

MICROCOPY RESOLUTION TEST CHART
NBS 1067 A

GAO

Report to the Chairmen, Subcommittee on Labor, Health and Human Services, and Education, Committee on Appropriations, U.S. Senate

1

October 1987

SOCIAL SECURITY

Payment Accuracy Rates Are Overstated

AD-A187 963



DTIC
SELECTED
NOV 19 1987
S D

DISTRIBUTION STATEMENT A
Approved for public release
Distribution unlimited



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

Human Resources Division

B-229146

October 29, 1987

The Honorable Lawton Chiles
Chairman, Subcommittee on Labor,
Health and Human Services, and
Education
Committee on Appropriations
United States Senate

The Honorable William H. Natcher
Chairman, Subcommittee on Labor,
Health and Human Services, and
Education
Committee on Appropriations
House of Representatives

This report discusses the accuracy of the Social Security Administration payments to more than 33 million retirees, survivors, and dependents. The report also describes how the agency monitors its payment accuracy and reports to the Congress on its performance.

Our review was made pursuant to the Budget and Accounting Act of 1921 (31 U.S.C. 712) and the Accounting and Auditing Act of 1950 (31 U.S.C. 3511).

We are sending copies of this report to the Secretary of Health and Human Services; the Commissioner of Social Security; the Director, Office of Management and Budget; and chairmen of congressional committees and subcommittees with oversight responsibility for the Social Security Administration.

Richard L. Fogel
Assistant Comptroller General

Executive Summary

Purpose

Over 33 million beneficiaries—retirees, survivors, and dependents—received about \$175 billion in Social Security benefits during fiscal year 1986. These beneficiaries rely on the Social Security Administration (SSA) to calculate and pay the correct benefits. SSA measures how accurately it calculates and pays benefits by reviewing an annual sample of payment cases. Because accuracy rates are used as barometers of how well SSA carries out its fiduciary responsibilities to beneficiaries and the trust funds, GAO sought to determine the adequacy of SSA's measurement procedures and the accuracy of the rates reported annually for the Retirement and Survivors Insurance (RSI) program, SSA's largest program. GAO also measured the effect of errors on beneficiaries.

Background

SSA's 10 regional assessment offices annually review a sample of about 2,000 payment cases to determine accuracy for the RSI program. This review is an imperfect barometer in that it is not designed to identify with certainty the many specific causes of errors. Other SSA samples are designed to analyze causes of errors and monitor changes in processing accuracy, but cannot be used to project overall RSI payment accuracy because they do not include certain cases. SSA uses the results of its annual payment accuracy sample to compare overall trends in RSI program accuracy over time.

In 1985, SSA began reporting an RSI payment accuracy rate to the House Appropriations Committee, which expressed concern about the impact staff reductions might have on payment accuracy. SSA reported that its payment accuracy rate, expressed as a percentage of total dollars paid, fluctuated between 99.5 and 99.6 percent for fiscal years 1981-86.

GAO evaluated SSA's processes for measuring payment accuracy, using the findings of SSA's quality review analysts for 1981-86 to verify the reported accuracy rates. GAO also statistically compared the performances of the SSA regional assessment offices to determine whether variations in the error rate detected might be due to something other than chance. Finally, GAO used SSA's 1984 payment sample to calculate the effect of errors on beneficiaries.

Results in Brief

SSA's sampling methodology and processes for measuring errors and determining annual payment accuracy rates are adequate. However, not all errors detected are included when accuracy rates are calculated because of how SSA interprets errors. Consequently, GAO believes actual error rates are about twice what SSA calculates. When SSA does not

include all errors in calculating and reporting accuracy rates, it overstates program performance and can mislead managers about the extent to which potential improvements might be realized.

When SSA reported its annual payment accuracy rates for 1981-86 to the House Appropriations Committee, it reported an accuracy rate based only on total dollars paid, not on the incidence of cases in error, which was less favorable. Further, SSA's reported accuracy rate did not include underpayments and certain overpayments, which SSA believed should not be reported as errors, nor did the rate differentiate between errors caused by SSA and those caused by others.

When GAO counted all detected errors for 1986, the case accuracy rate was about 87.3 percent (compared with SSA's 94.1 percent), which means that about 4.2 million persons were overpaid or underpaid benefits totalling \$1.1 billion. Recalculating error rates for 1981-86 to include all errors SSA detected did not change the overall downward trend in errors.

SSA payment accuracy rates measure overall RSI program accuracy, but do not measure how errors affect beneficiaries. When GAO calculated the total impact of all errors on individual cases over time, it found almost two-thirds of errors were underpayments; because of the nature of the errors, most would not have been likely to be detected by routine SSA processes.

Principal Findings

Incidence of Error Not Reported

For 1981-86, SSA reported an average 99.5 percent payment accuracy rate based on total program dollars and not the average 92.1 percent case accuracy rate based on SSA criteria. This gave an overly favorable impression of payment accuracy. Although both rates are appropriate measures calculated by SSA, the case accuracy rate is a better indicator of the incidence of error; the dollar accuracy rate is a better indicator of the overall impact of errors on program finances. SSA said that outside SSA, it reported the accuracy rate based only on total dollars paid because, in the past, it had reported this figure to those inquiring about overall payment accuracy. (See p. 13.)

Some Errors Are Not Reported

In reporting dollar accuracy rates, SSA understates the extent of errors because its criteria for errors do not include (1) underpayment errors, (2) errors of \$1 to \$5 per month, and (3) errors barred from correction because they occurred over 4 years ago. SSA has consistently excluded these categories from its overall accuracy rates: the first, because it does not represent dollars incorrectly paid, and SSA calculates a separate underpayment accuracy rate but does not report this outside SSA; the second, because the amount of the error is relatively insignificant; and the third, because payments that SSA can no longer correct are, according to administrative procedures, no longer incorrect payments. To fully disclose all errors occurring, GAO believes all errors detected should be included when calculating and reporting accuracy rates. (See pp. 12-14.)

Exclusion of errors detected significantly affects the number of cases reported in error. When GAO included all detected errors, the case error rate averaged about 15.0 percent or about twice what SSA calculated using its criteria. The case error rate steadily decreased from 19.7 percent in 1982 to 11.5 percent in 1985; this was primarily due to fewer cases with recent earnings being underpaid because SSA computed their benefit increases more quickly. The error rate increased in 1986 to 12.7 percent, but such increase could be due to the chance inherent in random sampling. Over the 6-year period, an average of about 4.0 million cases, representing an average of about 4.7 million persons, were affected by errors each month.

Although SSA determines whether errors it detects were caused by it or others, SSA has not reported this outside SSA. For 1981-86, the incidence of cases reviewed found to be in error and attributed to SSA varied from 14.4 percent in 1982 to 6.2 percent in 1985. It was 7.0 percent in 1986, a statistically insignificant increase. Reporting the errors caused by SSA would provide better comparisons of SSA's performance with earlier years. (See pp. 14-21.)

Some Errors May Be Undetected

Because two regional assessment offices may not be detecting all errors in sampled cases, RSI errors may be higher than indicated in this report. GAO combined 5 years of sample data and compared and analyzed the error detection rates of the 10 regional assessment offices. GAO found a disparity in detection rates for the Chicago and Denver regional offices that could not be attributed to chance. This disparity was not found by the SSA headquarters group, which samples and validates the regional assessments but has not done a disparity analysis. From the available

data, GAO could not estimate the extent of any undetected errors or their effect on error rates. (See p. 22.)

