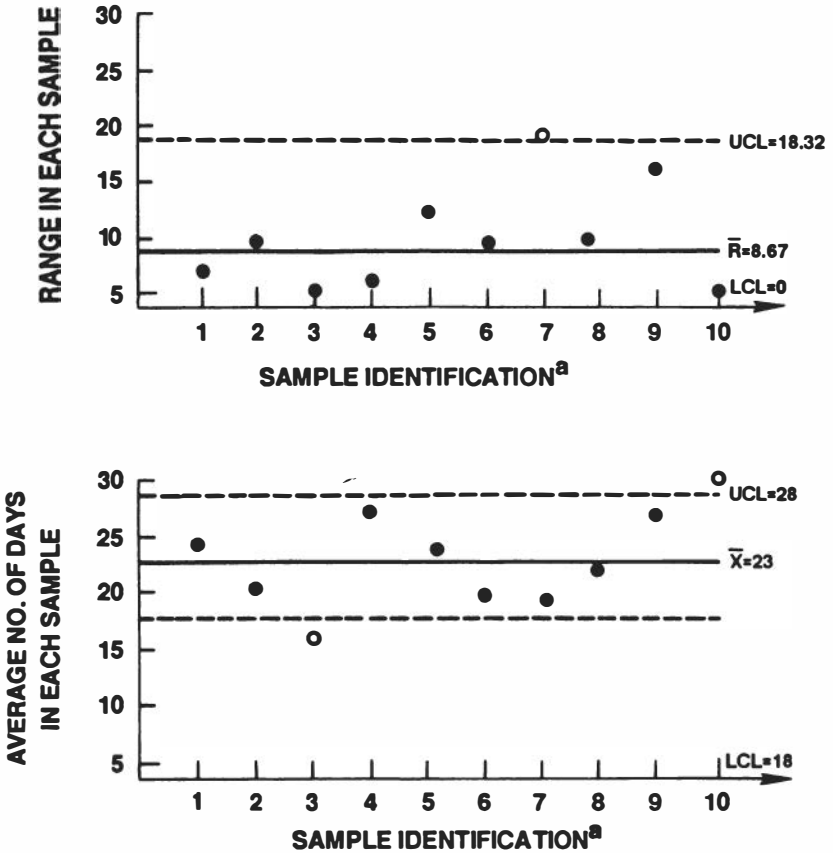


FIGURE 12 Hypothetical Control Chart for Timeliness of Disposition of Food Stamp Applications

Note: This example assumes that the samples and measures needed to establish the chart have been done. The assumed average time to disposition is 23 days, with an average range over the samples (longest less shortest disposition time in each sample of 5 items) of 8.67 days.

[a]The "sample identification" shows the sample for which the data are plotted. For a specific worker or work group, the sample identification could be the day or week on which the sample of 5 items was drawn.

or the unexpected and unplanned absence of an eligibility worker, could throw the process out of control.

To set up and use a control chart, one must identify a discrete and bounded process that ends in a measurable outcome that has been determined to be of importance. The choice of what units or for whom to create charts should reflect some likely relationship to the performance to be charted and controlled. And such choices should be sensitive to the costs of and benefits to be derived from the use of these tools, rather than others, to control quality.

TOOLS FOR MONITORING QUALITY: NEITHER CONTROL NOR IMPROVEMENT

Two specific tools have been used to monitor performance against criteria concerning quality: sampling inspection and quality audits. Both may promote process control and quality improvements, but neither one directly controls or improves a process. Both work most effectively when used in concert with tools for process control.

Sampling Inspection

Unlike a control chart, sampling inspection is an accept/reject tool for use on batches of intermediate or final product (see e.g., Bowker and Lieberman, 1959; Grant and Leavenworth, 1980). Sampling inspection uses probability theory to design sampling plans that rely on the inspection of samples of items from "lots" or "batches": decision rules that are incorporated in the design dictate the acceptance or rejection of the entire lot or batch on the basis of the results of the sampling inspection.

Many different designs have been developed for sampling inspection, often called acceptance sampling. Standards for sampling inspection designs are maintained by organizations such as the American Society for Quality Control (ASQC) and the American National Standards Institute (ANSI).

Sampling inspection is used by buyers in order to assure an "average quality level" among all the items in all the batches that are accepted. When written into contracts by buyers, sampling inspections can create pressures on manufacturers to control and improve their manufacturing processes. The consequences of rejection of lots may include tighter sampling inspection designs

or on-site monitoring of the supplier's production system by the buyer's inspectors. But sampling inspection alone cannot correct performance problems or promote improved production processes.

Even though it may be difficult to envision applications in the services field, sampling inspections have been used to review the work of claims processors in insurance. Analogous applications in the Food Stamp Program can be envisioned. Most of the work of eligibility workers is inspected by someone else, usually a supervisor; some states and localities require a 100 percent supervisory review and sign-off. As a practical matter, such reviews often are done on an ad hoc basis, focusing on only selected items, workers, and decisions. Sampling inspection offers a formal tool by which supervisors (whether in a local office or a state operation) could, with a sound basis in probability sampling and decision rules, design their inspections of workers' output. Supervisors could even direct the "reworking" of certain cases or "batches" of cases (see Gilford et al., 1983:Appendix F, for an example). As "quality control" textbooks indicate, however, sampling inspections work most effectively and at least cost when used in conjunction with control charts—inspection costs can be minimized with the assurance that a process operates in statistical control. And, as noted above, sampling inspection plans do not and cannot make quality improvements.

Quality Audits

Managers use quality audits to ensure that particular quality improvement systems are in place and working as planned. Such an audit, for instance, may verify that specified control charts have been established; that operators or clerks, and their supervisors, have been trained properly in the use of the charts; and that the charts have, in fact, been used and continue to be used. If, for the period under audit, a chart shows an indication of the introduction of an assignable cause of a performance problem, the auditors will look for the record of the search for trouble, its outcome or conclusion, and the effect—in the chart—of subsequent corrective action. Quality audits give managers feedback on the operations of the quality improvement system itself and may in turn suggest areas for improvement in quality control or quality improvement operations.

PROCESS IMPROVEMENT AND PRODUCT OR SERVICE DESIGN

Process control alone cannot give a manager a full picture of the potential capability of an improved process or the causes of variation in performance. Additional tools and information are needed for process improvement. In addition, a product's or service's "fitness for use" (as some quality improvement experts use the term; see, e.g., Wadsworth et al., 1986:15) cannot be improved without further extending the scope of quality improvement systems: in the industrial model, to the consumer or user of products (services); in the public model, to the achievement of program objectives as expressed by the policy makers.

Furthermore, what may be appropriately considered a common cause of poor performance or systemic problem at one level of analysis may—at another level—become a problem with an assignable cause. For instance, suppose that a state agency does not effectively train new Food Stamp eligibility workers. From the perspective of each local office, eligibility determination processes may well be in statistical control, because all workers will be relatively evenly—albeit badly—trained. In statistical control terms, the achieved performance will reflect poor training, which is part of the common causes. However, suppose that in one local office, a supervisor uses an effective on-the-job training strategy with all new workers. Comparisons of performance across local offices would show differences across offices. Although the common cause of poor state training affects performance statewide, the better performance in one local office may indicate an assignable cause for it. A search for the cause of this variation might then discover the local office training and provide a key for improved statewide performance through improved training.

A comprehensive quality improvement system, therefore, extends beyond process control and beyond other means of monitoring performance. Such a system incorporates analysis to determine causes of the achieved level of quality as well as variations in performance around that level. A comprehensive quality improvement system also will include monitoring tools and evaluation studies that focus on process capability and product or service design questions. Finally, a comprehensive quality improvement system also will incorporate the means by which to plan, test, and take action based on performance data.

Determining Causes of Poor Quality

Control charts cannot show the causes of poor quality and may not indicate where in the process problems will be found. Similarly, sampling inspection and quality audits cannot explain achieved performance levels or point to improvements.

People must take information given by control charts, or from sampling inspections and other reports, and diagnose the cause or source of the problems indicated. An experienced user of such charts who also knows the process well can probably make informed judgments in beginning the search for the cause of a problem. But even experienced users of control charts who are knowledgeable about the process in question will need to look carefully at the process, sort out possible causes, and test their effect on performance before determining the cause of the problem. Only then can appropriate action be planned. Other tools can be used to systematically analyze the possible causes of a deficient output. Indeed, a creative manager with a blackboard and chalk can draw a flowchart of the process in detail and query himself or herself, colleagues, and subordinates, for probable causes. Quality improvement experts, however, offer a variety of formal tools that can be used by the relatively unsophisticated on-line manager and a group or task force in a search for causes of problems in performance.

One such tool is called the Ishikawa diagram, cause-and-effect diagram, or "fishbone" diagram (Wadsworth et al., 1986:310-313). (See Figure 13 for an example.) The ribs of the "fishbone" identify the major steps in the process and the observed problem. The group then identifies the possible causes along each rib and continues this process until sufficiently discrete and testable causes are identified. An exercise with this kind of diagram may lead to a number of such fishbones, each focused on a more specific problem than that with which the exercise began.

The fishbone or cause-and-effect diagram also may help to set priorities for testing alternative explanations of a problem, which in turn precedes the development and testing of alternative corrections. Another tool is the Pareto diagram (Wadsworth et al., 1986:306 ff), which ranks causes or sources of error by their relative effect and displays the results graphically. The ranking alone suggests targets of opportunity for maximum effect in improving performance. Simple scatterplots also may aid in searches for the

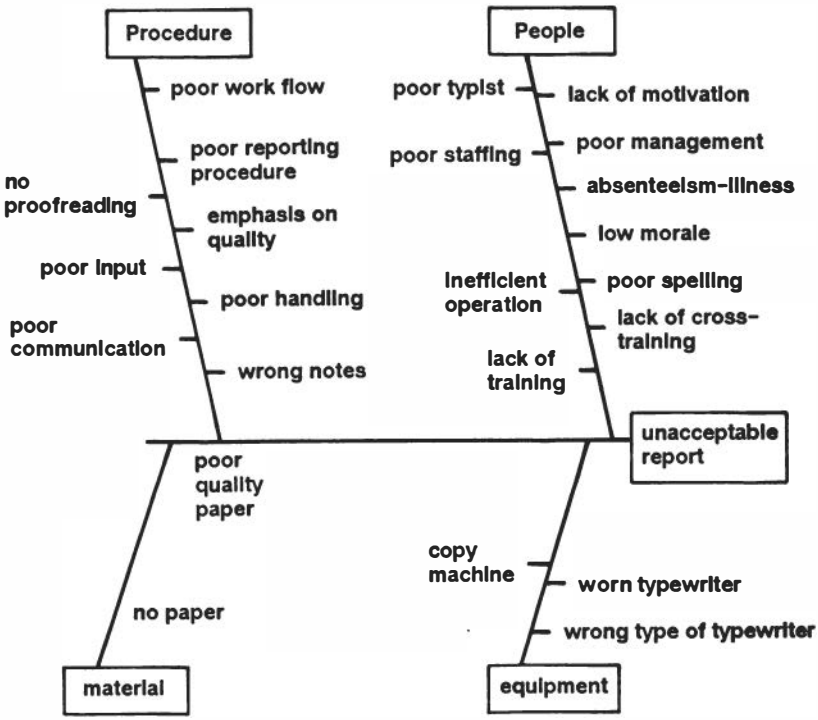


FIGURE 13 Cause-and-Effect Diagram

Source: Ingle and Ingle (1983:303).
Reprinted by permission of authors.

causes of problems, if and as appropriate data are available or collected. There are many other graphical displays and techniques that also may be useful (see, e.g., Wadsworth et al., 1986:Ch. 9).

Monitoring and Evaluating Performance

Quality improvement systems incorporate the tools of statistical quality control, but such systems do not rely solely on process

control, for which those tools were first developed. At more aggregate levels, performance must be monitored and tracked, often including summaries of control chart results.

Some important measures of performance may not be amenable to routine reporting and analysis, if, for instance, such measurements are expensive to make. Routine reports may not indicate where or how processing capabilities may be improved. Special studies are used, therefore, to both elaborate upon routine performance assessments and to help track down both assignable and common causes of problems in the quality of products or services delivered or to identify or evaluate alternatives for quality improvement. If, for instance, program managers or policy makers want to study performance measures other than payment accuracy, or if they wish to expand analyses of program effort to include, say, the effect of program interactions between receipt of Food Stamps and other services, special studies or traditional program evaluations would be appropriate.

One type of special study is a controlled experiment tailored to the needs of quality improvement activities (see e.g., Taguchi, 1981). Whether intended for improved design of a processing system or for improved product or service, controlled experiments can provide managers with sound information concerning what alternatives may work best. When appropriately designed and carried out they may also provide sound information concerning cost and benefit trade-offs among alternative actions.

Special studies and controlled experiments need not be elaborate and expensive. Small studies can be conducted within single offices or work groups. Some studies, and especially experiments, can be done cooperatively, among groups of offices (or states). For the latter kind of effort, it may be necessary to have a coordinating and technical assistance staff located centrally in order to facilitate and promote such efforts and to disseminate their results. Such an approach may be particularly appropriate for the Food Stamp Program, since FNS oversees program administration in more than 50 state-level jurisdictions.

The larger view of quality improvement, and the design and implementation of quality improvement systems, rests on the explicit linkage of all parts of the larger process to the highest goals or purposes of the organization. It reinforces what is only implicit in the use of the tools of statistical quality control, that discrete

steps in a process are intended to make operational the organization's larger mission. Processes operating in control, and working very efficiently, mean little if the product or service delivered does not meet the goal or purpose of its use.

The analog in public programs is conceptually simple. Programs are designed and administered in order to achieve stated purposes. The services delivered should be delivered as efficiently as is possible, according to prescribed procedures, but they also should be delivered with effect, to achieve the purposes for which the program exists. In this sense, then, quality improvement systems would also consider the larger objectives served or intended. Quality improvement systems can be useful mechanisms for holding appropriate people accountable, throughout public programs, for improvements in performance within their control.

Taking Action

Action is essential to quality improvement. It requires the testing of explanations of and alternative solutions to identified problems, which in turn often requires the participation of people from different parts of an organization or different participants in a policy process. Whether the effort is to correct a problem with an assignable cause or to effect systemic improvements, multiple units are likely to be involved. Typically, problems will arise with respect to organizational boundaries and different processing responsibilities (see, e.g., Ishikawa, 1985).

