

MEASURING UI BENEFIT PAYMENT ACCURACY UNDER UI PERFORMS: PROPOSED CHANGES TO BENEFITS QC

BACKGROUND

Since 1987, States have been required to measure the accuracy of benefit payment determinations by operating a Benefits Quality Control (BQC) program. BQC was modelled closely on the Random Audit (RA) program developed in the early 1980s, which was the Department of Labor's first attempt to estimate each State's rate of dollars overpaid and underpaid by systematic review of all aspects of paid claims. The RA approach which used random sampling methodology revealed more "accurate" error rates, totalling nine times those indicated by less rigorous current methods. As was the case under the random audit approach, QC investigators verify all information pertaining to each sampled case by reviewing agency records and contacting the claimant, employers and third parties. The object is to determine what the payment would have been had the agency been aware of all pertinent information and correctly applied State law and policy. To form the basis for program improvements and related programmatic analyses, over 100 data elements relating to the nature of the case, types and causes of any errors found are coded into a computerized database. In order to inform interested parties and involve them in seeking improvements where warranted, States must publically release summary error rate statistics annually.

Although the samples--ranging from 480 cases per year in the seven smallest States to 1800 in California--are tiny fractions of the number of paid claims, BQC is considered to be more resource intensive than other UI data collections. This is because the accuracy data are collected by interview or other contact and cannot be assessed from data which are computer-generated byproducts of agency operations or even from simple records reviews. Each case requires about 10.5 hours to investigate. In FY 1995, the cost of BQC was \$26 million, 7 percent of the total UI integrity budget and 1.3 percent of total UI administrative costs.

From its onset, BQC has faced numerous criticisms, among them:

- ! Some States, particularly smaller ones, consider BQC to represent a disproportionate fraction of their overall UI staff;
- ! States believe that, at current administrative funding levels, they cannot control or even substantially influence the employer and claimant actions causing most of the mispayments BQC identifies, yet the Agency is considered responsible for them;
- ! Some States have asserted that BQC tends to find the same type and level of errors year after year, implying both an irreducible minimum of error not subject to correction and also no need to measure accuracy as frequently as BQC does;
- ! Some stakeholders feel that since DOL never implemented a QC measurement of denied claims accuracy, BQC findings are unbalanced (they

necessarily understate the true underpayment rate, reflecting only payments that were too small, and not the more important erroneous denials);

! States claim that BQC's public data release often induces them to eliminate or weaken useful (but error-causing) features of their laws and policies to reduce reported errors;

! Despite Departmental efforts to encourage program improvement activities by reducing sampling and providing grants, States feel too much of BQC's resources are still devoted to measurement.

In Spring 1993, Vice President Gore's National Performance Review (NPR) singled out the BQC program as a DOL issue (Appendix 1). Many of the above criticisms were cited in the issue. In response, the Department agreed to "reexamine the present mix of systems for improving the performance of the unemployment insurance program and devise a unified strategy that improves its effectiveness." This review would include a "substantial focus" on BQC and "how its resources can best be divided between measurement, analysis, and direct support for program improvement." DOL also agreed to review whether BQC should look at denied claims, and whether it should include other "activities designed to strengthen the quality of the unemployment insurance program."

In October 1993 a joint State-Federal Performance Enhancement Workgroup (PEWG) held its first of sixteen meetings. Its primary agenda was the first part of the NPR charge--"reexamining the present mix of systems for improving the performance of the unemployment insurance and devise a unified strategy that improves its effectiveness" in the light of new principles for State-Federal cooperation. With that completed, however, it reached consensus in September 1995 on proposals for changes in the scale and focus of measuring benefit payment accuracy in the new "unified strategy" which it terms UI Performs. In considering further the implications of the UI Performs initiative as a whole, the Department has refined the budgetary aspects of the PEWG's proposal. The rest of this paper explains those changes and their implications.

APPROACH TAKEN

The PEWG began this task by first identifying and agreeing upon a set of interests. It then brainstormed a series of concrete options, which members took as the basis for further staff development. Staff analyzed the brainstormed options and identified thirteen dimensions which were deemed critical to assessing approaches to measuring benefit payment accuracy, and thus to reconsidering BQC.