Effect of Errors on Beneficiaries—Another Perspective

To measure the effect of underpayments and overpayments on beneficiaries, GAO used SSA's January 1984 sample. GAO found one of every six cases in error for an average of 5 years. About 60 percent of the error cases were underpaid, on average \$591 (median \$194). The remaining 40 percent were overpaid, on average \$1,069 (median \$251). About 10 percent of the error cases were judged likely to be corrected by one of several routine processes that verify or recalculate benefits. (See p. 27.)

Matter for Consideration of Congressional Committees

If data about the average and median amounts, as well as the duration of overpayments and underpayments, are desired, the Senate or House Appropriations Committees may want to so advise SSA.

Recommendations

GAO recommends that the Secretary of Health and Human Services (HHS) direct the Commissioner of Social Security to make a more complete disclosure of RSI payment errors. Specifically, the case error rate and the dollar error rate should be reported. In addition, the reported rates should (1) include all cases with detected errors and (2) indicate the incidence of errors attributed to SSA. SSA should use such error rates to better assess program performance and to decide the extent of improvement it deems attainable.

GAO also recommends that the Secretary direct the Commissioner to improve monitoring of regional assessment reviews and determine the following: the extent to which future regional reviews are not detecting errors, any actions needed to ensure all errors are detected, and the effect on payment accuracy rates of any errors not detected.

Agency Comments

HHS generally agreed with the thrust of the report. HHS said it would (1) review overall procedures for reporting payment accuracy data and reexamine its definitions and practices, (2) report case as well as dollar accuracy in the future for both overpayments and underpayments, and (3) report the extent of error attributed to SSA. HHS also said it planned to determine the reasons for the disparities in regional detection rates.

Contents

Executive Summary		2
Chapter 1		8
Introduction	Many Factors Can Affect Benefits	8
	How SSA Measures Payment Accuracy	9
	Objectives, Scope, and Methodology	10
Chapter 2		12
RSI Payment Errors Higher Than Reported	SSA Does Not Report All Detected Errors	12
	Reporting Error Incidence and Causes Would Better Indicate SSA Performance	14
	Payment Accuracy Reviews May Not Be Detecting All Errors	22
	Conclusions	25
	Recommendations	25
	Agency Comments and Our Evaluation	26
Chapter 3		27
How Errors Affect Beneficiaries	Almost One in Six Cases Had Outstanding Underpayments or Overpayments	27
	Erroneous Benefits Continued for an Average of 5 Years	28
	Conclusions	29
	Matter for Consideration of Congressional Committees	30
Appendixes	Appendix I: Scope and Methodology	32
	Appendix II: Comments From the Department of Health and Human Services	35
Tables	Table 2.1: Causes of Payment Errors (Fiscal Years 1981-86)	16
	Table 2.2: Dollar Error Rates Over 6 Fiscal Years	20
	Table 2.3: SSA Regional Assessment Office Error Detection Rates by Sample Year	22
	Table 2.4: Comparison of Regional Assessment Offices' Likelihood of Detecting Payment Error	23
Figures	Figure 2.1: Case Error Rates (Fiscal Years 1981-86)	17
	Figure 2.2: Cases in Error (Fiscal Years 1981-86)	19
	Figure 2.3: Dollars in Error (Fiscal Years 1981-86)	21

Contents



Approved For	
NTIS CRA&	✓
DTIC TAB	
Unannounced	
Justification	
By	
Date	
Authority	
A-1	

Abbreviations

- GAO General Accounting Office
- HHS Health and Human Services
- PSC program service center
- RSI Retirement and Survivors Insurance
- SSA Social Security Administration

Introduction

Social Security beneficiaries—retirees, survivors, and dependents—rely on the Social Security Administration (SSA) to accurately calculate their benefits. To measure the accuracy of its payments, SSA samples payments made and calculates an accuracy rate, which is reported to the Congress and others as indicative of how well SSA does its job.

The importance of SSA's payment accuracy has increased as staff cuts and other efficiency initiatives are considered and implemented. In 1985, the House Committee on Appropriations directed the Commissioner of Social Security to report to the Committee on the accuracy of payments, along with other performance indicators.¹ The Committee stated that these data would be particularly important in the future because of planned changes in SSA's staffing, field structure, data processing, and telecommunications.

This report (1) assesses the validity of SSA-reported accuracy rates for the Retirement and Survivors Insurance (RSI) program, SSA's largest program, and (2) measures how inaccurate payments affect beneficiaries.

Many Factors Can Affect Benefits

Measured by the number of beneficiaries and the dollar amount of payments, RSI is the largest program that SSA administers. In fiscal year 1986, SSA paid over 33 million beneficiaries about \$175 billion in benefits. RSI benefits are paid to partially replace earnings lost when workers retire or die. SSA must consider many factors related to the worker when determining eligibility for and the amount of the initial benefit, for example: amount of lifetime earnings, time period worked, date of birth, family relationship to others (such as spouse and children), entitlement to other benefits, and military service. After the benefit payments begin, events—postretirement earnings, marriage, death, and the birth of or changes in the age of children—could affect subsequent benefits. SSA cannot totally control the accuracy of initial or subsequent RSI benefits because it depends on beneficiaries and others, such as employers, to report accurate and timely information. Nevertheless, precise payment accuracy measurements are important because they can disclose trends in accuracy over time and can help identify the types of payment errors occurring.

¹For fiscal year 1986, SSA provided quarterly reports on service data directly to the Congress. For fiscal year 1987 and at the request of the House and Senate Committees on Appropriations, we are evaluating service data compiled by SSA and providing reports to the Committees on the effects of staff reductions on service.

How SSA Measures Payment Accuracy

The Office of Assessment within SSA reviews, evaluates, and measures the accuracy of Social Security programs, including the RSI program, and makes recommendations for corrective action. For the RSI program, the Office of Assessment uses three routine studies that address different program quality aspects: First, the Awards and Disallowances Study, which samples actions weekly and reviews about 60,000 cases annually, measures compliance with policies and procedures for approving initial claims. Second, the Postadjudicative Study, which also samples actions weekly and reviews about 30,000 cases annually, measures the accuracy of actions taken after a person becomes eligible for benefits. And third, the Payment Accuracy Study, which annually samples and reviews about 2,000 cases, estimates the overall accuracy of Social Security payments (benefit payments to people residing outside the country are not sampled as part of this study).

The Payment Accuracy Study sample includes both cases with actions during the year (such as initial claims and post-adjudicative actions) and cases for which no actions have been taken. It is, therefore, the only study from which an overall annual accuracy rate can be projected. Consequently, its value is primarily that of an overall accuracy indicator. This study also collects information about the causes of errors, but the sample size is too small for statistically valid conclusions to be drawn. More detailed data on the causes of payment errors are derived from the other two studies.

The Payment Accuracy Study currently samples January benefit payments, and the accuracy of that month's check is reviewed. A sampled check may include benefit payments for more than one beneficiary because multiple persons may be entitled to benefits under a single Social Security account number. The address of the sampled beneficiary determines which of 10 regional assessment offices will be assigned the case for review. At a regional assessment office, one or more quality review analysts will determine the correctness of the case by reviewing the claims folder, maintained and submitted by one of six program service centers (PSC). If required, an analyst will interview the beneficiary and others, such as employers, and gather corroborating evidence from different sources, such as state Bureaus of Vital Statistics.