Many techniques for tackling the problems of intra-organizational boundaries have been developed. Among the more popularly known are quality circles. Quality circles bring together workers and managers from different parts of an organization to review performance and to explore opportunities for improvement. Typically, circles extend throughout an organization, cross organizational boundaries, and overlap adjacent management levels. Recommendations for action and analyses of options are passed to the level at which decisions can be made. Quality circles meet on regular schedules, usually during working time (see, e.g., Crocker et al., 1984). Other types of quality improvement teams, established on an ad hoc basis or systematically, like quality circles, also have been used. Many state welfare agencies, for instance,

use corrective action panels or task forces to plan and oversee actions intended to correct identified performance problems. Such groups usually consist of representatives from different parts of the agency, each of which is involved in the problem and the plan.

These techniques are related to participative management strategies, management by objectives, and matrix management—approaches from the disciplines of public administration and management science. Two basic ideas lie behind such approaches. The first idea is that no single department or unit can effectively address many performance problems. No matter how an organization is organized, its structure will create some barriers to the identification of problems and the development of solutions to them. Some means of cross-cutting those boundaries is required.

The second idea is that people throughout an organization, at all levels, have some experience and expertise that can and should be brought to bear on quality improvement efforts. The responsibility for making improvements lies with those people who have the ability and resources to do so. Workers on the line are not responsible for large, systemic performance problems. Quality experts estimate that anywhere from 60 percent to 90 percent or more of the causes of performance problems are systemic problems, not the responsibility of line workers (see Deming, 1982; Ishikawa, 1985). However, managers cannot and should not take direct responsibility for problems specific to line workers or work groups. Managers are responsible for putting administrative systems into place, including quality improvement systems, and for providing those who must use these systems with the needed tools, including training. Proposed action plans, therefore, must be presented to groups that have the resources and authority to act.

One example of how organizational structure can create problems for the effective use of quality improvement systems concerns the accounting of costs and benefits of those activities. Typically, budget and accounting systems focus on the units by which an organization is structured. Yet many of the benefits gained from effective use of quality improvement systems may show up in department budget line items or in accounts different from those in which the costs of the quality improvement activities occur. This distribution of costs and benefits complicates cost-benefit assessments. Without explicit attention from managers at a high

enough level to see the costs and benefits broadly, unit or department managers may make decisions that are quite unjustified on cost-benefit grounds (see Grant and Leavenworth, 1980:8-14, for a particular example).

The panel notes that problems for quality improvement that are created by organizational boundaries are not unique to quality improvement. Most administrative or staff functions within a large agency face similar difficulties, e.g., budget and finance, accounting, personnel, and planning. Without systematic establishment of the function and its importance for all in the organization, such efforts also founder.

The actions required for quality improvement will reflect the nature of the problem that is to be addressed and the level of the process at which the problem exists. Some causes of performance problems in the Food Stamp Program could be in the law and require the attention of the high-level policy makers to effect improvement. Other causes could arise from the federal oversight activities or the transmission of policy interpretations through regional offices. Depending on the specific problem, action might be required at the highest FNS levels. A quality improvement system should include the means by which to identify the assignable and common causes of performance problems and locate responsibilities for action at appropriate levels.

ORGANIZING FOR QUALITY IMPROVEMENT

When carried out effectively, all three elements of a quality improvement system—process control, process improvement, and product or service improvement—share certain features:

- specification of purposes or objectives for which a process has been designed and of measures that relate directly to those objectives;
- collection and analysis of performance data;
- feedback to appropriate people of the information obtained from the measurement of performance and the analysis of those data;
- action based upon the results of such analyses; and
- follow-up of actions to assess their effectiveness.

Quality improvement is purposive. Whether for process control, process improvement, or product or service improvement,

quality improvement starts with specific goals or objectives that can be transformed into measures of what is desired or required. To further an organization's or program's mission or purpose, the measures must relate directly to the goals of the organization or program. At the operational level, those goals are expressed through prescribed procedures and actions.

Quality improvement is driven by the analysis of data relevant to performance. It is not enough to specify objectives and measures. Quality improvement also requires that people collect, analyze, and report measures of achievement or performance against those goals and objectives. Data should provide information about the quality of the product or service as well as about the process by which the product is made or the service delivered. In successful quality improvement systems, such data gathering and analysis relies on a broad variety of both routine and special-purpose studies. Among the latter, controlled experiments permit the testing of alternative solutions to identified problems.

Quality improvement requires timely feedback of performance information. Quality improvement systems also include the means by which to feed back into the processing system relevant performance information. Information gained from performance measurement and analysis may suggest or even clearly indicate steps that can be taken to improve the quality of a product or service. However, unless such information is given to those who have an interest in using it, and done so in a timely manner, the information cannot effectively be used.

Quality improvement requires action. Even with appropriate feedback into the system, actions still must be taken. Action with purpose is the point of quality improvement systems, action based on hard and credible analysis. And since action aimed at quality improvement very often cuts across boundaries within the organization's structure or, in public programs like Food Stamps, even across program boundaries, it requires the cooperation of people from different organizational units. It follows that different levels of management must be involved, actively, if such cross-unit actions are to be planned, carried out, monitored, and assessed. Such actions require decisions from those with the authority and resources to undertake them. Managers at all levels must value the quality improvement work to be done so that they devote their time, energy, and attention (and that of their subordinates) to it.

There must be incentives from the highest levels of management in the organization, or else quality improvement systems that cut across organizational boundaries cannot have their fullest effect.

CONCLUSION

The lesson from the experts is that if quality improvement is to be pursued, its importance must be established at and by the highest levels, and its work must pervade the operational units. There quality improvement requires central staffing with appropriate resources for technical assistance as well as the provision of needed information, training, and other assistance. It also requires moving as high up in the organization (or program) as is necessary to effect needed changes. In the intergovernmental structure in which the Food Stamp Program operates, that can mean—depending on the nature of the problem and the nature of the improvement—anywhere from line staff at the state or local operating level to the federal regulatory and oversight level or even to the basic laws that determine program design and program requirements.

The Food Stamp QC system contains some elements of a comprehensive quality improvement system and lacks others. It serves primarily a federal monitoring purpose and measures only payment accuracy while other program objectives to which program managers and policy makers should be held accountable—like participation rates, nutritional gains among recipients, timeliness of service provision—are absent.

Finally, the QC system does not now distinguish assignable from common causes of poor performance, and it cannot, therefore, attribute to appropriately located managers and policy makers the responsibilities for quality improvement actions. Sustained systemic quality improvement requires such attribution in order to permit and promote effective action.

REFERENCES

Bowker, A.H., and G.J. Lieberman
1959 *Engineering Statistics.* Englewood Cliffs, N.J.: Prentice Hall, Inc.

Crocker, O.L., S. Charney, and J. Sik Leung Chiu
1984 *Quality Circles: A Guide to Participation and Productivity.* New York: American Library.

Deming, W.E.
1982 *Quality, Productivity, and Competitive Position.* Cambridge, Mass.: Center for Advanced Engineering Study, Massachusetts Institute of Technology.

Gilford, D.M., D.P. Affholter, and L. Ingram, eds.
1983 *Family Assistance and Poverty: An Assessment of Statistical Needs.* Committee on National Statistics. Washington, D.C.: National Academy Press.

Grant, E.L., and R.S. Leavenworth
1980 *Statistical Quality Control*, 5th ed. New York: McGraw-Hill, Inc.

Ingle, S. and N. Ingle
1983 *Quality Circles in Service Industries: Comprehensive Guidelines for Increased Productivity and Efficiency.* Englewood Cliffs, N.W.: Prentice-Hall, Inc.

Ishikawa, K.
1985 *What is Total Quality Control: The Japanese Way.* Englewood Cliffs, N.J.: Prentice Hall, Inc.

Montgomery, D.C.
1985 *Introduction to Statistical Quality Control.* New York: John Wiley & Sons, Inc.

Taguchi, C.
1981 *On-Line Quality Control During Production.* Japanese Standards Association. (Available from American National Standards Institute, New York.)

Wadsworth, H.M., Jr., K.S. Stephens, and A.B. Godfrey
1986 *Modern Methods for Quality Control and Improvement.* New York: John Wiley & Sons, Inc.

6
Recommendations for a New System

The quality control systems for family assistance programs have evolved over more than 20 years. The Food Stamp QC system has taken its basic form and approach from the AFDC QC system that was implemented in 1970—developed to address particular concerns of that time—and it evidences both strengths and weaknesses. The weaknesses derive in part from the narrowness of focus; the strengths derive in part from what has already been learned from a long and well-studied history. As changes are made to this system in the future, the panel endorses the notion that the strengths that have been built over the 20-year history need not and ought not be lost as the weaknesses, shortcomings, or disadvantages are addressed. The recommendations that the panel makes for a new quality improvement system attempt to build on the current system, but to serve the larger set of interests represented in the current debate. This chapter first addresses the strengths of the current QC system and then outlines a strategy for promoting quality improvement, encompassing and in response to the two charges in the congressional mandates.

STRENGTHS OF THE FOOD STAMP QC SYSTEM

The current QC system for the Food Stamp Program produces and reports annual estimates of payment error rates, by state and for the nation as a whole. The sources and types of errors most frequently found are also reported. Before the QC system was in place, this kind of information was unavailable except through limited special studies. The Food Stamp QC system does a very large job of measurement, analysis, and reporting, sampling and reviewing more than 60,000 cases annually. With the exception of the final steps in the administrative process, it does that job on schedule.

Probability samples are drawn monthly, under procedures that are monitored for adherence to the sampling design. Cases are reviewed intensively, many measures are recorded and reported, and specified estimates of performance are generated annually. Since samples are drawn for review every month and reported at least that frequently, federal and state managers have access to interim data throughout the cycle. The performance data are analyzed in prescribed ways, and the results are put into a process that intends corrective action. Measured performance is related to consequences that are designed to promote improved performance.

Substantial Resources Dedicated to QC

At both federal and state levels, agencies invest substantial resources in the QC measurement, analysis, and reporting system. FNS devotes considerable staff resources to QC: in fiscal 1986, there were 150 full-time-equivalent QC staff positions at a cost of $4.9 million (for salaries, benefits, and travel). That effort is comparable to official estimates of the federal resources devoted to each of the QC systems for AFDC and Medicaid. Furthermore, FNS has used and continues to use outside experts to provide advice and guidance on a variety of aspects of its QC system.

Every state welfare agency has a staff dedicated to QC reviews and often another, smaller staff assigned to corrective action planning. Many other program staff, at all levels, are also involved in corrective action. In federal, state, and even some larger local welfare agencies, there are staff in QC and related activities who are trained in statistical methods and in such disciplines as

public administration, management analysis, research design, and program evaluation.

The QC system has reinforced an interest in doing good work throughout the Food Stamp Program, from the Congress and the federal agency executives to local Food Stamp workers. Perhaps most important, and as noted earlier in this report, measured error rates have shown improvement since 1980. Whatever the actual effects on performance of the QC system itself, and through whatever means, the QC system has documented and reported those performance measures and their improvement over time.

Debates abound concerning the meaning of "quality" and the utility, value, and effects of the current QC system for different purposes and as it now operates. Nonetheless, the panel believes that the investment in resources for QC, the emphasis throughout the delivery system on producing quality work, and the success of the system in measuring and reporting performance all represent strengths upon which to build.

Comparative Strength

In its initial work plan, the panel undertook to consider other federal programs in order to glean from them lessons for QC in the family assistance programs. Those selected were thought to have activities that were in some ways comparable to those of the family assistance programs. The panel found few lessons on which to draw.

The investment in QC resources in the family assistance programs seems to be greater than in other public programs. In some of those programs, there appear to be little or no designs for QC-type work. The panel found some lessons for quality improvement tools and techniques from the Internal Revenue Service, which will be developed further in the panel's analysis of the QC systems for AFDC and Medicaid.

The panel does not rely upon such comparisons in recommending improvements to the Food Stamp QC system. Three points, however, about QC in the other programs are worth noting. First, some programs relied on the family assistance programs for models of QC systems. The Supplemental Security Income (SSI) Program and the Unemployment Insurance (UI) Program, for instance, took the family assistance QC systems as their model

for QC system development. Although it is not clear that family assistance QC systems provided the model for QC in the Old Age and Survivors' Insurance (OASI) Program—also known as Retirement and Survivors' Insurance, commonly known as "Social Security"—the OASI QC operations appear to be a simplified version of AFDC QC: cases are sampled and subject to intensive review for the estimation of payment error rates.

Second, other public programs have much more limited QC efforts in comparison with the family assistance programs. The OASI Program, for instance is reported to sample only 2,000 cases annually for review and for only one month in the year; the Food Stamp QC system samples more than 60,000 active cases annually. Other programs may be years behind the family assistance programs in using QC tools. In the UI Program, state QC systems are not mandatory, although the U.S. Department of Labor intends to make the system mandatory in the future. For major educational loan programs (guaranteed student loans and direct student loans), there is no QC system now in place to monitor payments on loans, although formative work currently under way may lead to such a system in the future.

Third, payment error rates for the Food Stamp Program compare favorably with those of other programs. There are no official estimates of national error rates (or other performance measures) for the UI and educational loan and grant programs; unofficial estimates suggest payment error rates of up to 25 percent. The SSI program—with fewer earners than Food Stamps and direct federal government administration—reports payment error rates lower than those for Food Stamps (but above those for Medicaid). The OASI Program reports payment error rates under 1 percent per year, but the recipients of OASI benefits are retirees with little or no earnings (or their dependent survivors) whose eligibility has been established over a long period of time during which they (and their employers, usually) have made contributions to the fund. And both OASI and SSI are administered directly by the federal government, in contrast to the more complex intergovernmental administration of the family assistance programs.

The Internal Revenue Service produces national estimates of IRS performance, although their comparability to QC-based error rates for family assistance programs is open to question. One estimate is that of a total "tax gap"—the difference between the total

amount that should be collected from all sources and the tax revenues that are collected. Another estimate is that of a "voluntary compliance level"—the tax liabilities voluntarily reported and paid by individuals—(corporations and nonfilers are excluded) as a percentage of the liabilities that would be assessed if all such returns were thoroughly audited. The IRS (1986) forecast a "tax gap" for 1986 of 18.5 percent and estimated a voluntary compliance level for 1982 of 91.8 percent (or an 8.2 percent noncompliance complement).