1. Interests. Interests can be seen as criteria for evaluating a program, statements about what the program is to accomplish or constraints to which it must conform. Briefly, the interests the PEWG agreed on are summarized below:

a. Objectively Measure the Accuracy of Benefit Payment Determinations (both positive and negative).

* This information is useful to and desired by groups both inside and outside the UI system.

b. Improve the Accuracy of Benefit Payment Decisions Where Warranted.

c. Balance Costs of Performance Measurement and Improvement With Costs of Conducting Operations

2. Design and Evaluation Dimensions. The committee used the following dimensions in reviewing the existing BQC program and arriving at a proposed reconfiguration.

- a. Program Coverage
- b. Nature of Accuracy
- c. Type of Decisions Covered
- d. What to Review: Decisions/Outcomes or the Process?
- e. How Parties are Contacted
- f. Program/Review Cycle
- g. Sample Size
- h. Nationally Standard Methodology?
- i. Accuracy Data: Macro or Micro?
- j. Types of Data in Accuracy Database
- k. Quality Assurance (QA): Sampling
- l. Quality Assurance (QA): Investigations
- m. Restrictions on Use of Freed-Up Staff

Staff also developed dimensions for measuring costs and, by using the PEWG interests, for evaluating the revised program. Where possible, evaluation factors were quantified.

SUMMARY OF THE PROPOSAL

In brief, it is proposed that payment accuracy be measured very much as it is now through BQC, with the following changes:

! Sample sizes will be reduced from an average of 810 per State annually and a range of 480 to 1800, to 360 in the ten Smallest States and 480 in all others;

! States will have the option of verifying all claims information by phone/mail/fax;

! Denials will be investigated, probably using the paid-claims audit methodology, after pilot testing is completed and its results analyzed;

! The existing BQC data record will be reviewed carefully with the object of reducing its size;

! Data will be reported in "Building Block" fashion to help users more readily see the effect of different State laws and policies on overpayment and underpayment rates;

! States will no longer be required to release payment accuracy (overpayment and underpayment rates) in the public media; and

! The formulation and allocation of resources for BQC will change somewhat as a result of the changes in required sampling. In addition, the formulation and allocation of these resources will be done within an umbrella category of UI Performs. In this budget category, resources for a broad range of measurement, validation, and other activities pertinent to bringing about continuous improvement in States will be budgeted.

The following section explains which dimensions were left unchanged and which were altered to construct the proposed new accuracy measurement vehicle, and why the decisions were made.

CONSTRUCTION OF THE PROPOSAL

This section uses the design dimensions to outline what was changed and what was retained in reconfiguring BQC, and why. Table 1 shows the differences in design dimensions at a glance.

1. Unchanged Dimensions.

a. Coverage. This refers to the UI programs or program areas whose accuracy is assessed. BQC currently covers the intrastate portions of the largest permanently-authorized UI programs: the regular State program plus Unemployment Compensation for Federal Employees (UCFE) and Unemployment Compensation for Ex-Servicemen (UCX), and including Combined Wage Claims (CWC), intrastate benefits whose monetary amounts are based on wages earned in more than just the paying State. Use of these payment categories comprised approximately 95% of all UI payments in 1994 and effectively allowed the review of some five sixths of all eligibility determinations (most of the unreviewed actions are intrastate denial actions). The main continuing "program" not covered is the Interstate Benefits program, omitted because of the complexity of the review design. No change is proposed along this dimension.

b. Nature of accuracy. BQC's standard of accuracy, as noted in the introduction, is whether the agency reached the correct, *fully-informed* decision. That is, it assumes that an accurate decision is one which complies with written State law and policy,

reached by a SESA staffer aware of all pertinent information from all parties. The revised system retains this notion of accuracy, for two main reasons:

TABLE 1

	DESIGN DIMENSIONS	Current BQC	Recommended Change
1	COVERAGE: UI, UCFE, UCX & CWC	YES	YES
2	NATURE OF ACCURACY		
	Fully Informed Verifications	YES	YES
3	TYPE OF DECISIONS COVERED		
	Payments	YES	YES
4	WHAT TO REVIEW		
	Review Payment Decisions	YES	YES
5	HOW INVOLVED PARTIES ARE CONTACTED	Mixed Mode	Flexible
	Mixed Mode	YES	YES
6	PROGRAM CYCLE		
	Continuous	YES	YES
7	SAMPLE SIZE	480-1800	360-480
8	STANDARD METHODS		
	Specified National Standard	YES	YES
9	TYPE OF DATA COLLECTED		
	Micro	YES	YES
10	INFORMATION CATEGORIES		
	Accuracy and Error Details	YES	YES
	Information for Cont. Improvement and Policy	YES	YES
	Sampling Demographics	YES	YES
11	QUALITY ASSURANCE - SAMPLE	YES	YES
12	QUALITY ASSURANCE - INVESTIGATIVE	YES	YES

Recommended Change: 10 smallest States get 360 cases; all others get 480. DCI Elements reviewed with intent of reducing overall number. Smallest defined as those with fewest weeks claimed over past seven years as reported on ETA 5159 report.

! The Federal Government has a strong interest in program payment accuracy in general, and thus in ensuring that the UI system has information on the extent to which the system is making fully informed decisions in accordance with law and policy. The UIS adopted the Random Audit program in the early 1980s after pilot tests revealed the extent of mispayments made because SESA records contained incomplete information.

! The nature of the UI system is such that claimants (and, to a lesser extent, employers) control some of the information critical to making a correct eligibility determination. SESA administrative resources are insufficient to verify this outside information routinely. Nevertheless, SESA and other public policy-makers have a responsibility to know the extent to which they are tending to get biased information; even if SESAs cannot control the quality of all information, they can often influence and improve it. SESAs are not likely to improve what they do not know about.

This choice dictates the nature of the accuracy review. It requires an attempt to validate all information pertinent to the cases it reviews, doing new and original factfinding to ensure that its judgments about the correct payment reflect full information. This involves contacting claimants, employers and third parties in addition to reviewing SESA records. Over the past two decades, this has become the standard approach used for quality control reviews of major social programs.

c. Type of Decisions Covered. BQC currently samples only paid claims (implying that all decisions leading to the decision to pay a week were likewise positive). The committee agreed that ultimately denials should be assessed if feasible. The Department is now designing a pilot for measuring denials, the results of which will guide the decisions on the best way to conduct this review and estimate its cost. The current assumption is that denials accuracy will be assessed the same way as payment accuracy.

d. What to Review (Decisions or the Decisionmaking Process.) The ultimate object of accuracy measurement is to determine whether processes need improvement or not. All aspects of benefit quality measures used by UI have all sought to do this by making inferences from the review of *estimation* samples of the processes' ultimate *outputs*. If significant numbers of the outputs are inaccurate, one infers a deficiency in some part of the process and then seeks to determine whether it ought to be changed. This approach is retained in the proposed redesign. It is of proven statistical validity, enables calculation of an error rate which is of interest to many stakeholders of the UI program, and provides a wealth of related information for a variety of analyses.

e. Program Cycle. BQC could be said to be on an annual cycle, taking weekly samples, year after year. Taking intermittent accuracy measurements--measuring the accuracy of payments every other year or every third year instead of continuously--has been suggested as a way to conserve resources. The PEWG deemed it preferable to save resources instead by taking continuous measurements, but at a lower sampling rate. All statistical process control programs oriented toward continuous improvement presume continuous measurement. Continuous improvement is one of the cornerstones of UI Performs.

f. Nationally Standard Methodology? BQC has always operated according to a nationally uniform set of procedures, intended primarily to ensure investigative integrity and minimize nonsampling error. Although the interpretation of the error rate varies from State to State because the provisions of State law differ, the overall meaning of accuracy is the same in all States because procedures are uniform, and it is sound to compute a national "roll-up" of dollars overpaid and underpaid. Under UI Performs, accuracy measures will continue to be obtained according to standard national methodology, following the example of the other performance data gathered as part of the UI Performs system.

g. Accuracy Data: Macro or Micro? Most UI data obtained from the States are macro data--a single summary percentage or count. The BQC database, on the other hand, contains a wealth of data, including the investigative findings, on every sampled case. Such data are routinely produced by the accuracy investigation and are useful for a variety of analyses by both State and Federal staff. The cost of entering the data, transmitting them and maintaining the database are quite small and their analytical value much greater than macro data. Thus, it was agreed States will continue to collect data for analysis at both the State and Federal levels.