For each case, an analyst then codes the findings, such as the dollar amount of the payment error. SSA has set criteria for determining an error during a case review and for determining one for reporting purposes. For instance, the criterion for an error during a case review is the following: the sampled monthly benefit amount differs from what it

would have been had the case been handled correctly from the beginning. Detected instances of these differences are coded and entered into the data base. The Office of Assessment then applies different criteria (for example, the dollar amount of the error must be \$5 or more) to determine the errors in the data base that are "reportable" (see p.12).

Other findings the analyst codes are the cause (139 specific types) and the responsible party (SSA, beneficiary, or other); the analyst enters these findings by computer terminal into the payment accuracy data base maintained at SSA headquarters. These findings may be re-reviewed through a paperwork evaluation at SSA headquarters, where staff use a subsample to validate the findings of the regional case reviews. From the data base, the Office of Assessment produces payment accuracy statistics that are expressed as dollar and case error rates. The dollar error rates are projected to estimate the amount of dollars underpaid or overpaid.

The Office of Assessment uses the statistics generated from the data base to prepare reports on RSI payment accuracy. These reports are to be used by SSA managers to provide an indicator of how well SSA is attaining overall quality objectives.

Objectives, Scope, and Methodology

Our objectives were to determine (1) whether SSA accurately measures and reports RSI payment error rates and (2) the total amount and duration of payment errors. We focused on the error rates SSA reported to the House Appropriations Committee since that was the most current information provided to the Congress that specified RSI error rates.

To determine if payment accuracy rates that SSA reported represented the errors SSA had found, we obtained SSA's payment accuracy data base for fiscal years 1981-86, generated statistics using SSA's criteria for reportable error, and compared the statistics with those found in SSA's reports to the Committee and in reports within SSA (internal Office of Assessment reports to the Commissioner and other SSA managers). Where we identified instances of errors that SSA had detected and recorded in the data base but did not enter into the calculation of its accuracy rate, we recalculated the payment accuracy statistics.

To determine if SSA quality review analysts might be missing some errors, thereby understating the incidence of error, we analyzed 5 fiscal years (1981-85) of payment accuracy rates. Using a statistical model, we compared all possible pairs of reviewing regions to see if errors were

detected by each region at the same rate. This comparison was possible because each region reviews some cases from every one of the six RSCs. Differences in rates of error detection were then tested for statistical significance to evaluate whether the differences were in fact due to chance rather than to some other factor, such as the adeptness of the analysts.

To measure how errors affect beneficiaries, we reviewed 2,242 cases (in SSA's fiscal year 1984 payment accuracy sample) in detail and compiled cumulative statistics about the errors in each case, for example: total amount of payment errors, total benefits paid, total number of months paid erroneously, and total number of months paid. The 1984 payment sample was chosen because it was the latest sample completed when we began our review.

In this report, we present information from our review and analysis of SSA records (particularly its payment accuracy data base), interviews with SSA officials, statistical projections of error rates, and technical analysis of regional differences in detecting errors. In May 1987, we briefed the staff of the House Appropriations Committee about our review of SSA's payment accuracy. Subsequently, the Committee included language on this matter in its 1988 appropriations report (no. 100-256). The Committee advised SSA to disclose information on its incidence of errors and those attributable to SSA, as well as the percentage of dollars paid erroneously.

We did our work at SSA headquarters in Baltimore; a regional assessment office and satellite office in Philadelphia; and the Mid-Atlantic and Great Lakes RSCs. To understand the three routine RSI studies (mentioned earlier) used to monitor program quality, we obtained information on the quality assurance process. We also gathered and evaluated more detailed information concerning these studies. Additional discussion of our scope and methodology can be found in appendix I.

We did not review computer controls relating to the generation of statistics from the data base. This was unnecessary because, as part of our detailed work, we reviewed the data in 1984 case files and compared that data, on a case by case basis, with the information in the data base; using SSA estimating procedures, we were able to compute the statistics reported at SSA (internally) and outside SSA (externally). We did our work from March 1985 through October 1986 in accordance with generally accepted government auditing standards.

RSI Payment Errors Higher Than Reported

SSA's measurement system for the RSI program is adequate in that it can provide accurate and reliable estimates on the amount of error in the program. This measurement system uses the annual Payment Accuracy Study (mentioned earlier) to draw a sample, following a specified sampling methodology, and to provide basic data. These data are projected to the universe by specified statistical techniques. Although the measurement system is technically sound, we believe the data could be better presented to (1) more fully disclose the amount of error in the RSI program and (2) enable managers to better gauge the extent to which potential improvements might be realized.

During fiscal year 1986, SSA reported to the House Appropriations Committee an RSI accuracy rate based only on dollars paid, even though SSA's Office of Assessment had also measured the RSI program with a rate based on cases paid. SSA said it reported an accuracy rate based only on dollars paid because that was what it had reported in the past to those inquiring about this matter. An RSI rate based on dollars paid generates a high accuracy rate and, conversely, a low-error statistic because the program payment errors that do occur in monthly payments tend to be small relative to the total benefit. The case accuracy rate is, therefore, a better indicator of the incidence of error. Both rates are valuable for describing the RSI program and should be used to fully disclose the extensiveness of payment errors.

Dollar errors detected by SSA analysts are about twice what SSA's reports to the Committee indicated. SSA counted certain overpaid dollars while excluding several categories of detected errors because it did not consider these as reportable errors. In addition, SSA did not count underpaid dollars as errors even though they were detailed in internal reports.

For fiscal year 1986, all RSI payment errors detected by SSA analysts indicated that over \$1.1 billion was underpaid or overpaid during the year and about 4.2 million persons were affected by payment errors each month. RSI payment errors may be even greater because analysts in some regional assessment offices may be better at detecting payment errors than analysts in other regional assessment offices.

SSA Does Not Report All Detected Errors

SSA reported to the Committee that the RSI accuracy rate ranged between 99.5 and 99.6 percent from fiscal years 1981 through 1986. This figure includes only benefit dollars associated with certain overpayments

occurring in the sample month.¹ It does not include all detected errors recorded in the data base, namely net underpayments, errors below \$5 per month, and errors barred from correction.²

Although SSA only reported certain overpaid dollars to the Committee, internal SSA reports specify dollars attributable to both overpayments and underpayments. SSA does not consider underpayments to be erroneous payments because underpayments involve instances where a payment was not made. SSA said it did not report underpayments because historically it had reported only overpayments. In a benefit program such as RSI, we believe that underpayments should be reported as errors.

An Office of Assessment official stated that the minimum amount for reporting a monthly benefit error, set at \$5, was based on the quality assurance criterion used in the Aid to Families With Dependent Children program. This \$5 amount represents a negotiated criterion agreed to by the federal government and the states for determining state fiscal liability for incorrectly paid benefits. SSA decided to adopt this criterion for the Payment Accuracy Study, even though determining fiscal liability is irrelevant to this study.

A further justification given by SSA for using a \$5 exclusion criterion was that inclusion of errors of less than \$5 per month would not significantly affect the error rate. We found SSA's position to be correct about the dollar error rate (see p. 20); however, inclusion of errors ranging from \$1 to \$5 a month affected the case error rate by an average 30 percent (see p. 17); therefore, in our opinion, these lesser amounts should be included. Further, SSA recognizes that these are payment errors even though not reported; they must, therefore, be corrected unless they are barred from correction. We did not consider benefit computation errors below \$1.00 to be payment errors because SSA's rounding procedures for benefit payments require rounding computations downward, not upward, to the nearest dollar.