THE CHALLENGE OF QUALITY IMPROVEMENT IN THE FOOD STAMP PROGRAM

The two charges in the congressional mandate to which this report responds reflect the goals of making QC systems serve state management purposes and of holding state agencies accountable for payment accuracy. Imbedded in the current controversy, in the congressional request for both this study and a study from USDA and in the nature of the specific charges for those studies, is the recognition that many interests are at stake in trying to control and improve the performance of the family assistance programs.

Those interests include accountability for the expenditure of public money, the achievement of program objectives, and the attainment of program efficiency. These interests are important to all levels of government engaged in the delivery of family assistance. Serving these interests is made more difficult, and at times more acrimonious, by the complexity of intergovernmental jurisdictions, intermingled program responsibilities, and mixed policy prerogatives that characterize the administration of the family assistance programs.

As detailed above, program objectives in the Food Stamp Program seem straightforward enough, compared for instance to AFDC: improving nutritional status seems a smaller task than attaining self-sufficiency. But the Food Stamp Program objectives carry within them their own peculiar challenges to quality control and quality improvement. For most of its life, the program has experienced repeated and aggressive attempts to expand coverage; the aim of the QC system has been to tighten eligibility determinations and minimize improper food stamp issuances. These need not be completely contrary aims. But as a practical

matter, it may be difficult for program managers to undertake aggressive outreach and intake procedures, absent significantly increased resources, without at the same time loosening up eligibility procedures.

Improving performance in the family assistance programs is further complicated in other ways. The policies and administrative procedures intended to implement the programs are promulgated at different times by all three branches of the federal government and then by governments at the state and local level. Program rules themselves are intricate and knotty because Food Stamps, like other family assistance programs, has been the target of contention among different interest groups and therefore the focus of repeated legislative and regulatory changes and litigation. And the Food Stamp and AFDC Programs become intertwined because at least 40 percent of Food Stamp participants are served by the AFDC Program according to its rules as well. In addition, program procedures cannot be expected to integrate smoothly at the service delivery level, since the state and local agencies that administer these and related programs have been encouraged to innovate, to deliver different kinds and levels of benefits and services to potentially different recipient groups. As Food Stamps interact with other programs at the delivery level, changes in those other programs are bound to affect it. For other services, like those designed to satisfy the work requirement, the Food Stamp Program must interact with other service delivery systems at the state and local level.

The complexity of objectives and administrative interactions should encourage, not deter, the efforts to improve quality. Quality control ought to be viewed as a vehicle to monitor performance over time against the different and sometimes competing interests and objectives that the program is designed to serve. QC should be a vehicle to improve the program to better serve those different interests. Instead, the current system has amplified conflicts among interest groups and directly led to the current controversy and mandate for this study.

CHARGE 1: IMPROVING STATE MANAGEMENT

The first charge to the panel is to recommend ways to develop information from QC systems to serve state management purposes.

Although that charge could imply that improved state management is only a state interest, the complexity of program objectives and intermingling of program administration suggests that improving state management is of interest to both state and federal policy makers. State and federal officials also share an interest in broader measures of quality than simply payment accuracy, although the federal government has a clear interest in ensuring payment accuracy.

The findings and recommendations below reflect the desire to exploit the different capacities of the federal and state levels, as well as the mutual interests of each in program efficiency and program effectiveness. The panel also understands that program improvements at the service delivery level require attention to details of program operations and attention to details of program interactions. This report proposes a framework for quality improvement strategies; the panel may have more to say on program interaction at the service delivery level in its second report.

Restructuring the Current System

The quality control system now used for the Food Stamp Program serves, primarily, part of a traditional audit function, that of ensuring accountability for payment accuracy. This is an important function, indeed understood by federal and state administrators as essential to sound management practice and to any system of quality control in programs that involve intergovernmental delegations of spending authority. As discussed below under Charge 2, it does not serve that purpose well.

Beyond this important function, the current QC system is limited in what it measures and therefore does not serve well the purpose of policy makers or program managers toward continued quality improvement and the achievement of program objectives or program efficiency. Because it emphasizes one narrow concept of quality, it may even distort the incentives for program managers to make improvements. From the legislative history and accounts of those who helped develop the current system, it appears that the design developed for AFDC, and adopted for Food Stamps, was meant in part to respond to concerns about growing program size and cost. It therefore focused on causes of ineligibility among recipients, including fraud and abuse. Such a focus might offer a

proper safeguard against misdirected benefits, but it cannot affect other aspects of program cost. Some program growth has been largely deliberate—aimed at increasing well-being, alleviating the effects of poverty, and increasing self-sufficiency. Increasing self-sufficiency, thereby reducing dependency in the long term, has proven to be an extremely complicated task.

Although the current system offers one positive incentive to reduce payment error—so-called "enhanced funding"—few states have been able to earn the reward. The dominant aspect of the current system is a highly contentious financial sanction mechanism, with few other tools to develop meaningful state-level data to assist in quality improvement or corrective action. The current system does not incorporate some essential elements of a quality improvement system. Even though the evidence shows that many states have succeeded in reducing error rates, the panel does not believe that the current QC system can effectively promote continued performance improvements.

The QC system now in use provides state program managers with some information necessary to plan and implement changes that will improve operations. But the current system does not provide the detailed data that are needed for identifying and correcting problems at the service delivery level. The current QC system serves a federal monitoring purpose by providing a vehicle to hold states financially accountable for one type of payment inaccuracy. However, the arbitration and appeals processes created to serve that federal monitoring system are protracted over long periods of time, making it exceedingly difficult for state managers to accommodate the final results of QC findings for their own management purposes. Also, in carrying out the monitoring function in the current QC system, state QC reviewers at times find themselves in difficult positions: they serve their state agency employers and they stand in for the FNS/USDA in what may—and has—become an adversarial position to those state employers.

Finally, the current QC process raises serious problems, such as the meaning and propriety of the "regression estimator" and disputed interpretations of federal policy that become masked as findings of error in QC reviews. Such questions have led to substantial tension between state and federal officials and continue to promise a highly litigious relationship, diluting the potential utility of such activity for federal monitoring purposes.

The panel's recommended solution to the difficult federal-state relationship resulting from the current QC system is to separate sampling and review for the federal purpose of monitoring payment accuracy from state quality improvement activities. The details of this split as it relates to levying financial sanctions and incentives for state performance are detailed below in the discussion under Charge 2. However, it should be understood that the monitoring structure embodied in recommendation 1, below, resembles what has been called a "one-tier" system: that is, one federal review sample, one review of case records by federal QC reviewers, and one estimate of state error rates. The panel does not find the term "one-tier" completely descriptive, however, because its recommended federal monitoring system allows for state review of findings in contention and because it recommends other activities for quality improvement in which both federal and state governments have roles to play.

> **Recommendation 1**: The FNS/USDA sampling and measurement activities for monitoring issuance accuracy in the Food Stamp Program should be made independent of state-level sampling and measurement activities for quality control, quality measurement, and quality improvement.*

Discussion. Sampling, review, and estimation should be done independently for the purposes of monitoring issuance accuracy in order to maintain the integrity of the review process and to avoid diluting that purpose in attempting to make the activities fit another purpose. The USDA may conduct such reviews with its own staff, or it may contract for such reviews by independent contractors. The panel notes that estimates and reports from federal monitoring activities should be tracked over time and provided to state managers because such time-series data can and should be used in quality improvement efforts.

The establishment of an appropriately independent federal review function will mitigate a number of problems in the current QC system. It will eliminate the confusion and tension around the regression estimator of an official payment error rate since there would be no double sampling and re-review process. It will eliminate the inherent tension created by having state employees

*A panel member dissents; See Appendix A.

(QC reviewers) represent a federal government interest that is potentially adverse to the state agency's interest. It will promote financial integrity and accountability and remove the problem of excessive emphasis on misexpenditure from the currently misnamed "quality control" system, allowing a quality improvement system with a larger and more appropriate mission.

The panel notes that the arrangement it recommends may add new tensions. But those tensions are inherent in similar activities, such as those found in federal audit procedures, in which parties on different sides of an audit are accountable to different interests. Those tensions properly maintain accountability to those respective interests. The new quality improvement systems envisioned may also cost more in the short-run than current QC systems, since more activities and more and different information (described below) will be necessary.

A Comprehensive Quality Improvement System

A comprehensive quality improvement system, that would serve both federal and state purposes, requires, at a minimum: a system for ensuring statistical process control on whatever measures of quality the programs choose to monitor; a system for making improvements in the design of that process, based on problems discovered through routine monitoring and special studies; and a system for making improvements in the overall program design based on what is learned from routine measurement as well as special studies of program effects.

Process Control

The current QC system offers no adequate way in which states can address basic statistical process control—of payment accuracy or of the performance of other desired program behaviors, such as timely determination of eligibility, effective delivery of nutritional information, effective communication of client responsibilities that might reduce reporting errors, or effective work referrals.

The need for effective process control cannot be overemphasized. At the simplest level of process control for maintaining payment accuracy, the Food Stamp Program presents formidable challenges. Because of the volatility of the caseload—comprised in part of families and single wage earners who enter and leave the

program, who mix earnings and benefits, and whose income status is in any case peculiarly difficult to monitor—just calculating the correct benefits is a difficult task. The problem of determining the assets or correct benefit for other groups—for the elderly or for those who live or eat in group settings, where boundaries of available resources may be murky—is no easier. Ascertaining household composition in order to determine allotments presents its own potential confusions. The task of monitoring eligibility is further complicated because eligibility is often episodic and for short durations. And because substantial numbers of participants in the Food Stamp Program also participate in other assistance programs, the local welfare office must often apply different sets of rules to the same person or family.

Decisions about eligibility involve attempts to codify complex social phenomena: for instance, about the stability and permanence of living arrangements, and so, who should be assumed to be financially responsible for whom in a domestic relationship. The family assistance programs are not the only programs that are faced with such decisions; many service programs in both the private and public sector face the same task. Solutions—codifications—are always arbitrary and should be recognized as such. The QC systems in the family assistance programs mix the search for "truth" about these social phenomena with the goal of what would otherwise be viewed as basic process control over program operations.

The family assistance programs have been under almost constant scrutiny by policy makers at every level of government since their inception, out of frustrations about failures to achieve intended objectives, over the cost to administer them, and around concern about fraud and abuse. The environments in which the programs must operate—the generalized political environment as well as the ambience of local offices—do not make management easy. Policies and rules are often in flux, staff turnover is often very high, and adequate training is frequently a precious commodity.

These conditions argue strongly for a system in which performance data can be routinely collected so that managers can know how they are doing, know when their units are operating at the limits of their capability and when not, and discover what is within their capability to fix. The current QC system is not useful for this basic process control. It does not produce routine, timely data on

how correctly procedures are followed by workers, which would be provided by statistical quality control tools. It does not facilitate the prevention of incorrect eligibility findings before they become sources of QC error, even when major sources of error are known. As noted in Chapter 2, for instance, QC findings consistently show income-related errors to be about two-thirds of all errors made. It does not permit identification of assignable causes of error at any level in the system on a routine basis—whether from an individual worker, work group, local office or, except serendipitously, from rules and procedures. And it does not permit analysis of measures of quality other than payment accuracy for current federal payment error determinations.

Process Design Improvements

Process design improvements start at the most basic level of process control, using the information gained from process control activities as well as other analyses. Process design improvements can focus on a range of program operations, for example, on improving eligibility determination toward the end of improving payment accuracy or on improving job referral techniques toward enforcing the work requirement or attaining self-sufficiency. Other topics also seem to be likely candidates for analysis for process design improvements. It would be useful for each state to examine in a systematic manner, for example, work flows, forms design, staffing and training of staff, time and frequency of service referrals, clarity of regulations, and technical support systems that may lead state administrators to improve processes. Also, to date, quality improvement has focused principally on techniques that have lowered payment errors in one jurisdiction without understanding and analyzing how their introduction in a different jurisdiction may affect the eligibility and redetermination process.

Although the panel anticipates more discussion of local operations in its second report, it emphasizes the use of planned and controlled experiments to generate sound information on ways to improve program administration. Thus, experiments could test different verification standards and different verification techniques, different levels and kinds of worker training, or different ways of creating service resource linkages within communities.

Program Design Improvements

As elaborated in Chapter 5, a quality improvement system should develop information to clarify objectives, hone expectations, and refine—through increasingly better measurements—the estimates by which program performance can be assessed. In order to accomplish these objectives, it is necessary to create mechanisms for both routine and long-term performance measurement aimed at both micro administrative techniques and broad policy changes and at identifying both "assignable" and "common causes" of poor performance. A quality improvement system also must feed the appropriate information back into the program operations and ensure action based on the information.

The primary motivation for the quality control systems in the family assistance programs seems to have been concern with caseload growth and program costs, therefore, they have focused on measuring eligibility and payment accuracy. However, caseload and cost growth may have little to do with program quality. Indeed, such growth may indicate that the programs are accomplishing certain purposes well. The evidence from the experience with the quality control systems now in place is that they have not reduced caseload growth and costs.

The ultimate challenge of quality measurement is to create a hierarchy of program objectives—including timeliness, coverage, operational efficiency, payment inaccuracies, and other broad measures of program effectiveness—and a set of quality indicators that measure achievements of those objectives. Whether these measures are appropriate to hold states financially accountable—as discussed in Charge 2—they are indicators against which program managers and policy makers can chart their progress toward improving program quality.

Program design improvement requires the same kind of data as traditional program evaluation and in addition the means for acting on the results of analyses of those data. It argues for creating a variety of measures over time and a system that allows for routine monitoring and special studies, with the ability to feed the knowledge gained at each step back into program operations toward achieving continued improvement.

Recommendation 2: Measures of quality should be set against broad program objectives, beginning with the

language of intent in the Food Stamp Act. Federal performance measures should begin with payment accuracy but should be added to over time and as data from monitoring and evaluation activities permit greater breadth and specificity.*

Discussion. The Food Stamp Act of 1977, as amended, describes a wide variety of research and evaluation activities, some of which are mandated and some of which are suggestive of research efforts that the federal government may support. They are clearly intended to ensure that the federal government monitor its progress on achieving a variety of program objectives, including for instance, the measurement of nutritional status among the eligible population. The panel has not made an assessment of the research and evaluation activities that have already or are currently being undertaken by FNS or if they are being used to monitor the achievement of program objectives.