h. Quality Assurance--Sampling. BQC carefully monitors all major aspects of its sampling process to ensure their accuracy and integrity. Its quality assurance process is highly automated and once programmed, very inexpensive to operate. The present approach will continue.

i. Quality Assurance--Investigations. BQC attempts to minimize investigative (nonsampling) errors through reviews of State QC process and regular case reviews by DOL National Office (NO) and Regional Office (RO) staff. These have major resource implications, especially for DOL RO staff. In July 1995, based on joint NO-RO work, required Regional Office case and procedures reviews were halved so that staff could emphasize data use and program improvement with their States. National Office staff are likewise deemphasizing oversight in favor of promoting program improvement in their contacts with State and Regional staff. These emphases will continue under UI Performs.

2. Dimensions that Involve Change.

j. *How Involved Parties are contacted.* Incorporating results of a 4-State "telephone pilot", BQC adopted a "mixed mode" approach in July 1993 designed to save resources while maintaining the thoroughness of the verification. Most phases of the investigation are done by mail/fax/telephone, with the claimant interview and work search verifications still done mostly in person.

! Change: *The revision proposes allowing the State complete flexibility in how the verifications are to be done, realizing that some States will do all verifications by mail/fax/telephone, whereas others will continue the present mix of approaches. The tradeoff--potential savings of some 3 hours per case investigation time, plus travel costs, and no need for extra positions in 7 States to verify data pertaining to other States' cases, at a cost of roughly 1/6 fewer errors detected--seems acceptable. Allocations assume that States would continue to use the present methodology.*

k. *Types of Data in the Accuracy Database.* The current BQC data record (Data Collection Instrument, or DCI) contains over 100 elements, in the following major categories: data on payment accuracy and information on the types, causes and responsibilities of errors; other data to support continuous improvement (CI); data to support program analysis (often the same as needed for CI); data on demographics for sample selection and validation; and other demographic data. The size of the record is not a major resource issue as data entry is a small part of the total case time. A larger data record does imply more Federal programming effort and monitoring.

! Change: *DOL will initiate a review, involving State staff, of existing elements from the standpoint of their past and prospective use and extent to which they support the role of accuracy data in the UI Performs system, and the costs of making changes. The intent will be to reduce the number of elements if possible.*

l. *Sample Size.* BQC samples average 810 cases per State per year and range from 480 in the seven smallest States to 1800 in California. Sample size is a major resource issue. In varying sample size, one trades off resources against precision. Also, at small sample sizes, certain data elements which refer to infrequent occurrences may become useless if they require making inferences from very small subsamples.

! Change: *The proposal calls for reducing samples for measuring the accuracy of paid claims to 480 cases per year in the 42 largest States and 360 in the ten smallest. This sampling rate requires 198 investigator staff vs. the 354 now required for BQC, no extra supervisors and--if States decide not to take advantage of the flexibility on investigative methods--about half the travel dollars.*

m. *Use of Freed-Up Staff.* Under BQC, if States are approved to conduct a special study or program improvement study they may theoretically free up all but the staff needed to investigate 400 cases annually. In practice, however, over the past few years, only a small fraction of staff has been diverted for special or PI studies.

*! **Change:** Under the proposal, staff freed up because of sample reductions or changes in how verifications are conducted will be available for investigating denied claims and other UI Performs activities, including taking other performance measurements.*

3. Publication and Release of Data.

Although not strictly a part of how the measurement program is designed and operated, two important changes in how the data should be used in the public arena are proposed. Under BQC, each State is required to release its yearly average overpayment and underpayment rates through the public media. One month later, DOL publishes a compendium of all State releases, along with additional analyses of QC data. The "annual report" error rates are simple totals, excluding only a few classes of actions (affecting about one quarter of the States) from the definition of error, although States are free to elaborate on or qualify the standard release as they wish.