¹In this report, we categorize payments to nonentitled individuals and excessive payments to entitled individuals as "overpayments". SSA categorizes such payments as "excess payments" because, in its view, nonentitled individuals are erroneously paid but not overpaid.

²An SSA administrative procedure bars errors from being corrected if the decision on which the payment was based occurred more than 4 years earlier. This procedure, termed "administrative finality," is intended to allow individuals to rely on SSA decisions. Such a determination may be reopened and changed for specified reasons such as fraud or clerical error unfavorable to the claimant. Most errors barred from correction are overpayments.

The Payment Accuracy Study procedures state that if a payment case is barred from correction by an administrative procedure, it is to be reported internally as a "non-dollar impact deficiency." SSA believes that reporting cases barred from correction as errors is inappropriate because the cases are deemed "correct" by an administrative procedure. SSA also believes that to report such cases as payment errors would contradict this procedure. We did not evaluate the merits of the administrative procedure, which also affects SSA programs that we did not review. However, in our opinion, cases barred from correction should be reported as payment errors because the payments for these cases are different from the payments that would have been made had the cases been handled correctly from the beginning.

SSA's including all errors in calculating and reporting payment accuracy would better indicate the RSI program's overall performance; it would enable SSA to better gauge the extent to which potential improvements might be realized. An overstated high accuracy rate could lead managers to conclude that further improvements would be marginally cost-effective and therefore should not be explored. Conversely, management may have more incentive to improve an accuracy rate of 85 percent, for example, than one of 92 percent. Fully disclosing the extent of all errors occurring would enable managers to better decide whether the causes should be more closely examined so as to reduce future incidence. Likewise, fully disclosing all errors to the Congress should enable it to better carry out its oversight responsibilities.

Reporting Error Incidence and Causes Would Better Indicate SSA Performance

SSA reported dollar error rates to the Committee and in internal Office of Assessment reports. However, the internal reports also included the related case error rates and data on the extent of error attributed to SSA and others; such rates and data, which were not reported to the Committee, provide a measure of the extent to which beneficiaries are affected by errors and the extent to which SSA is responsible. These reports are not routinely distributed outside SSA. According to an SSA official, neither the Secretary of Health and Human Services nor the Congress is sent the reports because SSA believes neither desires such detail. We were told that internal reports have been distributed outside SSA when requested or when SSA felt the reports answered a specific inquiry.

Case accuracy rates are a better indicator of SSA's ability to pay benefits accurately because these rates show how many people are being correctly or incorrectly paid, whereas the dollar accuracy rates are a measure of the size of dollar errors in relation to total dollars paid. Although

the latter are an appropriate measure, they mask the frequency of error. For example, if half of all beneficiaries were underpaid \$10 per month for 1 year, the dollar accuracy rate would be correctly calculated as 99 percent (based on an average payment of \$500) and the case accuracy rate as 50 percent. The dollar accuracy rates show that the trust funds are not significantly affected; the case accuracy rates, that many beneficiaries are. Both rates are valuable for program oversight.

Case error rates for fiscal years 1981-86 doubled from an average of 7.9 percent (as reported in internal SSA reports) to 14.9 percent, if all errors identified by SSA analysts and recorded in the data base are counted. Similarly, dollar error rates doubled from an average of about 1/2 of 1 percent, as reported to the Committee, to almost 1 percent (if underpayments, as well as certain other overpayments not now reported, are considered erroneous). Including all errors identified when calculating accuracy rates would not incur additional cost because all errors detected, regardless of whether they are reported as errors, are recorded in the data base to provide a record of what was found.

SSA's Payment Accuracy Study is not intended to provide data on the causes of case errors, and the size of the annual sample does not permit projecting specific causes with statistical certainty; however, for internal reporting purposes, SSA does categorize the causes it identifies. This categorization distinguishes between errors caused by SSA and others. Generally, SSA causes most underpayments, ranging from a low of 72.1 percent in fiscal year 1986 to a high of 96.1 percent in fiscal year 1982; others, such as beneficiaries, cause most overpayments, ranging from a low of 50.0 percent in fiscal year 1986 to a high of 74.7 percent in 1983. Reporting and comparing the incidence of error caused by SSA each year would more precisely indicate how well SSA fulfills its responsibility to accurately pay benefits.

Taking the errors detected by SSA in the annual sample, we calculated who was responsible for payment errors. Overall, most errors were attributed to SSA; the case error rate decreased between fiscal years 1982-85, showing a slight but statistically insignificant increase in 1986 (see table 2.1).

Table 2.1: Causes of Payment Errors
(Fiscal Years 1981-86)^a

Cause of error	Fiscal years					
	1981	1982	1983	1984	1985	1986
SSA	11.5	14.4	10.0	8.1	6.2	7.0
Beneficiary or other	2.0	2.1	2.9	3.7	2.6	3.0

^aExcludes errors made in cases barred from correction by SSA's administrative procedures

Some errors could be categorized as being caused by either the beneficiary or SSA, depending on how one interprets the findings in a specific case. For example, a beneficiary who was overpaid because he did not report earnings in excess of the maximum allowed alleged he had not been told that such earnings had to be reported. SSA says it advises all beneficiaries of this requirement at the time of application and sends periodic subsequent reminders; in addition, this requirement is stated in the information provided to applicants and beneficiaries. SSA attributed the error in the example above to the beneficiary. In reviewing SSA's basis for categorizing the causes of errors, we found that its determinations, generally, were reasonable and based on the information in the file.

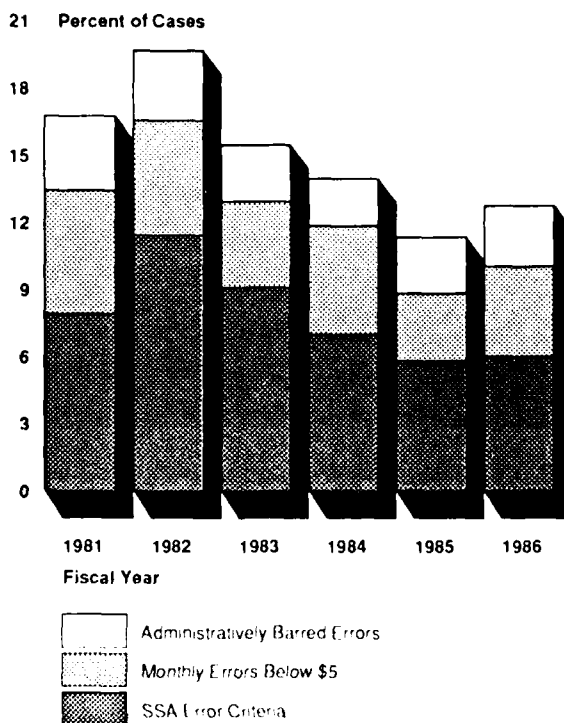
Portraying the incidence of error, as shown in table 2.1, gives a clearer indication of SSA's performance over time; however, because of the nature of the annual Payment Accuracy Study, incidence of error is not the best indicator for detecting recent changes in accuracy. The study measures cases being paid erroneously as of the month of the study; therefore, it primarily reflects errors made in earlier years. Consequently, the usefulness of the data in measuring how accurately SSA handled actions affecting benefits in 1986 versus 1985, for example, is limited. Better measures for that purpose are studies of process accuracy—the Awards and Disallowances Study and the Postadjudicative Study. Such measures, which are compiled as a 6-month rolling average and reported quarterly, would be better indicators of the effect staff reductions could be having on RSI program accuracy. Both the annual payment and quarterly process accuracy rates are included in our three reports on Social Security.³