As described in Chapter 2, one indication of how effectively the program is meeting the nutritional needs of low income people is how many such people participate. Participation in the Food Stamp Program of those who are eligible has been a perennial policy issue since the program's inception; therefore, a Food Stamp quality improvement system should provide periodic estimates at both the national and state level of the coverage of the eligible population. The microsimulation models based on data from the Survey of Income and Program Participation support national estimates of coverage; no surveys or analytic models now provide such estimates at the state level. On the basis of national microsimulations, FNS has estimated that about 60 percent of the eligible population participates in the Food Stamp Program and that about 80 percent of those with incomes below the poverty line participate.

The panel recognizes the difficulty of controlling the level of participation and therefore using it as a measure of program quality. Many panel members believe that specific policies, at both the federal and state level, affect levels of participation in the Food Stamp Program. Some panel members believe that coverage estimates should never be included in performance measures for Food Stamps for which states may be held financially accountable

*A panel member dissents; See Appendix A.

because of the many factors that influence participation that are beyond a state's control. Other panel members believe that performance standards for coverage can be set, and therefore that states should be held financially accountable to them. For instance, standards could be expressed as a function either of a national average or as a function of state caseload and demographic characteristics.

Building the States' Quality Improvement Systems

A comprehensive system of quality improvement, which includes process control and process and program design improvements is difficult to design and implement, but necessary to maintain and improve quality of service. States should be encouraged to develop and use such comprehensive systems, and the federal government should assist in that effort with financial resources, technical assistance, and the creation of a federal policy environment that recognizes their importance.

The panel recognizes that with the implementation of a separate federal monitoring system and the considerably reduced state review responsibilities, current QC resources may get diverted. State agencies need protection from having their current quality control resources diverted to competing purposes in state budgets. Other recommendations that the panel proposes specify active federal involvement in state planning to ensure that funds are indeed redirected toward quality improvement efforts.

> **Recommendation 3**: Congress should mandate that state welfare agencies maintain current levels of effort in their quality control operations for 5 years and that state quality control resources be redirected toward the development of quality improvement systems.

There is little basis in current evidence to specify an appropriate level of support for quality improvement systems. The level is probably at least as large as current outlays to prevent overpayments. Federal and state officials should use the opportunity afforded by a 5-year maintenance-of-effort period to assess the need.

Recommendation 4: Congress, FNS, and the states should each set in place evaluations to determine appropriate levels of support for quality improvement at the end of the 5-year maintenance-of-effort period.

In order to assure that the development of state quality improvement systems serve the Food Stamp Program's larger objectives, these systems must also begin to build on the current performance measures. The panel notes that objective standards may be defined and expanded as better measurements can be developed.

Recommendation 5: Congress and FNS should mount evaluation efforts during the 5-year maintenance-of-effort period to develop objective standards for expanding the scope and specificity of quality measurements to which states can be held accountable.

Federal Role as Facilitator

The panel recognizes that much of the contention around "quality control" has been the perception on the part of state officials —in these and other programs—of too much criticism and too little meaningful technical assistance or financial support from the federal government. The problem may be exacerbated by inadequate support at the state level. The panel believes that states should be encouraged to develop their own quality improvement strategies, which should include process control and process design improvement components. The panel also believes that the federal government should take very seriously its responsibility to provide, directly or through others, financial support and technical assistance for quality improvement activities. Quality improvement strategies should be shared among states, and the federal government should act as facilitator, coordinator, and source of financial support and technical expertise for such activities.

Recommendation 6: The federal government should help states to develop quality improvement systems. It should provide technical expertise and financial and regulatory support and should disseminate information about best practices. FNS should support activities that include multistate, randomized experimentation with innovative

practices and interagency cross-program efforts. State agencies should develop quality improvement strategies and share information with other states with similar needs or interests.

Discussion. States need to know how to make their monitoring and evaluation efforts more rigorous, more capable of generating persuasive and generalizable information. The necessary expertise is frequently not easily accessible at the state level, both because of budget constraints and because of fewer people with experience in complex and sophisticated research techniques. Consequently, the federal government should facilitate the sharing of information among states, such as through the use of multistate cooperative efforts for experimentation with innovative service delivery strategies, and make research design and statistical expertise available to state agencies. The states should pool resources with states that are nearby or that have similar caseload or design problems, in order to create generalizable research findings and to use more sophisticated research designs. Multistate pooling of resources may also offer economies of scale in the research.

Management Strategies for Quality Improvement

Quality improvement requires the ability to facilitate administrative and policy improvements from the information gained in monitoring and evaluation activities. Quality improvement also requires nonroutine mechanisms to respond to major program failures or performance shortcomings.

The ability to make routine fixes requires an efficient system of feeding information back into the system—in part by just recognizing the use of the data and analyses generated from basic process and program design activities. Feedback may also be served by a variety of administrative mechanisms that encourage cross-unit, intra-agency problem solving. Quality improvement teams, matrix management schemes, special task forces, or any other approaches that prove useful in problem solving ought to be encouraged.

It is more difficult to develop such management structures across agencies and across programs. But such structures should be encouraged because of the special interdependence of the family assistance programs with other programs with similar objectives. In state and local government, such arrangements can be achieved

only with commitment and intervention at the gubernatorial or mayoral level. Quality improvement funds should be available for such purposes.

Quality improvement also requires an environment in which commitment to quality is pervasive, from federal policy makers, to state and local administrators, to service delivery staff. In the family assistance programs, quality performance objectives—with respect both to eligibility and promised services—also must be understood by potential and current program participants. The perceived commitment to payment accuracy may assist in minimizing client error. The expectation of delivery of promised services may promote the perception of program legitimacy and encourage client compliance, both with program rules and in participation in desired services. Some states have devoted considerable attention to developing intake procedures designed to communicate clearly the opportunities for, as well as the hazards of, specific forms of participant compliance. Such efforts should be carefully evaluated.

The process described is developmental and iterative. Program objectives shift in emphasis over time; quality measures can and should be continually refined on the basis of improved information and assessment. At least equally important, some processes can be successfully institutionalized, while others should be expected to change over time, to encourage innovation and avoid staleness.

The panel recommends no standard national effort for monitoring and evaluation. Rather, it suggests a system that permits a variety of measures; multiple measurement techniques; routine and nonroutine studies; formal evaluations, including carefully designed pilot and demonstration projects; controlled experiments, quasi-experiments, and naturally occurring experiments; and other observational studies.

The details of the quality improvement system that the panel recommends, and some measure of the capability for improvement and the time required for major changes, are subjects for the panel's further exploration.

CHARGE 2: FINANCIAL CONSEQUENCES FROM QC MEASURES

A quality improvement system as described above is a long-term goal for development within the Food Stamp Program. The federal monitoring function will and should continue. For it to be effective, it must have consequences—whether through financial sanctions or through other mechanisms. And if it has consequences, it should be perceived to be fair. The panel's recommendations for holding states accountable to quality measures reflect that perspective.

The QC system for the Food Stamp Program is not described by FNS as an audit system and neither the panel's charge nor its work includes a review of FNS audit functions. Staff in FNS report that FNS maintain an audit capability outside the QC system but that the resources devoted to it are quite small in comparison to the Food Stamp Program responsibilities. Federal agency auditors—from FNS or from other parts of the USDA—do not routinely audit state agencies' certification of the eligibility of, and issuance of coupons to, food stamp recipients. The QC system obviously functions like an audit system for that purpose: it provides state-level estimates of payment accuracy, and based upon those estimates, establishes state liabilities related to certain "misspent" federal funds.

The panel recognizes the evident congressional intent to relate measured performance of state welfare agencies to differential federal sharing in the costs of the Food Stamp Program, to create incentives for state managers to improve the performance of their operations. The panel finds that the current QC system primarily serves the federal monitoring purpose, and that the QC system can be substantially improved for this purpose. Freed of the difficulties inherent in trying to make this system serve both state and federal quality improvement, and hence serving neither well, state and federal officials' attention can be directed toward making the current QC system better serve that federal monitoring role.

The panel also recognizes that program administrators accept audits as legitimate exercises of federal oversight, and the panel reflects that sentiment in its recommendations for preserving financial accountability for payment accuracy. However, the panel views QC review activities from a quality improvement perspective, not one of audits. In order to adapt the current QC system

to better fit the federal monitoring role, the panel makes a number of recommendations to improve the efficacy of incentives for improvement that are tied to measured performance through financial consequences. It also makes recommendations to improve the measurement of state agency performance and to quickly clear the backlog of Food Stamp sanctions from fiscal 1981 and subsequent years.

Setting Standards

The panel precedes its recommendations on how to set standards for performance and hold states financially accountable to them—recommendations 7 through 11—with a caveat. As is detailed in the discussions about these recommendations, the panel cannot set precise thresholds, or target error rates, or the precise magnitude of sanctions or rewards to be tied to achievement of those targets. Certain empirical evidence necessary to frame the choices precisely is not available. The panel also believes that these are policy choices and that they should change over time to reflect changes in capability. However, the panel does wish to make clear its conclusions about the direction and magnitude that these choices ought to take in order to promote desired changes from the current system and the panel therefore offers these five recommendations as a whole.

Picking a Threshold

A numerical performance standard or target that reflects the capabilities of the Food Stamp Program cannot now be set empirically. But as detailed in Chapter 2, the panel notes that the measured error rates in Food Stamps have declined since fiscal 1980. (Such a pattern is often seen with the introduction of quality control systems: overall performance improves as the most serious assignable causes of poor performance are corrected, leaving relatively greater "common" or systemic causes of performance problems, which are more difficult to solve.) Nonetheless, for fiscal 1985, FNS has sent notices of sanction liabilities to 48 jurisdictions, based on measured performance.

The panel concludes there is something wrong with a QC system that "finds" that nearly all of the operational units are

performing poorly in the face of a 5-year record of improved national performance. Either the standards of performance are too tight, or the QC system has failed to put in place the mechanisms that could achieve those standards in that period of time.

A comprehensive quality improvement system should offer adequate resources and incentives to prod states to improve to their fullest capability. In the traditional view of "quality control," a system for a program operating within attainable levels of performance should sanction and reward only those performers at the extremes of performance. If this view were applied to the Food Stamp Program, the sanction and reward mechanism would identify state agencies that are unusual because their performance is extremely good or extremely bad on the distribution of states' annual performance summaries; only those at the extremes of performance would be subject to sanction or reward. Recommendation 8 and its discussion (below) describe how the financial consequences would be calculated to sanction or reward those states whose performance was measured to be outside the standards.

Accounting for State-to-State Differences

Federal law and regulations establish eligibility criteria and benefit standards. If those national program requirements generate state-to-state differences in performance, states should not be held to uniform performance standards. But the question of how to document, and then how to account for, any such state-to-state differences that may be generated is a difficult one. The panel believes that the solution to this problem lies in taking account of known sources of error. State-to-state differences in caseload mix are one source, when those differences reflect cases that are differentially prone to error.

State-to-state differences in caseload mix may influence measured performance in ways beyond the state's control. Households with working recipients, for instance, are known to be much more error prone than other recipients. Routinely and over many years, the single largest type of error found in Food Stamp QC reviews is that associated with income and with appropriate deductions from income. A state welfare agency has no control over such factors as the demographic characteristics or economic circumstances of the state itself. Either or both could influence the proportion

of a state's Food Stamp caseload that consists of recipients with income, the working poor.

Another possible example concerns elderly recipients. The second largest type of error reported by FNS over time is that of excess resources. Although FNS does not know if these errors are associated with elderly recipients of Food Stamps, it is a plausible hypothesis: research on Medicaid, for instance, indicates that a large proportion of the resource errors encountered are those among the elderly participants in the program. A state agency has no more control over the proportion of the elderly in its Food Stamp Program than it does of the proportion of the working poor.

State-to-state differences in caseload mix with respect to portions of the caseload that are differentially error prone can be easily accommodated. State performance targets or thresholds can be set as weighted averages of national standards or thresholds set separately for these groups. Each state's weights are the proportions of its caseload in each group. (The identification of such groups may change over time and should be reviewed periodically for such changes.)

> **Recommendation 7**: FNS should set performance standards to reflect state-to-state differences in caseload composition with respect to groups that are differentially prone to error. FNS should initially examine the possibility of using at least three groups as likely candidates for such corrections at this time: all households with reported earnings, all nonearning households with one or more elderly recipients, and all other households.

Discussion. A simple hypothetical example illustrates how such correction factors would work. Suppose that the national error rate standards are 15 percent for households with earners, 6 percent for households with elderly people and no earners, and 2 percent for all others. Suppose that two states have the following caseload mixes: state A, 20 percent earners, 20 percent nonearners with elderly, and 60 percent all other; state B, 50 percent earners, 20 percent nonearners with elderly, and 30 percent other. State-specific performance standards or thresholds would be:

$$\text{State A: } .2(.15) + .2(.06) + .6(.02) = 5.4 \text{ percent}$$
$$\text{State B: } .5(.15) + .2(.06) + .3(.02) = 9.3 \text{ percent}$$

There is an alternative procedure that would produce the same arithmetic results but that the panel does not recommend. FNS could simply set national standards for each recipient group, and reward or sanction states for performance with respect to each group in the state caseloads. The panel rejects the latter approach for two reasons.

First, the approach of holding states separately accountable to performance standards for subsets of the participant population would, the panel believes, require sample designs that are stratified by those groups. Questions of equity and precision (see recommendation 14) would then need to be addressed for each stratum, each group. As long as state caseload counts or independent estimates provide the relevant proportions needed for state-specific weights, single state-specific standards can be set, and sample designs—for purposes discussed later (recommendation 14)—need only focus on the precision of the estimates of the states' total error rate. Second, the proposed calculation, as described, is both easy to understand and makes clear to a program manager his or her state's final cost sharing position.

FNS should establish a routine analysis of QC data to determine whether identifiable groups—such as those described above—are more prone to error and are differentially represented in states' caseloads. Unless such populations are both differentially prone to error and differentially represented in states' caseloads, such adjustments would have no effect on the performance standards to which state agencies would be held accountable.

Standards Over Time

The panel notes that performance standards need not be set in concrete, fixed for all time. As performance improves, standards probably should be reset to reflect improved capabilities. The aim of a quality improvement system should be to promote continued improvement; it should not rest with the achievement of one specified set of performance standards. Also, different performance measures should be added to the set upon which differential rewards and penalties are established. Indeed, as the Food Stamp Program evolves, it seems quite likely that different policy and program goals will be set or those existing ones might be reordered to reflect changing conditions and priorities. A Food Stamp quality

improvement system should incorporate a regular cycle of reevaluation of standards against achieved performance, with periodic resetting of standards.