*! **Changes:** The following are proposed:*

- 1. States would not be required to release accuracy findings through the public media; and*
- 2. The Department would publish accuracy findings in a compendium of all UI Performs data, using a components or "building block" form so that readers could be more readily see the effect of various State differences, particularly in law and policy, on measured overpayment and underpayment rates.*

4. Allocating Resources Under UI Performs.

Since its beginnings as a mandatory program in 1987, BQC has been a separately identified program with its own budget formulated and allocated to ensure States received sufficient resources to meet its required workload. UI Performs presumes a much different environment with changed emphases and therefore different budgetary implications. This proposal contemplates reducing resource demands for measuring paid claim accuracy by up to 50%; at the same time, measurement effort will be increased for denied claims and benefits timeliness and quality, and greater effort is assumed for analysis and program improvement. The resources previously identified as "Benefits QC" must thus be subsumed into a new UI Performs budget construct, and allocations designed to provide all States with equitable shares of resources intended to fund the measurement, analysis, planning and actions which will foster the UI Performs goal of continuous improvement.

IMPLICATIONS OF THE CHANGES IN ACCURACY MEASUREMENT

The implications of these changes *from the standpoint of the traditionally-conceived "BQC Budget"* are displayed in the table below. They indicate how the proposal would change the demands on that set of resources. The most significant are as follows:

1. Costs for Measurement. BQC currently requires 481 staff nationwide to measure paid claim accuracy. The proposed sample sizes can be accomplished with an estimated 179 fewer positions with no change in methodology, and will also require fewer travel resources. Offset against these savings are a currently estimated 104 positions nationwide to measure denials accuracy. This program is expected to start in three years. Not shown, but as noted above in this paper, UI Performs also implies staffing and travel to obtain the new benefits timeliness and quality measures being implemented in 1996 to replace the Quality Appraisal measures and to conduct data validation. Measuring tax administration accuracy, timeliness and quality measures is still estimated to require about one staff year per State.

2. Resources Available for Program Improvement. The changes in sample size proposed above will make available to States resources for the rest of UI Performs activities, including program improvements. The PEWG recommended that allocations for measuring benefit accuracy assume that States will *not* exercise their full flexibility. Exercising the option to verify payment accuracy totally by telephone, fax and mail methods would release some additional resources for other UI Performs. It is anticipated that the ultimate allocations of UI Performs resources will provide both positions and travel for continuous improvement activities after allowance for the other measurement and validation activities which UI Performs will require. In addition, in the three years before the measurement of denied claim accuracy begins nationwide, the resources ultimately needed for verifying the accuracy of benefit denials will be available for program improvements (or for conducting the denials pilot project).

3. Reduced Investigative Thoroughness and Precision. The proposed reductions in sample size will reduce the precision of all estimates and make inferences about certain subcategories of errors riskier. The 1990 Telephone pilot project indicated that doing the claimant interview and worksearch verifications by telephone would lose about half of their error detection power, so if States eliminate in-person contacts entirely the estimated error rate will fall by up to 15 percent. The review of the DCI may eliminate some data elements but the extent of the lost potential for analysis is not known now and may never be quantified.

a. Effect of Flexibility on Measured Error Rate. The following table compares the present national average error rate and its major cause components with what would be expected, based on the 1990 Telephone Pilot, with verifications conducted by all phone/fax/mail. The all-phone results can be seen as the probable lower limit or "worst-case scenario" of full flexibility in methodology.

Percentages of Dollars Overpaid, 1994

	<u>Actual 1994</u>	<u>Estimated, All Phone</u>
Benefit Year Earnings	2.30	2.07
Base Period Wages	1.03	1.02
Separations	1.52	1.40
Work Search	1.54	0.99
Other Eligibility Issues	1.81	1.59
All Other Issues	<u>0.34</u>	<u>0.30</u>
Total	8.54	7.37

b. Effect on Precision. The States are distributed by percentage increases in sampling error as follows:

<u>Increase in Sampling Error</u>	<u>Number of States</u>
0-10 percent	4
11-20 percent	13
21-30 percent	10
31-40 percent	10
41-50 percent	6
51-60 percent	7
61+ percent	2

For all but three States, samples are smaller and estimates thus less precise. The average (median) State will see its error of estimate rise by about 30 percent. The smallest States now sampling at 480 cases per year whose samples will fall to 360 cases will see sampling error rise by about one seventh. The greatest rise in sampling error would be experienced by the present large-sample States. For the eight largest States, increases in sampling errors will range from 50 to 100 percent.