³Social Security: Staff Reductions and Service Quality (GAO HRD-87-66, Mar. 10, 1987; GAO HRD-87-97BR, June 2, 1987; and GAO HRD-87-139BR, Sep. 17, 1987)

RSI Cases in Error

SSA's detected and reported case error rates over 6 years are shown in figure 2.1. Case error rates, reported internally, would increase if cases with a monthly dollar error below \$5 and cases with errors barred from correction by an administrative procedure were included. For instance, in fiscal year 1986, the case error rate increased from a reported 5.9 to 12.7 percent when all detected and recorded errors were counted. Each case represents one or more beneficiaries paid under one Social Security account number. The categories represented by the bar graphs in figure 2.1 include both overpaid and underpaid cases because SSA includes both in its internally reported statistics.

Figure 2.1: Case Error Rates (Fiscal Years 1981-86)



From fiscal years 1981 through 1986, SSA internally reported case error rates that averaged 7.9 percent. When errors below \$5 per month and errors barred from correction are added, the average error rate increases to 14.9 percent. The error rate decreased from 1982 to 1985 largely because of fewer cases being underpaid.

We tested the changes in case error rates for statistical significance to determine whether the changes could be due to chance, inherent in random sampling. We found that the increase from 1981 to 1982 and the decrease since 1982 were not due to chance, but we could not rule out that possibility for the small increase from 1985 to 1986.

The underpaid case error rate dropped from 14.8 percent in 1982 to 6.7 percent in 1985. SSA attributes the underpayment decrease to more timely processing of benefit increases from earnings after retirement.⁴ The error rate increased slightly in 1986 because the number of underpaid cases below \$5 per month went up. SSA was responsible for 72.1 percent of the underpaid cases in 1986, primarily because of computer processing problems. The percentage of underpayments caused by SSA has generally declined from a high of 96.1 percent in 1982.

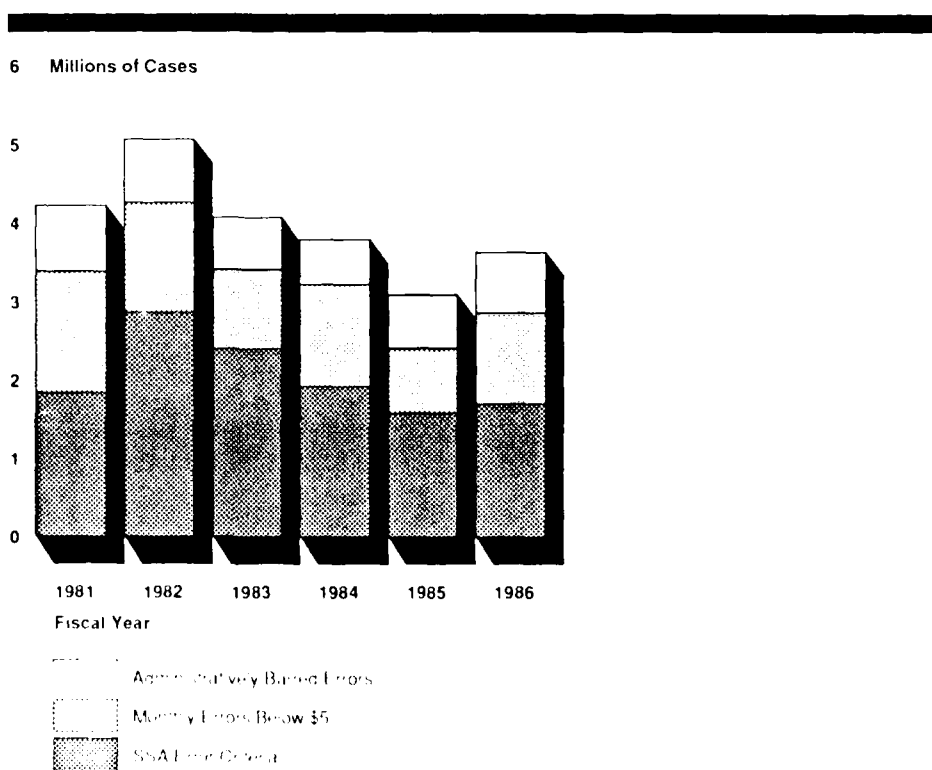
From 1981 through 1986, overpaid cases fluctuated between 4.7 and 5.6 percent of cases paid. Before 1986, beneficiaries had been responsible for the greater proportion of overpaid cases. This proportion peaked in 1983, when beneficiaries were responsible for 74.7 percent of the overpaid cases, but declined to 50.0 percent in 1986. Of the cases it analyzed and reported, SSA found that the reason for this decline was (1) that beneficiaries were more accurately reporting earnings to comply with the annual earnings test and (2) the phasing out of student benefits. SSA attributed the improved accuracy in earnings reporting to its sending reminders to beneficiaries about estimating earnings.⁵

⁴In order to know if RSI benefits should be increased for those entitled to them who earn wages or have self-employed income, the Social Security Act requires SSA to recompute their benefits. SSA's policy is to annually review benefit computations for beneficiaries with recent earnings. A partially automated process, Automatic Earnings Reappraisal Operation, is used to recompute benefit amounts. The time it took to complete this process increased from 16 months in 1978 to 42 months in 1980; after 1980 the time required has gradually been reduced. The SSA time criterion for a payment sample error is that the process should be completed no later than 21 months after the end of the earnings year. GAO's Delays in Recomputing Social Security Benefits Cause Underpayments for Extended Periods (GAO HRD-84-71, Sept. 13, 1984) discusses problems SSA experienced in this area during 1978-82.

⁵Social Security benefits are meant to replace, in part, earnings lost to a beneficiary or beneficiary because of retirement, disability, or death. Therefore, the amount of benefits that a retiree under age 70 and dependents will receive each year depends on whether the beneficiary is partially or fully retired. The earnings test (also known as the retirement test) is used to measure the degree of retirement and to determine the amount, if any, to be deducted from monthly benefits. Annual exempt amounts vary depending on whether the beneficiary is under age 65 or at least age 65. Amounts also vary from year to year to keep them in ratio with earnings levels. Beneficiaries are required to estimate earnings for the coming year and, if necessary, benefits are adjusted based on the estimate. Following the end of the earnings year, beneficiaries are required to file an annual report specifying the exact amount earned during the previous year and, if needed, a reconciliation of benefits received is made.

Using the case error rates just discussed, we estimated the number of cases paid in error each month in a given year, from 1981 to 1986. The estimated number of cases paid incorrectly over 6 years is shown in figure 2.2.

Figure 2.2: Cases in Error (Fiscal Years 1981-86)



Using SSA's criterion for reporting an error, an average of 2 million cases was erroneously paid each year over the 6-year period. However, when other categories of recorded errors are included, the average rises to almost 4 million cases per year. Using the average number of beneficiaries per RSI case, we estimate that the average number of persons affected by payment errors was about 4.7 million per year for the 6-year period. For 1986, an estimated 4.2 million persons were affected by payment errors.