Setting the Magnitude of Sanctions and Rewards

For those states subject to sanction or reward, the panel believes that the appropriate dollar amount of sanctions or rewards should consist of a proportion of the value of the benefit dollars in error. The panel cannot scientifically or empirically set the numerical value of that proportion. It also notes, as discussed below, that the proportion for reward could be made less than that for sanction. But if a state agency's performance is far out of line, with expectations based on capability, the penalty ought to be substantial. And that penalty should be tied directly to the measures upon which performance is judged to be deficient. The panel notes also that in any case the proportion should be established such that a state's liability should not be substantially larger than that which would have been incurred under the current system.

To understand the calculation in algebraic terms, let

X	=	measured difference between the estimated error rate and the sanction threshold,
X(state's total benefit dollars)	=	maximum dollars at risk for sanction,
S	=	sanctions proportion, and,

therefore,

S(maximum dollars at risk)	=	sanction liability.

Similarly:

Y	=	measured difference between the estimated error rate and the reward threshold,
Y(state's total benefit dollars)	=	maximum dollars subject to reward,
R	=	reward proportion, and,

therefore,

R(maximum dollars subject to reward)	=	reward possibility.

Recommendation 8: Sanctions and rewards should be calculated as a proportion of the benefit dollars in error above or below the performance standards. Because the base for calculating sanctions will be different from that under current procedures, the sanction proportion will need to be revised.*

Discussion. The panel does not believe that rewards for excellent performance need to be as substantial as sanctions for poor performance. In part, this reflects a recognition that some incentives already are in place, such as higher rates of federal funding for the development of certain automated data processing and information retrieval systems and for the investigation and prosecution of fraud. The panel suggests that higher rates of federal funding also be considered to support special studies for quality improvement. In part, this judgment reflects the recognition that state and local officials share with the federal government an interest in effective and efficient program administration.

In recommending that performance-based sanctions and rewards be established as a function of benefit dollars, it is not the panel's intention to substantially increase or decrease the magnitude of the sanction liability that a state may incur in comparison with what it would be under the current system for an equivalent discrepancy between estimated performance and the sanction threshold. Nor is it the panel's intention to create a large windfall for states that may receive a reward bonus. In choosing appropriate proportions for the sanction and reward functions, Congress and FNS should consider the magnitude of the sanctions or rewards in the current system.

For example, a state whose error rate currently is in excess of the standard by exactly 5 percentage points (a 10 percent error rate) will today incur a sanction of 35 percent of the federal share of the state's administrative costs. Since the federal government pays one-half of such costs, the sanction represents about 17.5 percent of the state's budget for administration. If the state's administrative costs are 10 percent of the benefit value of food stamps issued, the sanction under the current system is about 1.75 percent of the total value of food stamps issued.

---

*A panel member dissents; See Appendix A.

Under the panel's recommendation, the state with an estimated error rate of exactly 5 percentage points above the threshold, as above, could risk as much as 5 percent of the total value of food stamps issued. If the sanction proportion chosen is one-half of the total dollars in error, then 2.5 percent of the value of food stamps issued would be sanctioned; if the sanction proportion is one-third, then 1.67 percent of total benefits would be sanctioned; if the sanction proportion is 0.7, the sanction would be 3.5 percent of the total benefits, or twice the sanction under the current system.

It should also be noted that the proposed sanction and reward system eliminates a problem in the current system that applies both to its sanction and incentive mechanisms. The current incentive provision is so-called enhanced federal sharing in administrative costs for state agencies that exceed the current threshold and that meet certain other criteria. The enhanced rate is fixed and applies equally to any performance better than the specified performance threshold no matter how much better. Therefore, there is no incentive for further improvement. Similarly for sanctions, once a state has passed a sanction threshold, there is no further disincentive to bad performance until the next full percentage point in error rate is reached. The panel's recommendation, tying rewards and sanctions to a fraction of benefit dollars at any point between measured performance and the threshold, eliminates the problem for both rewards and sanctions.

As part of the question of setting thresholds for sanctions or rewards, the panel also considered in detail the question of purported "bias" in the use of the point estimate for taking sanctions, which some critics have charged arises as a result of sampling error in the estimate. The panel uses the term "statistical fairness" rather than "bias" since the latter term has another meaning in statistics. A technical discussion of this question appears in Appendix B. The panel defines a system to be statistically fair if the expected consequences of measured performance based on performance estimates from probability samples do not differ from what those consequences would have been if the estimates had been based on the entire universe or population. The panel's analysis in Appendix B assumes an unbiased estimator of payment error.

As the analysis in Appendix B shows, statistically unfair outcomes under the current Food Stamp QC system will occur because

of the step function that determines sanctions. Under the current system, states are more seriously disadvantaged as they approach the upper end of each step for which there will be a change in the sanction (e.g., 5.8 or 5.9 percent). States will, upon occasion and as a result of sampling error, be sanctioned at the higher step. At the same time, FNS is disadvantaged in small ranges just over each step consequence (e.g., at 5.1 percent) and is disadvantaged further the worse the state's performance. As a result of sampling error in the state's estimates, FNS may fail to collect the full sanction that might have been collected. The actual magnitude of the relative statistical unfairness to either state or federal government also depends upon the precision of the estimate of payment error.

If FNS were to use a lower bound of confidence interval rather than the point estimate for determining the amount of the sanction, then states would be less disadvantaged, in some cases much less so. Yet the federal government would be much more disadvantaged. Under a statistically fair sanction (or reward) system, the appropriate estimate would be the point estimate for making both decisions—that is, whether and how much to sanction (or reward)—not the lower bound of an interval estimate.

Federal Response to Poor State Performance

The panel recognizes that, as a matter of accounting convenience, the FNS might need to collect sanctions as offsets to its share of administrative costs since the state never actually receives federal benefit dollars that it might pay back when penalized, and other state funds may not be accessible for the purpose of paying sanctions. However, to sanction administrative funds from poor performing states deprives them of the very resources that they need to make further improvements. The panel is also concerned that other dysfunctions will be directly encouraged by the burdens of sanction. Therefore, the panel also suggests that mechanisms be put in place that ensure that the federal government monitor states' responses to sanctions and assist where appropriate in designing strategies for improvements.

There are few mechanisms within the QC system outside of fiscal sanctions to provide effective oversight of and assistance to poor performers. Yet federal response to poor state performance ought to be helpful rather than merely punitive and not

constrained only by findings of payment issuance inaccuracies or other "sanctionable" offenses.

The chief mechanism for such a federal response is to hold a state out of compliance with federal law (or, in the case of AFDC, with the strictures of its approved state plan). Federal administrators tend to resist such findings because they generate serious system responses, need sound legal basis, and necessitate burdensome follow-through activities.But the merits of something resembling a non-compliance response which offers both carrot (assistance) and stick (federal intrusion) is especially appealing. A similar, but not identical strategy has been used elsewhere: The Job Training Partnership Act (JTPA) stipulates a mechanism for forcing a governor to reorganize local program operations when the local entity fails to meet that program's performance standards, as specified in law and regulation.

The panel cannot specify such standards for the Food Stamp Program. A special Secretary's discretionary fund for federal or other technical assistance and for on-site federal monitoring, when measures from a number of evaluative and monitoring tools indicate the need for special action, could be quite useful. In the extreme case of legal noncompliance, current authority provides for federal intervention.

> **Recommendation 9**: FNS should monitor the consequences of sanctions to ensure that corrective action is undertaken. Congress and FNS should take steps to minimize potentially harmful effects in program administration and program objectives, for example, in the structuring of payment schedules for sanctions. In addition, as a nonroutine response to major deficiency findings, USDA should establish a special Secretary's discretionary fund for aggressive technical assistance and on-site monitoring when a state Food Stamp agency has been determined to have serious performance problems.

Discussion. The panel is especially sensitive to the potential for desired quality improvement efforts to promote dysfunctional side effects. Allegations of case churning offer such an example (discussed above). It offers its recommendations for sanctions and incentives with the expectation that federal monitoring efforts will

be appropriately increased and focused to guard against such responses. It also offers them with the expectation that under the new monitoring system, fewer states would be subject to sanction and so careful monitoring of the consequences of sanction is feasible.

Establishing Performance Measures

As indicators of quality, policy makers can choose to measure anything that reasonably serves the purpose of monitoring administrative efficiency or program effect: payment accuracy, timeliness, administrative costs, coverage of target populations, reduction of dependency, might all be candidates. (As noted above, there are divergent views among panel members about whether coverage could or should be used as a measure of performance against which states might be sanctioned.) Estimates of payment accuracy—of Food Stamp issuances—provide one clear measure of effectiveness and are an entirely proper component of a quality improvement program. The history of welfare policy and the widespread frustration with caseload growth and rising program costs make the current QC focus on overpayment inaccuracies entirely understandable. And the federal interest in controlling delegated spending authorities makes it obligatory that payment accuracy retain an important place in a larger quality improvement system.

But there is no logical or necessary boundary between payment accuracy and a variety of other indicators of program quality. The goals of the family assistance programs are surely broader than payment accuracy. The complexities of these goals and potential conflicts among them require that the objects of quality measurement and study also be broader. Measures of timeliness and administrative efficiency, for example, are useless as indicators of program quality unless they are tied to program objectives and work in tandem with other program objectives. Case churning again offers a useful illustration: it is possible to imagine that the desire to process cases timely to satisfy QC demands, uncoupled from concerns about certain applicants' difficulties in producing documentation, has led to turning eligible applicants away when relaxing the timeliness criterion only slightly might avoid their sacrificing a whole month's needed benefits.

The panel recommended above that federal agencies collect measures of performance that reflect the broad, diverse, and sometimes conflicting goals and objectives of the Food Stamp Program. Such measures should inform a variety of quality improvement activities at all levels of family assistance policy. Performance estimates with serious consequences for state welfare agencies, however, should be both directly relevant to program objectives and reasonably precise.

The panel can as yet identify only measures of issuance inaccuracy as both easily collectable with reasonable precision and directly related to the objectives of the Food Stamp Program. During the panel's recommended 5-year effort to develop effective quality improvement systems (discussed above), FNS should develop and test other measures related to the objectives of the Food Stamp Program. Such measures should be added, in the future, to the set of performance measures for which differential cost-sharing will be used to structure incentives for improved performance.

The current official payment error rate includes only payments to ineligible recipients and overpayments to eligible recipients. Underissuances and negative case errors deserve a place with overissuances in an official payment inaccuracy measure, since one clear objective of the Food Stamp Program is the accurate provision of benefits to those who are eligible and who apply. The panel agrees with the use of dollar-based rather than case-based measures: the dollar values indicate the magnitude of the inaccuracy for each inaccurate case.

Recommendation 10: The USDA should use estimates of issuance inaccuracy in the Food Stamp Program for the purpose of linking financial consequences to state agencies' performance, based on the absolute dollar value of each of four distinct types of errors: issuances to ineligible recipients, overissuances to eligible recipients, underissuances to eligible recipients, and issuances that should have been made to applicants or recipients who were wrongly denied or terminated. Issuance inaccuracy *rate* estimates should be made for each type of error by dividing the estimates of the total dollar value of errors for each component by the total value of issuances actually made. The performance estimate upon which to make sanction

and reward decisions should be based on a weighted average of the four estimates of issuance inaccuracy rates. Although the panel believes that the determination of weights is a policy decision, it believes that each of the four types of error is important and that the weights should reflect this.*

Discussion. The panel acknowledges that policy makers may choose different weights to be attached to each component of issuance inaccuracy, weights that reflect policy choices. It may be desirable, for instance, to give errors made in issuances to eligible recipients less weight than those made in issuances to ineligible recipients or than incorrect denials and terminations. (It is an empirical question whether the same cases may be overpayments one month and underpayments the next, thereby cancelling the monetary effect of each other's payment errors over time.) Or, some people may argue on the basis of economic and social consequences that benefits wrongly denied are the most serious kinds of issuance errors, and therefore should be given the greatest weight.

The panel notes that mathematically and financially equivalent results could be constructed by creating separate sanction and reward functions for each element or component of issuance inaccuracy. The panel believes that its recommended solution—a weighted average, with a single reward and sanction function of the total issuance inaccuracy estimate—is simpler and more appropriate. First, it requires only one function that relates measured performance to financial consequences. Second, it makes explicit the policy choices reflected in the weights, which would only be implicit in the formulation of separate reward and sanction functions. Third, separate functions would complicate sample designs in light of the panel's recommendation for equity with respect to relative risks of sanction across states (see recommendation 14 below).

The weighting of components of error is found quite commonly in the world of statistical quality control. The quality measurement plan used at Bell Communications Research and at AT&T, for instance, draws on a 50-year history of differentially weighted components of errors in its inspection systems. The military standards for sampling inspections also distinguish between classes

*A panel member dissents; see Appendix A.

of errors for inspection purposes, according to their magnitude or seriousness for the use of the product in question. Statistical quality control texts also make such distinctions in their discussions of quality inspection measures, whether for control charts or sampling inspections.

The panel recognizes that its recommended statistic for estimating inaccuracy rates is somewhat flawed conceptually. Ideally, the denominator should include the dollar value of both underissuances and nonissuances that should have been made. The difficulty with this concept is that the denominator is an unknown quantity—the value of all erroneous underissuances and nonissuances cannot be precisely determined for the entire caseload.

The panel also recognizes that policy makers will set standards, rewards, and sanctions as a function of the policy process, not necessarily or only on the basis of scientific evidence or determination. The panel notes that the empirical consequences of using the simple statistic will be of small magnitude: currently, negative case action and underissuance error rates are usually less than 5 percent, often less than 2 percent. The simpler statistic, therefore, is not likely to have much if any effect on the consequences that derive from estimated performance. The panel believes that the simplicity of the proposed statistic makes a sufficient case for its use.

The panel believes that all four components of issuance inaccuracy should be estimated and reported separately, since each represents a different kind of error. Actions to improve performance with respect to one may not affect any of the others or may work to the detriment of any of the others. If they are combined and reported as a single performance measure, important information could be needlessly lost, particularly with respect to changes or trends over time.