State-by-State detail on precision is not shown but is available on request from Burman Skrable at (202) 219-5220 [fax: (202) 219-8506].

c. Ability to Guide Program Improvements. The smaller samples, giving reduced precision and fewer error cases to analyze, will make program improvement actions somewhat riskier. That is, they make it likelier both that some problems will be underestimated (and not addressed) and others may be overestimated (and thus perhaps be acted on when they should not be.) On the whole, the proposed changes give reasonable precision in identifying the two problems causing overpayment errors.

As a rule of thumb, it is assumed that in the aggregate a State would have to accumulate roughly 100 cases to conclude it had a major new accuracy problem. The time to accumulate 100 cases thus indicates, approximately, how quickly its accuracy measurement system can respond. Under BQC at present, this time

varies from three weeks in the largest State to 9 weeks in the smallest. Under the proposed revision it will grow to 9 weeks in the 42 largest States to 14 weeks in the 10 smallest.

The ability to conduct program improvement actions may also be affected by decisions on the size and content of the case record. This review has yet to be undertaken.

4. The Effect of Building Blocks. The table below shows, for national totals, a possible "building block" scheme for reporting 1994 BQC error rate data. The broadest possible rate (which includes as errors formal warnings given for worksearch violations and cases under appeal) is not now reported. All other categories are reported in the Annual Report rate (Rate 2). Three others are shown: errors on which States may legally act; legally actionable errors less worksearch; agency responsible only; and fraud only. Their national averages are:

1994 National Average Overpayment Rate

1.	All Actions, Causes, Responsibilities	9.04%
2.	1 Minus Formal Warnings & Appeals (Ann. Report Rate)	8.54%
3.	2 Minus Not Legally Actionable by State	6.76
4.	3 Minus Work Search Errors	5.70
5.	Agency Solely or Partly Responsible	1.88
6.	Fraud Only	1.84

Other divisions could be more advantageous and the Department continues to consider alternatives; it is hoped that comment on this proposal may suggest others. The effects of these different "cuts" vary considerably from State to State. Although State-specific data are not included in this proposal, the data will be provided on request. Contact Burman Skrable at (202) 219-5220 [fax: (202) 219-8506].

TIMING OF IMPLEMENTATION

The Department intends to begin implementing the changes as soon as comments can be considered; however, most of what is proposed cannot be effected immediately. The proposed changes in methodology and sample size are significant enough that under the Paperwork Reduction Act the Office of Management and Budget (OMB) must first clear them. Although OMB has been briefed on the proposed changes and has approved them in principle, the formal "paperwork process," which may take 6-8 months, is required. Changes involving the Data Collection Instrument and the use of a "building block" approach to data release require further consideration and possibly programming changes. Denied claims accuracy measurement will be pilot tested before nationwide implementation. With these factors in mind, the following target dates have been tentatively set:

Elimination of State Release	June 1996
Sample size reductions	Summer 1996
Verification flexibility	Summer 1996
Denials Accuracy (Pilot Begins)	October 1996
"Building Block" Data Breakdown	January 1997
DCI Review/Changes	January 1997
Denials Accuracy (Nationwide Start)	October 1999

Change the Focus of the Unemployment Benefits Quality Control Program to Improve Performance

BACKGROUND

Each State Employment Security Agency (except the Virgin Islands) is required to conduct a benefits quality control (BQC) program. Several state agencies have reported problems with this program indicating that it uses scarce staff resources to identify the same types of payment errors--quarter after quarter, year after year--and that it measures only one aspect of the quality of the unemployment insurance program, i.e., benefit payment accuracy. These agencies also believe it fails to improve the quality of the program. The states maintain that little attention is paid to correcting errors identified through the BQC program.