RSI Benefit Dollars Paid in Error

The dollar error rate calculated by SSA, used to report payment accuracy to the Congress, averaged less than 1/2 of 1 percent over 6 fiscal years. We compared this rate with the dollar error rate we calculated that included additional SSA detected errors (table 2.2). Consideration of these additional error categories increases the average dollar error rate to almost 1 percent. A dollar error rate of 1 percent seems low, but currently this rate amounts to \$1.5 billion in erroneous benefits each year. Of the three types of errors SSA detected but did not include in its dollar error rate calculation, including underpayments as a dollar error increased the rate by the largest increment. Overall, from 1981 to 1983, the dollar error rate steadily increased because of increasing underpayments; the rate then decreased, although not constantly, with the lowest total dollar error rate for the 6 years in 1986.

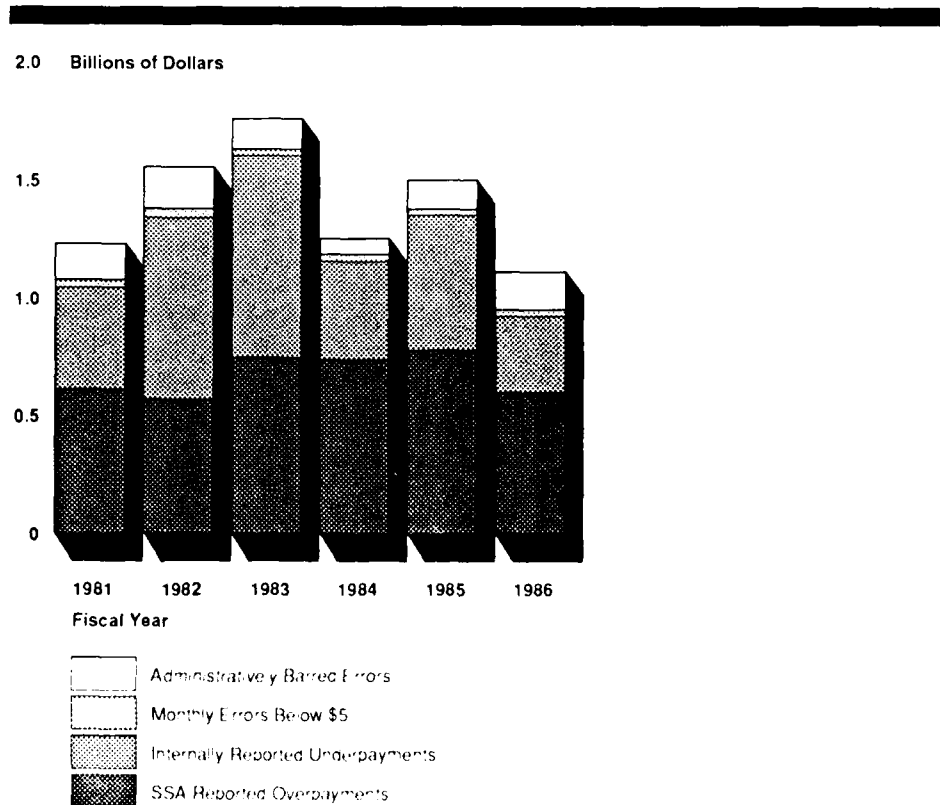
Table 2.2: Dollar Error Rates Over 6 Fiscal Years^a

	Fiscal years					
	1981	1982	1983	1984	1985	1986
In percent						
Recorded errors						
Calculated by SSA	0.52	0.43	0.52	0.49	0.48	0.35
Additional calculations by GAO						
Underpayments	0.37	0.58	0.59	0.27	0.36	0.19
Below \$5 per month	0.03	0.02	0.01	0.02	0.02	0.02
Barred	0.13	0.14	0.09	0.04	0.07	0.09
Total	1.05	1.17	1.23	0.83	0.96	0.66

^aWe calculated these error rates using SSA's data base of detected and recorded errors. The total error rates are weighted estimates, not the arithmetic sum of the individual components.

The effect of differing error-reporting criteria on projected error dollars is illustrated in figure 2.3. For example, in fiscal year 1986, the estimated dollars paid in error grew from \$602 million (\$0.6 billion), using the accuracy rate SSA reported to the Committee, to \$1.1 billion when all detected and recorded errors are included. In this figure, underpayments are illustrated separately because SSA did not include them in its report to the Committee. SSA's 1986 internal report shows underpaid dollars and overpaid dollars in a similar manner. As shown in figure 2.3, the addition of errors of less than \$5 per month does not materially affect the projected dollars in error, even though the addition of these errors does significantly increase the projected cases in error as was seen in figure 2.1.

Figure 2.3: Dollars in Error (Fiscal Years 1981-86)



According to SSA, it was responsible for 77.1 percent of the underpaid dollars in 1986. Incorrect computations have replaced late recomputations (as mentioned earlier, to increase benefits as a result of earnings after retirement) as the leading underpayment problem.

In 1986, SSA reported that beneficiaries were responsible for 53.7 percent of the overpaid dollars. Although the annual earnings test was still the highest single reason for overpayments (21.0 percent), this reason was at its lowest point since overpayments have been reported. Most overpaid dollars for 1986 resulted from problems that formerly had a much smaller dollar error rate, such as incorrect determinations of family relationships.

Payment Accuracy Reviews May Not Be Detecting All Errors

Quality review analysts determine payment accuracy rates by reviewing cases. We found that analysts in some regional assessment offices may be better at detecting errors than analysts in other regional assessment offices. Consequently, RSI errors may be higher than indicated in this report. We combined 5 years of payment accuracy results to compare the 10 regional assessment offices. Our statistical analysis showed significant differences, which could not be attributed to chance, in the numbers of errors detected by the various regions. These differences were not detected by the headquarters group, which samples and evaluates cases reviewed by the regional assessment offices. We applied an analysis methodology specifically designed to identify statistically significant differences in detection rates between pairs of regional offices. SSA did not make such an analysis.

Each regional office's error detection rates for the payment accuracy samples of 1981-85 are shown in table 2.3. For example, in 1981, the Boston Regional Office found 17.2 percent of the cases it reviewed to be in error. In the 1981-85 period, overall detection rates among the offices ranged from a low of 11.3 percent for Denver to a high of 19.6 percent for Philadelphia. Some of the differences in error detection rates are because the majority of sample cases assigned to a regional office for review come from one or two of the six PSCs; thus, these cases may not have the same frequency of errors as those cases assigned to another regional office, which receives cases from different PSCs. Stated another way, a regional office assigned cases with more errors should find more errors.