Criteria for Measuring Issuance Accuracy

Current QC practice determines recipients' eligibility status for the month under review, based on information available at the time of federal re-review. However, measures of issuance inaccuracies should try to get as much information about eligibility status and benefit determinations as is possible. Those who make policy and basic program design choices—that is, both federal and state program designers—need to know about such things as the extent

to which unreported income effectively drains dollars from the Food Stamp Program.

Federal monitoring serves another purpose, however, that of holding state agencies accountable. Standards to which state officials are to be held accountable must also be under their control. This raises questions of assignable causes of performance problems. For what errors *should* workers and state agencies be held accountable? As discussed in Chapter 5, quality improvement experts emphasize assignable versus common causes of performance problems. The responsibility for accomplishing certain corrective or remedial or improvement actions can then be located appropriately with the persons or levels within an organization or program that have the means—including authority and budget—for implementing the needed or desired changes. The general approach ought to be followed, and state agencies held properly accountable for performance within their control.

Consider, for instance, the case of an applicant who may report no earnings. The worker can find no evidence of current earnings through a prescribed check of past employers and computer matches against specified data bases. The applicant properly is certified to be eligible and is issued food stamps. Later computer matches by QC reviewers (with subsequent verification) show income to the recipient for the period in question. On the basis of that information, the applicant either should have been ineligible or should have received food stamps at a lesser value than was granted. But the information was not known and could not have been taken into account by the state worker when the application was acted upon. The most common data bases for checking for wages, social security withholding information and the wage reports to the state for unemployment insurance purposes, typically run a lag of from 3 to 6 months in making wage information available.

If the systems for income verification cannot provide up-to-date information that workers can use in their decisions, then the workers should not be held accountable for the error described above. The system-level problem may be one of either state or federal responsibility and needs to be addressed. But if the available system has been used, in accordance with specifications given workers and under procedures or guidelines given to or provided by the states, then the state agency should not be held responsible for

any resulting error. If the state has not made the required systems available, however, the state agency should be held accountable for such errors.

Measurement specifications or criteria are needed to distinguish between those problems or errors for which state agencies can and should be held accountable and those for which they should not. State plans offer an appropriate place to establish such specifications or criteria. Therefore, these plans should be subject to FNS review and approval, as they were until 1982. In essence, state plans serve as contractual agreements between FNS and state agencies concerning just what state agencies will do and against what procedural standards they can and should be held accountable.

> **Recommendation 11**: FNS should approve state plans for the Food Stamp Program, and the plans should constitute the criteria against which states are held financially accountable. Federal reviewers should identify those sources of error that are associated with violations of the state plan. Federal reviewers should continue to report on the sources or types of errors that are now measured through QC systems.

Discussion. FNS review and approval of state plans is well within federal oversight responsibilities. Questions about such issues as who can be held responsible for clients' misreporting should be bounded by the systems for verification and the procedures that support them that both federal and state officials have agreed to. However, since both federal and state governments should continue to be interested in searching for truth about payment inaccuracy, federal QC reviewers should continue to collect information about such inaccuracy that extends beyond errors that arise from failures by state and local workers to follow the state plan. Whether a client misreports information is still of interest to both state and federal program administrators even though it is unrelated to a state plan. As federal reviewers collect such information, FNS will build a useful data base for the estimation and analysis of systemic performance problems. Such a data base can be used in planning special studies aimed at process and program improvements. The information also can be used to contrast states' performance, and states with extremely good

or extremely poor showings on such measures might be targeted for special technical assistance or special studies to explore and explain such discrepancies. Either kind of activity might yield further information for making systemic improvements or for shaping individual state plans for quality improvement.

Technical Errors

AFDC and Food Stamp officials have designated certain kinds of errors as "technical errors": they are procedural errors that may or may not represent actual misspent dollars. Examples include the failure to provide social security numbers for all recipients, for instance, and the failure to register for WIN or to assign child support rights to the state in the AFDC program. The failure to document such procedural steps does not automatically render a Food Stamp benefit to be in error. (In AFDC, technical errors result in an automatic determination that the AFDC payment was in error; under Medicaid, federal law prohibits the finding of a payment error simply because of such a technical error.)

FNS does not include technical errors in calculating the basis for imposing fiscal sanctions or incentives—and the panel supports the FNS position—but certain technical requirements and the errors measured against them represent surrogates for important program objectives that should be systematically monitored. Procedural requirements generally provide some form of internal control, much like the separation of duties or the reconciliation of bank statements with ledgers in financial operations. Such internal controls may reduce the risk of financial loss, but the failure to follow such requirements does not, by itself, represent an actual loss—nor does the adherence to those requirements ensure the avoidance of an erroneous financial transaction.

In the family assistance programs, social security numbers for many recipients of assistance give agencies a useful tool for checking for unreported income or assets, through computer matching with bank records, income tax returns, and wage reports. However, under current rules, all recipients of AFDC and food stamps must provide social security numbers to the state agency, even for infants and small children. If the purpose here is to enumerate members of a household, other evidence could be used for this purpose—birth certificates, school records, and third-party corroboration, for instance. And they may inadvertently create

errors where their original purpose—e.g., for wage match checks—is applied inappropriately to other uses—e.g., the evidence of children in a household. The effectiveness of the use of social security numbers for young children, for internal control purposes, can and should be evaluated empirically.

Some technical errors appear to be intended as surrogates for the achievement of program objectives, not just as internal controls. To the extent that they have been demonstrated to provide meaningful surrogates for the achievement of program objectives, they should be retained and used to monitor state agencies' performance. If they serve only weakly as such surrogates for the achievement of other program objectives, different measures should be sought and built into monitoring and performance reporting operations. If they serve, in fact, only as internal controls with respect to program objectives other than payment or issuance accuracy, their effectiveness for that purpose should be evaluated.

A good example of the misuse of technical errors comes from AFDC. WIN registration of certain AFDC recipients is a requirement of eligibility, but the fact of WIN registration alone does not ensure that appropriate or effective training or work experience is received. Indeed, the program objective—relevant and effective training—may be subverted, through (for instance) the automatic registration of all covered recipients, when WIN registration does not ensure WIN participation. The panel is aware of at least one state that has taken this route, in order to avoid AFDC QC-based "payment errors," regardless of the effectiveness of such registration for recipients.

At least one other objective clearly identified in law and regulation has not been included in quality measurements. It is a requirement that applicants be certified eligible and served or certified ineligible in a timely manner—and so informed. In AFDC it is within 45 days of application; in Food Stamps, it is 30 days. Presumably, for serving the eligible poor, faster is better. No useful measure of timeliness has been developed and used to characterize performance. Furthermore, the panel has received analyses that argue that the timeliness standard itself has created a dysfunctional program response in order to avoid a payment error, case churning. No measures of churning have been developed, either.

In sum, federal reviewers should identify, report, and analyze technical errors, since they offer useful information on at least

some procedural controls that have been built into the programs. Analyses of technical errors should provide evaluations of their usefulness as controls; that is, further investigation into cases with technical errors can and should provide estimates of the risk of financial error or loss associated with technical errors.

State Review of Federal Findings in Case Reviews

As the USDA takes on the panel's recommended monitoring role, the state agency should be given a chance to review findings prior to the release of an issuance inaccuracy report and to dispute matters of fact in the federal reviewers' work. A system should be devised so that preliminary findings of issuance inaccuracy can be reviewed and disputed by state officials.

The system also should have a mechanism for the resolution of disagreements when federal and state reviewers cannot agree. As the new system is implemented care should be taken to preserve due process mechanisms. Such a system can be given regulatory shape, with specific schedules and deadlines. Ultimately, states may litigate the findings and financial liabilities that result from such reviews. By focusing only on findings of error in a federal audit and by pursuing only those cases in which federal and state reviewers continue to disagree, the process devoted to the issuance inaccuracy reviews would be substantially less time-consuming than is now the case under the "two-tiered" QC system.

> **Recommendation 12**: Under the new system that the panel recommends, FNS/USDA should establish a process to accommodate state differences with federal findings of issuance inaccuracies and should establish procedures for the timely resolution of differences through the national office with respect to those findings.

Discussion. With the use of federally approved state plans as the standards against which the state agencies are to be held accountable, case review disputes between the states and FNS should occur somewhat less frequently than they now do. Issues of policy interpretation should be taken up and resolved around the approval of state plans, not in disputes of case reviews for issuance accuracy. Nonetheless, some disputes will arise. An arbitration process, such as that now used and described in Chapter 3, could be used to resolve such differences. The panel notes that as long

as the potential for regional FNS variation exists, then a national office arbitration step is essential for the uniform application of national Food Stamp rules and policy.

Sample Design for Issuance Accuracy Reviews

The use of samples in conducting reviews for monitoring purposes is well founded. In audit applications, for instance, misspent funds identified through reviews of samples of records or cases provide a sound basis for estimates of total misspent funds. FNS should base its reviews of state welfare agencies' expenditures on monthly samples of active cases and of negative case actions (terminations and denials), as is now done under QC practice.

In any sampling application, the estimates derived for the population characteristics in question—here, the dollar value of issuance inaccuracies—contain room for error based on sampling alone. When samples are drawn probabilistically, the sampling error can be estimated: that is, a range of error around the sample-based estimate can be constructed in which the true population measure is expected to lie. When one makes judgments or decisions based upon estimates calculated from samples, risks of making the wrong choice necessarily accompany those decisions. Whichever direction the decision takes, some risk of being wrong is present. Sound statistical practice dictates that reports of such estimates also contain estimates of their sampling error, so that consumers of the information can evaluate those risks for themselves.

> **Recommendation 13**: FNS should routinely report estimates of sampling errors for all performance estimates based on sample data.

Discussion. If financial consequences are tied to sample-based estimates, differential risks across states and between state and federal agencies raise difficult and serious questions. Since the precision of state-level estimates varies across states, serious questions of equity are raised in the treatment of state agencies when such estimates are related to consequences of some import. The panel believes, therefore, that sampling error should be equalized across states, in contrast to the current QC practice.

The panel recognizes that a case for different precision across states can be made, in the interests of selectively targeting the use of federal monitoring resources. And for purposes other than

attaching financial consequences to estimated performance, such targeting should be done. Nonetheless, if all states are to be treated equitably with respect to the financial consequences of sample-based performance measures, then equal precision for those measures should be the aim.

> **Recommendation 14**: FNS should design samples for the purpose of estimating issuance inaccuracy rates to achieve equal precision across states.

Discussion. In order to design samples to control precision, information is needed about the distribution of each component of payment inaccuracy that is to be estimated. This information is not available: in the current systems, for instance, the valuation of negative case decisions made in error has not been done. Until such information can be obtained through federal monitoring efforts, the panel suggests designs for federal agency reviews that use equal sample sizes across states.

FNS should routinely and periodically review state agencies' performance, especially issuance accuracy performance. When financial consequences are attached to measured performance and the measures are sample-based estimates, some legal questions might arise concerning the statistical grounds for such uses. Sample designs that control the precision of estimates of issuance inaccuracies, across states, should serve to minimize and simplify any legal issues that might arise concerning equity in the treatment of states.

Financial Liabilities for Past Performance

The panel's recommended systemic approach to quality improvement is quite different from the current QC system. The panel recognizes that it would be impossible practically to revisit QC case reviews as far back as 1981 and rework the measurements and sanctions according to the panel's recommendations for the future. In order for state and federal officials to clear that backlog of liabilities, claims, and appeals as quickly as is possible in order to get on with the business at hand and for the future, the panel recommends solutions of expedience: the recommendations recognize the performance standards set in law, do not change the rank order of states according to estimates of performance, and respond

to some issues of statistical "fairness," (as discussed above and in Appendix B).

As part of the solution to the estimation for past years, the panel has examined the regression-adjusted estimator used by the Food Stamp QC system. The panel finds this estimator lacking for a variety of reasons; the panel is exploring alternative estimators and will report its critique as well as proposed solution(s) in its second report. All three family assistance programs use the regression-adjusted estimator. Any recommendation from the panel will affect all three programs' resolution of the backlog of liabilities, appeals, and lawsuits. The panel, therefore, will take the time afforded by the schedule for the AFDC and Medicaid QC studies in order to develop, as thoroughly as is possible, the estimators to be considered. The panel may recommend one estimator, or it may lay out several with a discussion of the strengths and weaknesses of each.

Once an estimator has been chosen, problems in the current QC system still will make its use difficult and troublesome in resolving existing disputes. Because of the step function that has been established in law, effective for fiscal 1983 and subsequent years, the sanction system is statistically "unfair" to both states and the federal government. It also sanctions against a standard of performance that apparently has no empirical basis—no measure of states' performance capabilities for the time periods in question. The panel believes that without asking Congress to retrospectively change the law, there is no statistically "fair" and reasonable solution to the question of how to use measured QC payment error rates to properly determine state-specific sanctions for years past.

Recommendation 15: For the purpose of determining whether a state agency should be sanctioned under the current QC system for fiscal 1981 through fiscal 1987, FNS/USDA should use a lower bound for the yet-to-be-recommended estimator. The width of the interval between the point estimate and the lower bound should be the same for all states. The width should be based on the estimated sampling error for the state agency with the largest sampling error in its payment error rate estimate. If the lower bound is at or below the sanction threshold,

then a state would not be sanctioned. If the lower bound is above the threshold, then a state would be sanctioned.

Recommendation 16: If the decision to sanction for years fiscal 1981 through fiscal 1987 has been made, the yet-to-be-recommended point estimate of payment error should be used to establish the magnitude of the sanction to be levied.

Discussion. The panel notes the correspondence of its first recommendation here with that made by a previous National Research Council panel, that the sanction decision reflect a very low probability that a state agency is wrongly found to be subject to sanction (see Gilford et al., 1983). The use of the lower bound and the largest sampling error to establish which states will be subject to sanction reduces the probability that a state will be wrongly subject to sanction due to sampling error. But once the sanction decision has been made, the point estimate reflects the best statistical estimate of actual performance, and, hence, should be the basis for establishing the magnitude of the sanction. The panel's recommendations mitigate some of the statistical "unfairness" without differentially penalizing larger states for their larger sample sizes.

On a separate matter, the panel does not specify in its recommendation a level of confidence to be used in determining the lower bound for the state with the largest confidence interval around the estimate of payment error. It notes that the choice of a confidence level should reflect the valuation of different sets of risk to both parties, federal and state. In the absence of that valuation, the panel notes that a 95 percent confidence level has a widespread use and traditional acceptance.