Under the BQC program, states are required to perform an in-depth audit on a sample of at least 400 paid claims annually. The Department of Labor (DOL) determines the actual sample based on a state's size and needs for precision. The audit involves interviewing the individual claimants and verifying base period wages, job search contacts during the week in question and all other aspects of their eligibility. Done almost entirely in person through June 1993, each audit averaged about 13 hours. A new auditing methodology, introduced in July 1993, combines telephone, mail, and facsimile communication with in-person contact. These procedures were designed to reduce the audit to approximately 10.5 hours. As data are verified, they are entered into a computer provided through BQC for analysis by the state, and then transmitted to DOL.

These audits are used to estimate each state's proper payment and payment error rates and to identify the party responsible and the reason for improper payments. The largest dollar errors identified are due to claimants' working while claiming benefits, their failure to meet work search requirements, and their being separated from their jobs for reasons that should disqualify them from receiving benefits. Because of their unique unemployment insurance laws, states have the primary responsibility for analyzing BQC data and using it to improve program operations.

Each June, states must report to the public the findings for the preceding calendar year, which DOL then releases in a compendium in July. Data release is intended to get interest groups involved in unemployment insurance improvement at the state level. Some states contend that DOL does make some use of the reports, but DOL has poorly explained how the information gathered applies to effective management.

A review of several state employment security offices provided differing opinions on the value of BQC data. Some office directors report that the data are necessary because of the potential for overpayments in a program distributing more than \$33 billion in funds annually.

Many states take the position that something must be done because the data are not being used in a manner that justifies the cost. Recommendations from some states range from reducing the number of cases reviewed to abolishing the program altogether. Others call for a combination of actions, including an explanation of how the data are improving the quality of the program.

DOL is aware of state dissatisfaction with the BQC program. The department agrees that, despite efforts to promote analysis and improvement actions through small funding grants, information sharing, training, and provisions for collecting more state-specific information, the anticipated level of BQC-inspired state analysis and corrective actions has not been forthcoming. As originally planned, some BQC resources are now being shifted from BQC activities. About one-tenth of the resources potentially available for BQC will be used to examine the quality of unemployment insurance tax operations through the newly developed Revenue Quality Control (RQC) program. However, DOL contends that BQC is indispensable, because it measures total payment accuracy in a way that is considered to be statistically valid. This measure is essential for state management and federal oversight. The Office of Management and Budget (OMB) considers the unemployment insurance accuracy rate to be an essential measure of the Unemployment Insurance program's accomplishments, which is required by the Chief Financial Officers Act.

ACTION

The DOL should reexamine the present mix of systems for improving the performance of the unemployment insurance program, and devise a unified strategy that improves its effectiveness.

This reexamination will include a substantial focus on the BQC program and how its resources can best be divided between measurement, analysis, and direct support to program improvement. In addition, DOL should consider the extent to which the BQC performance measurement will continue to keep its existing focus on paid claims in unemployment insurance core programs, or will include measurement of decisions to grant or deny claims or activities designed to strengthen the quality of the unemployment insurance program.

IMPLICATIONS

The \$34 million currently spent on the unemployment insurance quality control program now goes mostly to measure the accuracy of paid claims (\$2 million funds RQC, and roughly \$1 million is diverted to state-designed studies to craft program improvements). Some of the resources could be employed more directly in drawing implications from the performance measures for program improvement, or used to improve benefit payments processes or to measure other areas to complement existing measures, e.g., denied claims.

The net result is intended to be a shift in focus from error measurement to constructive use of the results to improve benefit payment quality and more effectively achieve the unemployment insurance program's goals.

FISCAL IMPACT

States were allocated approximately \$34 million to operate BQC programs in fiscal year 1993, which is about 1.7 percent of the \$2 billion states receive to administer the unemployment insurance program. These funds can be used more efficiently if DOL, in consultation with the state agencies operating the program, can identify the combination of activities that will lead to program improvements. The net result is expected to be a shift of resources from performance measurement activities to data analysis and process improvement. The magnitude of the shift cannot be projected in advance of the policy review. Thus, the fiscal implications cannot be estimated.

Note: This is the text of National Performance Review issue paper DOL21, published in From Red Tape to Results: Creating a Government That Works Better & Costs Less, Accompanying Report of the National Performance Review, Department of Labor. Washington, D.C.: Office of the Vice President, September 1993, pp 87-89. It has been retyped for electronic transmittal.