Table 2.3: SSA Regional Assessment Office Error Detection Rates by Sample Year

Regional assessment office	Sample year					5-year average
	1981	1982	1983	1984	1985	
Boston	17.2	21.5	20.7	17.4	16.7	18.6
New York	24.6	19.0	16.5	17.6	15.8	18.7
Philadelphia	19.1	25.6	14.7	17.4	17.0	18.8
Atlanta	15.6	19.7	18.1	14.7	16.7	17.0
Chicago	12.2	15.8	11.7	14.3	11.0	13.8
Dallas	18.9	23.6	11.2	14.1	15.8	16.7
Kansas City	24.7	19.7	16.1	17.2	14.0	18.4
Denver	11.1	14.9	11.5	12.7	11.0	12.2
San Francisco	13.8	16.7	16.1	16.8	11.0	15.7
Seattle	14.0	17.1	16.3	16.0	11.0	15.9
Overall average	16.6	19.5	15.3	13.9	11.3	16.1

The study results shown in table 2.3 do not confirm that one region is consistently better at detecting errors than another. Differences could be attributed to, as mentioned above, (1) some regions' being assigned more cases with errors than those assigned to other regions or (2) the normal and expected variation that arises from random sampling. Since we wanted to know whether some regions were better at detecting errors than others, we analyzed the 1981-85 Payment Accuracy Study data to determine if differences in error detection rates between regions were statistically significant.¹⁰ Were we to find any statistically significant differences, we would conclude that some factor, other than the sampling process, was affecting error detection. In our opinion, that factor would most likely be the ability of the regions to detect errors. Our statistical model controlled for other influences by mathematically removing the effect of some offices' receiving for review more cases with errors than others; we then evaluated the statistical significance of the remaining differences in error detection rates. Using the model, we compared the differences in rates between all possible pairs of regions and found comparisons with Chicago and Denver to be statistically significant. We found no other comparisons to be statistically significant. From the model, first, we conclude that Philadelphia, Kansas City, Dallas, New York, and Boston were more likely to detect an error than Chicago (table 2.4). For instance, Philadelphia would have an 85.3 percent better chance of detecting an error than Chicago; we are 90 percent confident that the true difference ranges between 82.3 percent and 88.4 percent.

Table 2.4: Comparison of Regional Assessment Offices' Likelihood of Detecting Payment Error

Regional assessment office	Greater probability of detecting payment error than Chicago	
	Estimate	Interval ^a
Philadelphia	85.3	82.3 to 88.4
Kansas City	84.8	76.7 to 91.6
Dallas	73.8	70.7 to 76.9
New York	68.7	59.3 to 78.7
Boston	51.8	20.6 to 91.0

^aIntervals are computed at the 90 percent level of statistical confidence.

¹⁰Differences were considered significant if they met the 90 percent level of statistical confidence—that is, if we drew 100 samples, we would expect to find error detection rate differences falling within the confidence interval in 90 of these samples.

Second, we found statistically significant differences in error detection between the Philadelphia and Denver offices. We estimated that Philadelphia had a 72.0 percent better chance of detecting an error than Denver if it reviewed the same case, and we are 90 percent confident that the true difference lies between 48.8 and 98.7 percent.

Although our analysis suggests that some reviewers are not detecting some errors, SSA headquarters' re-reviews of a sample of cases reviewed by regional assessment offices generally have not found that reviewers missed errors. SSA headquarters uses a subsample of the RSI sample cases to validate the results of RSI payment accuracy case reviews done in regional offices. After a case is selected for re-review by headquarters personnel, the claims folder and the quality assurance review folder (with evidence gathered by regional quality review analysts) are sent to SSA headquarters where personnel review the regional findings and the coding of these findings in the data base. Headquarters personnel have only paperwork available for their review, whereas the regional quality review analysts had the opportunity to interview each beneficiary and others, such as employers, and to gather additional evidence. If relevant evidence was not gathered during the regional review, the subsample reviewer is limited in the conclusions to be drawn.

For the period we examined (1981-85), subsample reports show that most error cases and a sample of nonerror cases were re-reviewed at SSA headquarters. The number of nonerror cases re-reviewed was reduced during this period because SSA believed that (1) review accuracy had increased and (2) the amount of other work done by the headquarters review group had increased. Recent cases re-reviewed have included most error cases and about 10 percent of nonerror cases; but, generally, the re-review has not identified any errors missed by regional reviewers.

SSA could consider whether it would be cost-effective to have cases from each office re-reviewed by an experienced reviewer in another region. If such cross-checks were done without the benefit of the re-reviewer's knowing the prior determination, the cross-check could provide an independent, technically sound verification. Current re-reviews by headquarters personnel are made with knowledge of the reviewer's determination; these re-reviews tend to emphasize ensuring that the reviewer's coding of findings is consistent with procedures rather than cross-checking.

Conclusions

SSA's payment accuracy for RSI benefits is not as good as portrayed by reports to the House Appropriations Committee. For the RSI program, SSA reported a payment accuracy rate no less than 99.5 percent for 1981-86. Our analysis of RSI payment accuracy indicates, however, that the reported accuracy rates overstate performance because SSA does not

- count all of the errors it detects in yearly studies, excluding errors involving (1) underpayments, (2) amounts below \$5 per month, and (3) cases precluded from correction because of administrative procedures and
- report accuracy rates on the basis of cases paid, which (including the above three categories) results in an average of 85-percent payment accuracy for 1981-86.

Not including all errors detected when calculating accuracy rates overstates program performance and can mislead managers and the Congress about the extent to which potential improvements might be realized. In addition, SSA needs to determine what additional monitoring or verification might be needed to ensure quality review analysts are detecting all errors.

A fundamental goal of SSA is to make accurate RSI benefit payments, protecting the interests of beneficiaries and the integrity of the trust funds. As such, the payment accuracy rate is an important indicator of management effectiveness and program design. Since payment accuracy can be measured in a variety of ways that portray wide differences in performance, SSA should provide as full a disclosure of its performance as possible to assist both management and the Congress in providing program oversight.

Recommendations

The Secretary of Health and Human Services should direct the Commissioner of Social Security to report all categories of detected RSI program errors. Such reporting should include

- case accuracy rates as well as dollar accuracy rates,
- all errors detected during the sample period, and
- the incidence of errors attributed to SSA.

The Secretary should also direct the Commissioner of Social Security to determine

- the extent to which regional assessment offices may not be identifying errors during the Payment Accuracy Study and the effect such undetected errors have on the payment accuracy rate and
- what additional verification or changes may be warranted in the monitoring of regional quality review analysts.

Agency Comments and Our Evaluation

In a letter dated September 25, 1987, the HHS Inspector General said HHS was in basic agreement with our report. (See app. II.) He added that overall procedures on reporting payment accuracy will be reviewed and, in the future, SSA will report on case as well as dollar accuracy.

Concerning our recommendation to report all categories of detected RSI program errors, the Inspector General said SSA's disclosure of payment accuracy data, both internal and external, could be more uniform and comprehensive. He noted that the House Appropriations Committee made specific mention of the need to disclose information on the incidence and source of errors in its report (no. 100-256) on 1988 appropriations. (See p. 11 for a discussion of the Committee's report.) The Inspector General said that historical definitions and practices about payment error data will be reexamined to determine specific changes that should be made. He said the reports on payment accuracy covering the period ending December 31, 1987, will reflect these changes.

Concerning our recommendation to examine how well regional assessment offices detect payment errors during the annual Payment Accuracy Study, the Inspector General said that the disparity in error rates we identified at two regions was not necessarily due to performance deficiencies. According to the Inspector General, other factors, such as regional variations in error rates or, possibly, demographic and economic causes, might be responsible for lower detection rates. Nevertheless, the Inspector General said that there is merit in the thrust of our finding and HHS will pursue the reasons for this disparity. He anticipates preliminary results before 1988.

We believe HHS's plans are responsive to our recommendations.