These recommendations may appear to be at odds with the panel's earlier discussion of statistical "fairness," which argued that the point estimate was the appropriate estimate to use for making sanction and reward decisions. That argument rests on the expectation for long-term, average results for any one state. Settlement of past disputes will occur only a few times for any state per program. The use of a lower bound to make past sanction decisions, but the point estimate to determine the sanction amounts, compensates somewhat for the "unfairness" inherent in the current system. The panel notes that some agencies, public

and private, make more extensive use of a lower confidence bound for recovering misspent funds based on an audit of sampled records with the logic that the party that designs the samples and directs the audit should bear the largest risks of sampling error.

REFERENCES

Internal Revenue Service
1986 IRS Measures of Voluntary Compliance. Fact Sheet FS-864. Internal Revenue Service, Washington, D.C.

Gilford, D.M., D.P. Affholter, and L. Ingram, eds.
1983 *Family Assistance and Poverty: An Assessment of Statistical Needs.* Committee on National Statistics. Washington, D.C.: National Academy Press.

Appendix A
Dissent: David B. Swoap

The panel has conducted a searching and thorough review of present quality control systems, and there is much with which I concur in its initial findings.

Its overall thrust, however, contains both premises and specific recommendations with which I must dissent. They are delineated below. I dissent from:

1. *The overly broad definition of quality control adopted throughout the report.*

The report contends that quality control should be expanded to include a number of additional policy goals and that "the current QC system is limited in what it measures and therefore does not serve well the purpose of policy makers or program managers . . ." (Chapter 6).

In fact, quality control in public assistance programs—as the term has been applied historically and properly is applied conceptually—should be defined as the effort to instill and measure precision in eligibility and payment determinations and to prevent the misallocation of taxpayer dollars resulting from erroneous determinations.

I specifically dissent from:

a. The belief that FNS should develop other important measures of performance for inclusion, such as timeliness or coverage of the eligible population.

b. In Chapter 2, the inference in "all three focus on recipient eligibility status and benefit payment accuracy. None focuses on the effects of the receipt of benefits or on precisely how those benefits are used with respect to the achievement of program objectives" that this division is inappropriate.

c. In Chapter 2, the inference in "QC systems now in place do not measure the effects of such [policy] choices or the achievement of program objectives other than accuracy" payment that the current emphasis is inappropriate; I believe it to be properly focused.

d. In Chapter 3, the inference in the description of the current QC system, similar to those described above.

e. The continuing emphasis, in the discussion in Chapter 6 contrasting the current system with the recommended enlarged system, on purposes beyond a proper QC scope.

f. The statement in the Summary and Recommendations that the "QC system lacks many of the elements of a comprehensive quality improvement system" and the statement in Chapter 6 that "the current system does not incorporate some essential elements of a quality improvement system."

g. Recommendation 1, that "FNS/USDA sampling and measurement activities for monitoring issuance accuracy should be made independent of" the others described. The former is integral to, if not identical with, the latter.

h. Recommendation 2, that "measures of quality should be set against broad program objectives, beginning with the language of intent in the Food Stamps Act. Federal performance measures should begin with payment accuracy but should be added to over time and as data from monitoring and evaluation activities permit greater breath and specificity."

2. *The effort to include underpayments and issuances not made or terminated with regard to eligibles.*

As stated above, quality control should be limited to the measurement of cases where tax resources have been misspent. Where an expenditure has not occurred, a misexpenditure has not occurred. However laudable defenses against underpayments

and issuances not made or terminated with regard to eligibles are (and I believe them to be,) they should not be included in proper quality control measurements nor weaken their application through an explicit or implicit "balancing" attempt.

I specifically dissent from Recommendation 10.

3. *The view that there should be "rewards" for good performance as well as sanctions for poor performance.*

Again, while on the surface it might appear more equitable or "balanced" to establish a system of rewards as well as sanctions, there should not be a need to expend additional tax dollars in "rewards" to achieve accurate or efficient administration. Rather, a sounder system of sanctions (which recognizes joint federal-state-local responsibility for error-causing provisions) should be pursued.

Specifically, I dissent from:

a. The view in Chapter 6 that differential sharing in the implicated benefit costs of errors should be structured to provide both positive and negative incentives, rewards for good performance as well as sanctions for poor performance.

b. Recommendation 8.

If the recommendation were modified to make it clear that the only rewards contemplated are only a *lesser* or scaled application of sanctions, or a retention of shifts in administrative cost-sharing, to recognize demonstrated or sustained performance, I would concur.

4. *The belief expressed that performance consequences should not be expected to accrue to large numbers of state agencies each year.*

The presence or absence of performance consequences should be an independent variable—that is, tied to the objective standard of performance and not to a subjective view as to expected incidence. To be sure, a continuing pattern of large numbers of states reporting errors or incurring sanctions year after year may be indicative that there is something wrong with the QC system (just as it led, in part, to the call for this study); it may equally be indicative of administrative performance that needs improving on a wide variety of fronts. It is a symptom that something needs

correcting; it should *not* be an automatic indictment of the performance measures nor something subject to arbitrary standards about expected incidence.

5. ***The view that technical errors should not be used for valuing misspent dollars.***

So-called "technical errors" can often be euphemisms indicators of programmatic problems and/or abuse. They should *not* be excluded from error computation.

Specifically, I dissent from:

a. The omission of this argument throughout the report.
b. The explicit rejection of this concept in Chapter 6.

6. ***The idea that coverage should be measured or used in quality control.***

With reference to the discussion in Chapter 6, I believe that it is inappropriate to use quality improvement systems either to measure coverage or to invoke performance measures relating to it.

The principal concern I have, and the reasons undergirding this dissent, are that I do not wish to see the *successes* of the current quality control system (See Figure 10, Chapter 2), which have been amply demonstrated since the establishment of the current sanction system, weakened or perhaps eliminated altogether by the much larger world view of quality control held by the remainder of the panel. Certainly the present system needs to be made more fair, more precise, and more reflective of joint governmental roles (in cause and effect with regard to errors). It does not need to be made unworkable or unproductive in the achievement of that goal that should be central: precision in eligibility and payment determination to prevent misallocation of tax resources.

Appendix B
Assessing the Statistical Fairness of the Food Stamp Quality Control System

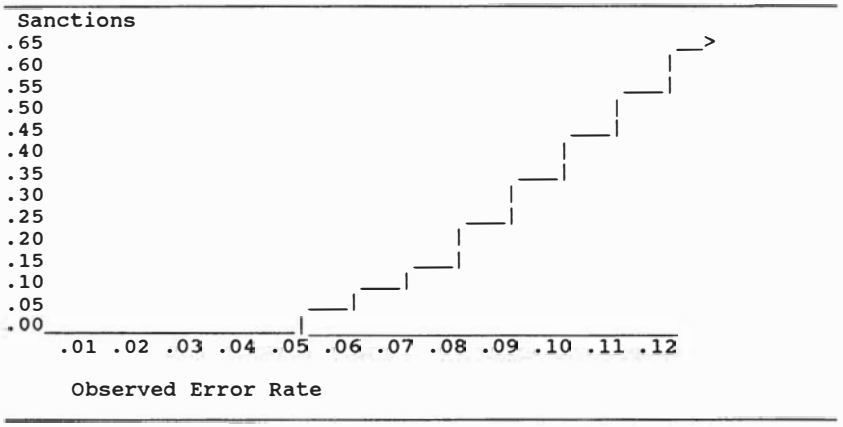
The Food Stamp Program makes use of a formula for imposing sanctions on states, based on estimates of state payment error rates from the quality control system. This formula is used when estimated payment error rates exceed a threshold error rate, currently set by law at 5 percent. To describe the formula precisely, let:

t	=	**the 5 percent threshold,**
D	=	**federally reimbursed state administrative costs,**
X	=	**estimated payment error rate in the sample, and**
S	=	**the sanction levied.**

Then:

$$
S = \begin{cases}
.05\,D & \text{if } t < X \leq t + .01 \\
.10\,D & \text{if } t + .01 < X \leq t + .02 \\
.15\,D & \text{if } t + .02 < X \leq t + .03 \\
.25\,D & \text{if } t + .03 < X \leq t + .04 \\
.35\,D & \text{if } t + .04 < X \leq t + .05 \\
.45\,D & \text{if } t + .05 < X \leq t + .06 \\
.55\,D & \text{if } t + .06 < X \leq t + .07 \\
.65\,D & \text{if } t + .07 < X \leq t + .08 \\
.75\,D & \text{if } t + .08 < X \leq t + .09 \\
.85\,D & \text{if } t + .09 < X \leq t + .10 \\
.95\,D & \text{if } t + .10 < X \leq t + .11 \\
1.00\,D & \text{if } t + .11 < X
\end{cases}
$$

Pictorially:



A statistically "fair" sanction percentage for this analysis is defined as the percentage that would apply if one could determine the true payment error rate from a state's quality control system and then apply the above formula to that true error rate. Since the true payment error rate is unknown, the estimated payment error rate is used in place of the true payment error rate to determine sanctions. The use of the estimated payment error rate results in a higher or lower sanction percentage to the extent that the estimate is higher or lower than the true rate. The question then arises as to whether the method of basing sanctions on the estimated payment

error rates is "fair," in a statistical sense: that is, over an extended period of time, would a state's sanctions be higher or lower than would result by the use of its true payment error rate? The panel notes that this conception of "fairness" is a very narrow one. It does not consider other grounds on which "fairness" may be questioned.

To answer this question the panel calculated, for selected true payment error rates and variances of estimated payment error rates for some hypothetical state quality control systems, the expected percentage of the federal reimbursement that a state would have to repay minus the percentage appropriate for the corresponding true payment error rate. That is, the panel computed:

$$[E(S(\text{est. error rate})) - S(\text{true error rate})] / D,$$

where E denotes mathematical expectation. These percentage differences, expressed as proportions, are presented in Table B-1. The various true error rates given there span most of the estimated error rates for states during the previous several years. Also, the standard deviations of the estimated error rates span the standard deviations that seem likely, given the state and federal sample sizes currently used.

For example, consider a state with true error rate of 11.9 percent (the third row from the bottom) and a standard deviation of 1 percent (.010) for its sample estimates, the second column. The sanction, if levied on the true error rate, would be 45 percent of the federal share of administrative costs. Because of the way that the step function is constructed, sampling error in this state's performance estimate will result in oversanction more often than undersanction: the next step for higher sanctions is just 0.1 of one percentage point above the state's true error rate; the next lower step is 0.9 of one percentage point below. The mathematically expected sanction percentage, then, is actually about 49 percent of the federal share of administrative costs, or 4 percent above the "fair" sanction. Similar calculations are the basis for the remaining entries. A minor simplification was used in the above calculations. FNS has a parallel incentive system for states that have estimated error rates below the threshold error rate, but these incentives were omitted from the computations. However, the magnitude of these payments are such that they would not substantially alter the main features of Table B-1.

TABLE B-1 Difference Between Expected and Statistically "Fair" Sanction Proportion (use of point estimate)

True Error Rate	Standard Deviation									
	.005	.010	.015	.020	.025	.030	.035	.040	.045	.050
.000	.00	.00	.00	.00	.00	.00	.01	.02	.02	.03
.015	.00	.00	.00	.00	.01	.01	.02	.03	.04	.06
.020	.00	.00	.00	.00	.01	.02	.03	.04	.05	.07
.030	.00	.00	.01	.01	.02	.03	.05	.06	.08	.09
.031	.00	.00	.01	.01	.02	.04	.05	.06	.08	.09
.042	.00	.01	.02	.03	.05	.06	.08	.10	.11	.13
.045	.01	.02	.03	.04	.06	.07	.09	.10	.12	.14
.053	-.01	-.00	.01	.02	.04	.05	.07	.09	.10	.12
.060	-.02	-.02	-.01	.00	.02	.03	.05	.07	.09	.10
.064	-.00	-.00	.01	.02	.04	.06	.07	.09	.11	.12
.075	.01	.02	.03	.04	.06	.07	.09	.10	.12	.13
.086	.01	.01	.02	.03	.04	.05	.06	.08	.09	.10
.090	-.05	-.05	-.04	-.04	-.03	-.02	-.01	.00	.01	.02
.097	.02	.02	.02	.03	.03	.04	.05	.06	.06	.07
.105	.00	.00	.00	.00	.01	.01	.01	.02	.02	.02
.108	.03	.03	.03	.03	.03	.04	.04	.04	.04	.04
.119	.04	.04	.04	.04	.04	.03	.03	.02	.02	.02
.120	-.05	-.05	-.05	-.05	-.05	-.06	-.06	-.07	-.07	-.08
.135	.00	.00	-.00	-.01	-.02	-.03	-.05	-.06	-.07	-.08

From the table, a few major points are clear. First, many of the situations are "unfair" to the state. Next, there are many combinations of true error rates and standard deviations for which the sanctions are "unfair" to FNS. In particular, for states with low standard deviations and error rates just above the threshold level or a step in the sanction formula, and for states with true payment error rates with sanction percentages near the maximum of 100 percent, the differences can be negative. Finally, it is apparent that there are realistic situations for which the difference can be substantial, certainly 10 percent or higher.

The differences in Table B-1 are moderate for some situations. If the standard deviations are small relative to the step width and the true payment error rate is fairly near the middle of a step, the differences are quite small.

One method that has been proposed for promoting "fairness" is to use the lower bound of a confidence interval around the error rate estimate to establish the magnitude of a sanction. This is essentially the lowest value for the error rate that one would

TABLE B-2 Difference Between Expected and Statistically "Fair" Sanction Proportion (use of lower confidence bound)

True Error Rate	Standard Deviation									
	.005	.010	.015	.020	.025	.030	.035	.040	.045	.050
.000	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.015	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.020	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.030	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.031	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.042	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.045	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
.053	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05
.060	-.07	-.09	-.09	-.10	-.10	-.10	-.10	-.10	-.10	-.10
.064	-.06	-.08	-.09	-.09	-.09	-.09	-.09	-.09	-.09	-.09
.075	-.05	-.10	-.12	-.13	-.13	-.14	-.14	-.14	-.14	-.14
.086	-.08	-.14	-.18	-.20	-.22	-.22	-.23	-.23	-.23	-.23
.090	-.15	-.22	-.26	-.29	-.31	-.32	-.32	-.32	-.33	-.33
.097	-.08	-.17	-.22	-.26	-.29	-.30	-.31	-.32	-.32	-.32
.105	-.10	-.20	-.27	-.32	-.36	-.38	-.39	-.40	-.40	-.41
.108	-.07	-.17	-.25	-.30	-.34	-.37	-.38	-.39	-.40	-.41
.119	-.06	-.16	-.25	-.33	-.38	-.42	-.45	-.46	-.48	-.49
.120	-.15	-.25	-.34	-.42	-.48	-.52	-.54	-.56	-.58	-.58
.135	-.10	-.20	-.30	-.39	-.47	-.53	-.57	-.60	-.63	-.64

find plausible given the estimate from the sample. To investigate this method, the panel repeated the above calculations using the estimated error rate minus two times the true standard deviation, which gives the lower limit of an approximate 95 percent confidence interval for the error rate. The differences calculated are presented, as proportions, in Table B-2. The fact that in actual use the variance would have to be estimated raises some practical difficulties, one of which is that estimates of the variance that are biased high would benefit the states in that the lower limit would be even lower. This problem is not evident in Table B-2 since the variances that are assumed to be true are used.

The results in Table B-2 are quite different from those in Table B-1. Instead of many situations being unfavorable to the states, now the majority of the situations are unfavorable to FNS. The states are not "unfairly" sanctioned under this system, but at a cost of excessive inequity when the true payment error rate is above the threshold rate. Therefore, instead of increasing "fairness," the change from using the point estimate to using the lower bound of a

confidence interval merely transfers the inequity to another party, and the inequity can be much larger than in the present situation. For this reason, once the decision has been made to sanction a state, the panel recommends using the point estimate to establish the magnitude of sanction liabilities (see recommendation 16).

However, the use of the point estimate results in relatively greater inequity to those states with true performance levels within a few percentage points of the lowest sanction threshold, 5 percent, for two reasons. First, states just under that threshold should not be sanctioned, yet at times they nonetheless will be sanctioned (see Table B.1 again). Second, the size of the penalty increases after the third step in the sanction formula. Therefore, for a state with a true payment error rate between 5 and 8 percent, an overestimate of performance that crosses the 8 percent step will generate an oversanction that is larger than the undersanction generated by an equal-sized underestimate. Therefore, the panel recommends the use of a lower bound to determine whether a state should be sanctioned (see recommendation 15).

Appendix C
Biographical Sketches of Panel Members and Staff

JOHN NETER is the C. Herman and Mary Virginia Terry professor of management science and statistics in the Department of Management Sciences and Information Technology, College of Business Administration, University of Georgia. He received a B.S. from the University of Buffalo, an M.B.A. from the University of Pennsylvania, and a Ph.D. in applied statistics from Columbia University. Previously he taught at Syracuse University and was chair of the Department of Quantitative Analysis, and he was a professor of management science at the University of Minnesota and chair of the Department of Quantitative Analysis. He has served as the editor of *American Statistician* and associate editor of *Decision Sciences*. A fellow of the American Statistical Association, he also served as director of the association from 1975 to 1980 and as president in 1985. He is also a fellow of the Decision Sciences Institute, which he served as president in 1978-79, and of the American Association for the Advancement of Science.

HENRY AARON is a senior fellow at the Brookings Institution in Washington, D.C. and professor of economics at the University of Maryland, College Park, Maryland. He received an A.B. from

the University of California, Los Angeles and M.A. and Ph.D. degrees from Harvard University. In 1978-79 he was chair of the Advisory Council on Social Security, and in 1977-78 he served as assistant secretary of planning and evaluation, U.S. Department of Health, Education and Welfare. He is a member of the Board of Trustees of Teachers Insurance and Annunity Corp. and the Board of Directors of ABT Associates, Inc. He is the associate editor for the *Journal of Economic Perpsectives* and the *Journal of Health Economics* and a member of the editorial board for *Public Finance Quarterly* and *Knowledge*.

MARY JO BANE is professor of public policy at Harvard University. She received a B.S. in foreign services from Georgetown University and an M.A. in teaching and E.D. from Harvard University. She has served as executive deputy commissioner for the New York State Department of Social Services; deputy assistant secretary for program planning and budget analysis for the U.S. Department of Education; and associate director at MIT-Harvard Joint Center for Urban Studies. Her research interests focus on public policies on poverty, welfare, and families.

LEON GILFORD is vice president and member of the board of directors of COBRO Corporation, Wheaton, Maryland. He received a B.A. in mathematics from Brooklyn College and an M.A. in mathematical statistics from George Washington University. His major professional interest is in the design of mathematical models for use in the control of social and physical systems. As a mathematical statistician at the Bureau of the Census, he designed and implemented the quality control procedures for the 1950 and 1960 censuses. Other positions he has held include special assistant for reliability in the Atomic Energy Commission, chief statistician and director of automated data processing at the U.S. Tariff Commission, and principal scientist at Operations Research, Inc. He is a fellow of the American Statistical Association and the American Association for the Advancement of Science and a member of the International Statistical Institute. He has been a member of advisory committees to the Energy Information Administration, U.S. Department of Energy, and to the National Center for Education

Statistics, U.S. Department of Education. He has served as president of the Washington Statistical Society and of the Washington Section of the American Society for Quality Control.

LEON GINSBERG is Carolina research professor of social work at the University of South Carolina. He received a B.A. from Trinity University, an M.S.W. from Tulane University, and a Ph.D. from the University of Oklahoma in political science. He was the recipient of the Distinguished Service award of the West Virginia Welfare Conference in 1970 and was named the West Virginia Social Worker of the Year for 1978. He was previously an associate professor at the School of Social Work, University of Oklahoma; professor, director, and dean of the School of Social Work at West Virginia University in 1968-77; West Virginia commissioner of human services in 1977-84; chancellor of the West Virginia Board of Regents for Higher Education from 1984-86; and Fulbright professor at the University of Pontificia Bolivariana, Medellin, Columbia in 1974. He was president of the American Public Welfare Association in 1984-85. Research interests include management of social welfare programs, Latin American welfare programs, and rural social work.

DANIEL HORVITZ is executive vice president of the Research Triangle Institute in North Carolina and adjunct professor of biostatistics at the University of North Carolina. He received a B.S. from Massachusetts State College and a Ph.D. in statistics from Iowa State College. His major work is in sampling theory, sample survey design and methods, measurement error research, randomized response, longitudinal surveys, demographic simulation models, and social program evaluation. Memberships in professional organizations include fellow of the American Statistical Association, which he currently serves as vice president, fellow of the American Association for the Advancement of Science, the International Statistical Institute, the American Public Health Association, and the Population Association of America. He chaired the National Research Council Panel on Statistics for Family Assistance and Related Programs in 1981-83.

GERALD LIEBERMAN is professor of operations research and statistics at Stanford University. He received a B.M.E. from

Cooper Union, an A.M. from Columbia University, and a Ph.D. in statistics from Stanford University. He was previously vice provost and dean of graduate studies and research, associate dean of the School of Humanities and Sciences, and chair of the Department of Operations Research at Stanford University. He served as a mathematical statistician for the National Bureau of Standards. His professional interests include quality control, reliability, risk, and mathematical models of stochastic systems, in general. He was a fellow at the Center for Advanced Study in the Behavioral Sciences in 1985-86. Among his professional affiliations, he is a fellow of the American Society for Quality Control (where he was awarded the Shewhart Medal in 1972), a fellow of the American Statistical Association (where he served as vice president in 1963-64), a fellow of the Institute of Mathematical Statistics, a member of the Institute of Management Science (where he served as president in 1980-81), and a member of the Operations Research Society of America. He is also a member of the National Academy of Engineering.

CARRIE MENKEL-MEADOW is professor of law at the University of California, Los Angeles. She received an A.B. in sociology from Barnard College and a J.D. from the University of Pennsylvania. She previously taught at the University of Pennsylvania Law School and was a staff attorney for Community Legal Services, Inc., specializing in government benefits and employment issues. Her research has focused on the delivery of legal services, specifically looking at public policy issues affecting legal services for the poor, dispute resolution, particularly negotiation and other alternatives to ligitation and issues involving social science and law. She has served on the Board of Trustees of the Law and Society Association, the Society of American Law Teachers, the Western Center on Law and Poverty, among other institutions. She has been a consultant to the American Bar Association, the Center for Public Resources, the Center for the Study of Dispute Resolution, the Legal Services Corporation and many other public and private organizations.

HELEN O'BANNON is senior vice president, University of Pennsylvania. She received a B.A. in economics from Wellesley, an M.A. in economics from Stanford University, and has completed course

work on a Ph.D. in business from the University of Pittsburgh. Other positions she has held include secretary of the Pennsylvania Department of Public Welfare; commissioner for the Pennsylvania Public Utility Commission; research economicst for the U.S. Comptroller of the Currency; fiscal economist for the U.S. Secretary of the Treasury; and research assistant with the U.S. House of Representatives Banking and Currency Committee. She has also taught and lectured at Bryn Mawr, the University of Pennsylvania, and Chatham College.

JAMES PRESS is professor of statistics at the University of California, Riverside, and was chair of the Department of Statistics in 1977-84. He received a B.A. from New York University, an M.S. from the University of Southern California, and a Ph.D. in statistics from Stanford University. He previously taught at the University of Chicago, the University of British Columbia, the University of California, Los Angeles, and Yale University. His research interests have centered on multivariate statistical analysis, Bayesian statistics, econometrics, and applications of statistics to the social sciences. He is a fellow of the Institute of Mathematical Statistics, the American Statistical Association, the Royal Statistical Society, the American Association for the Advancement of Science, and is a member of the International Statistical Institute, and the Bernoulli Society. He is currently a member of the Committee on National Statistics of the National Research Council.

SAMMIE LYNN PUETT is the associate vice president of university relations and professor of journalism at the University of Tennessee. She received a B.S. in business administration and an M.S. in educational psychology from the University of Tennessee. From 1980 to 1985 she served as commissioner of the Tennessee Department of Human Services, and she served as commissioner of the Department of General Services in 1979-80. She chaired the Governor's Cabinet on Social Services from 1983 to 1985 and was a member of the Tennessee Medicaid Medical Advisory Committee and the Tennessee Commission on Aging from 1980 to 1985. She also served on the Governor's Task Force on the Prevention of Mental Retardation from 1980 to 1983 and the Governor's Task Force on Healthy Children from 1983 to 1985. Prior to her state

government services, she was executive assistant to the Chancellor at the University of Tennessee, Knoxville, editor of *Tennessee Town & City* magazine, and a columnist for *Nation's Cities* magazine. She is an accredited member and a former national board member of the Public Relations Society of America. She serves on the executive committee of the White House Commission on Presidential Scholars.

BERNARD STUMBRAS is currently visiting fellow at the Institute for Research on Poverty at the University of Wisconsin and was previously assistant administrator of the Division of Community Services. He received a B.S. and M.S.W. from the University of Wisconsin. He was involved in numerous social work programs from 1955 to 1982 and was a consultant to the U.S. Department of Health, Education, and Welfare and various states on eligibility simplification and use of computer systems for income maintenance programs. He is a member of the Executive Committee of American Public Welfare Association, its Ad Hoc Committee on Welfare Reform, and its Board of Directors.

DAVID SWOAP is a partner of the government relations firm of Franchetti and Swoap (San Francisco and Washington). He is a member of the American Public Welfare Association and the United Council on Welfare Fraud. He received a B.A. in government from Denison University, an M.A. in government from Claremont, and a Dr. Sci. from the University of Osteopathic Medicine and Health Sciences. He was previously the director of the California Department of Social Welfare and Benefit Payments; chief of presidential transition at the Social Security Administration; staff to the U.S. Senate, Committee on Finance; the undersecretary of the U.S. Department of Health and Human Services in 1981-83; and secretary of the California Health and Welfare Agency from 1983-85.

JUDITH TANUR is professor, Department of Sociology, State University of New York at Stony Brook. She received a B.S. in psychology, a M.A. in mathematical statistics from Columbia University, and a Ph.D. in sociology from SUNY at Stony Brook. She is a fellow of the American Statistical Association and the American Association for the Advancement of Science and an elected

member of the International Statistical Institute. She edited the *International Encyclopedia of Statistics* and *Statistics: A Guide to the Unknown*. She is currently a member of the Committee on National Statistics of the National Research Council and served on its Panel on Statistics for Family Assistance Programs and chaired its Advanced Research Seminar on Cognitive Aspects of Survey Methodology. She is currently cochair of the Social Science Research Council's Committee on Cognition and Survey Research and book review editor of the *Journal of the American Statistical Association*.

DENNIS P. AFFHOLTER serves as the panel's study director. He received a B.G.S. from the University of Michigan and an M.A. in political science from the University of Wisconsin. He was previously director for planning and analysis with the Maryland Department of Human Resources; a senior fellow for the Center for Policy Research, National Governors' Association; and a research associate for the Panel on Statistics for Family Assistance and Related Programs of the Committee on National Statistics of the National Research Council.

FREDRICA D. KRAMER serves as the panel's associate study director. She received a B.A. in political science from the University of California at Berkeley, an M.U.P. from Hunter College, City University of New York, and has completed all but the dissertation for a doctorate in public administration from the University of Southern California. She has worked in social welfare policy in state, local, and federal governments, recently focusing on welfare and employment policy at the office of the assistant secretary of planning and evaluation at the U.S. Department of Health and Human Services and at the U.S. General Accounting Office. She was formerly with the Committee on Law Enforcement and Criminal Justice of the National Research Council.

GARY CAUSER is research associate for the panel. He received a B.S. and M.S. in mathematics from Indiana University of Pennsylvania, an M.S. in statistics from Ohio State University, and an M.S. in computer science from Johns Hopkins University. He is currently working on a Ph.D dissertation in mathematical statistics at George Washington University. Previously he was senior

statistician for the W.R. Grace and Company and statistical engineer for Corning Glass Works. In these positions he served as a company consultant in the areas of applied statistics and quality control.

JUDITH A. UHLMANN served as research assistant for the panel. She received an M.A. from Antioch University in planning and administration. She was previously project manager for the Department of Human Resources, State of Maryland, and consultant with D'Amico Associates.

ELAINE McGARRAUGH is research assistant for the panel. She received a B.S. from McMurry College in social sciences (education). Her previous National Research Council positions include staff associate for the Committee on Population; production editor for the Committee on Population and Demography; editorial assistant with the Assembly of Behavioral and Social Sciences; administrative associate for the Committee on Aging; and research assistant for the Institute of Medicine.