How Errors Affect Beneficiaries

Knowing the size and duration of payment errors to beneficiaries can supplement payment accuracy data. During our review, we expanded SSA's criteria for error, including errors outstanding before SSA's sample month. We did this because it is possible, as we noticed during our review, that beneficiaries can be correctly paid in SSA's sample month, yet have an uncorrected payment error from periods before the sample month. If an error was outstanding, we calculated the net benefit effect for each beneficiary as of the sample month (this effect included all uncorrected errors for those beneficiaries since the case began); we then calculated the sum of these payment errors in terms of monthly payments to beneficiaries. SSA does not routinely compile accumulated payment errors for the cases it samples.

Using an expanded view of uncorrected errors, we found that 16.8 percent (4.5 million cases) had an uncorrected error as of January 1984 and that the cases had been in error for an average of about 5 years. Underpaid cases averaged \$591.50, which was worth about 1.6 months of 1984 monthly benefit payments for those cases, and overpaid cases averaged \$1,069.50, which was worth about 2.4 months of 1984 benefits for these cases. When compared with total benefits paid in each case, error amounts averaged about 5 percent. Most of the error cases would not have been detected or corrected by routine SSA processes. Even if an error had been detected by SSA, apart from the sample of 2,242 cases, 14.6 percent of the error cases could not be fully corrected because they were either totally or partially barred from correction by an SSA administrative procedure.

Almost One in Six Cases Had Outstanding Underpayments or Overpayments

Of the almost one in six cases in error (16.8 percent), some cases had all underpayments or all overpayments and some cases had both. We computed a net error in each case and found that slightly more than 6 of 10 error cases had a total net underpayment. Dollar errors in cases with net underpayments ranged from \$1 per case to \$6,285.80,¹ with an average of \$591.50 per case; the median underpayment was \$194.55. When compared with the total payments made in each case, underpaid dollars ranged from less than 1 percent to 243 percent (the beneficiary received about one-third of the amount entitled), with an average of 4.4 percent. In comparison, dollar errors in cases with total net overpayments

¹The \$6,285.80 underpayment occurred because SSA did not include earnings after retirement in the monthly benefit computation. The beneficiary in this case had been underpaid for the past 7 years.

ranged from less than \$1 per case to \$21,367.20,² with an average of \$1,069.50 per case; the median overpayment was \$251.40. In terms of total payments in each case, overpaid dollars ranged from less than 1 percent to 100 percent (the beneficiary was not entitled to benefits received), with an average of 5.4 percent. The 16.8 percent error rate included cases with relatively small total errors.³ We computed the average without these small errors and found the error rate decreased to 13.1 percent.

SSA will detect and correct some errors by routine processes. For example, beneficiaries who continue to work after retirement may be eligible for an increase in their monthly benefits. SSA procedures require that this increase be made within 21 months after the end of the earnings year. Cases that exceeded this 21-month criterion were considered erroneous by SSA and us, even though a future routine process would eventually correct the problem. Of the cases in error as of the sample date, 10.8 percent were either corrected by the time of SSA's case review or, in our opinion, were likely to be corrected in the future by one of several routine SSA processes, such as the Automatic Earnings Reappraisal Operation, intended to review and update records and recalculate benefits.

As compared with cases that could be corrected by routine processes, an administrative procedure barred either total or partial correction of 14.6 percent of the error cases, as mentioned earlier. About three-fourths of these cases had one or more errors that were totally barred from correction. The other one-fourth had multiple errors, of which at least one was barred from correction. Cases with net overpayments were more than twice as likely to be partially or totally barred from correction than those with net underpayments. Generally, when the underpayment is not caused by the beneficiary, it can be corrected even after 4 years.

Erroneous Benefits Continued for an Average of 5 Years

The number of months a beneficiary received erroneous benefits, as of January 1984, ranged from a minimum of 1 month to a maximum of 257 months, with an average of 61.3 months and a median of 40 months. On the average, cases in error received the wrong monthly benefit amount for 70 percent of the months for which a benefit was paid or was due. Without small errors, the average was 66 percent.

²The \$21,367.20 overpayment occurred because SSA found that the beneficiary should have had benefits withheld because of earnings. This beneficiary had been overpaid for the past 2 1/2 years.

³We defined small errors as a case with below a \$50 total accumulated error and below \$5 a month average error. Cases had to meet both criteria to be considered small.

Errors Worth About 2 Months of 1984 Benefits

We compared the accumulated error totals with the 1984 benefit payment to get another perspective on the impact of errors on beneficiaries. To do so, we first had to express the accumulated errors in 1984 dollars because inflation had increased the value of earlier error dollars. We used the cost-of-living increases granted to Social Security beneficiaries to increase all error dollars to a 1984 base. We then measured the economic worth of the accumulated errors as of the 1984 sample date. Further details on this constant dollar technique can be found in appendix I.

Application of this constant dollar technique showed that the average net underpayment per case of \$591.50 was worth \$827.00 in 1984 economic purchasing power. We divided the escalated error totals by the 1984 benefit payment for each case and averaged the results. Using this technique, the average net underpayment was worth about 1.6 months of 1984 benefits for these beneficiaries. The range varied on an individual case basis from less than 1 month to almost 17 months worth of benefits. For a typical elderly couple aged 65 and over in 1984, this amount of economic purchasing power represented about 6 percent of total yearly expenditures.

Our constant dollar analysis showed that the average net overpayment per case of \$1,069.50 was worth \$1,342.00 in 1984 purchasing power (about 2.4 months of 1984 benefits). The range was again wide for individual cases, from a low of less than 1 month to a high of almost 35 months. For a typical elderly couple in 1984, the additional economic purchasing power was worth about 10 percent of total yearly expenditures.

Conclusions

Most uncorrected errors are small relative to the monthly benefit. These errors are not of the type likely to be detected by routine processes that review payments, update records, and recalculate benefits. Consequently, these errors are likely to continue for a long time and the accumulated error amount can be substantial. The total lifetime impact of errors on beneficiaries affected by errors not detected by routine processes or special studies, such as the annual Payment Accuracy Study, will be larger than the averages presented in this chapter. This is so because of the following: Had the cases not been included in the sample, the payment errors for most cases would have continued indefinitely and the amounts overpaid or underpaid would have accumulated further.

SSA does not currently compile data from its Payment Accuracy Study to disclose the size and duration of payment errors to beneficiaries. Periodically collecting and reporting the effect of errors on beneficiaries would provide an additional payment accuracy indicator, disclosing the extent to which errors are affecting beneficiaries' payments.

**Matter for
Consideration of
Congressional
Committees**

If data about the average and median amounts as well as the duration of overpayments and underpayments are desired, either on a periodic or as requested basis, the Senate or House Appropriations Committees may want to so advise SSA.

Scope and Methodology

We wanted to (1) determine if the sample design and methodology of the Payment Accuracy Study could accurately measure errors, (2) compare the errors SSA finds with what it reports, (3) determine whether all errors were being detected, (4) recalculate error rates to reflect any detected errors not included in SSA's calculation, and (5) measure the dollar effect of errors on beneficiaries.

Adequacy of Sample Design and Methodology

From a statistical standpoint, we evaluated the sample size and the procedures used for sample selection, case review, and estimating the extent of error; we also reviewed the application of the prescribed procedures.

Verification of Detected and Reported Errors

To verify SSA externally reported error rates, we obtained published SSA reports from fiscal year 1981 through 1986 and SSA reports to the House Appropriations Committee for fiscal 1986; we then compared data from these sources with the study's supporting data base. We limited our verification of the data base to a comparison of the coding on most 1984 cases and the findings recorded in the quality assurance review folder. Although a few differences were detected, the instances were not numerous enough, in our opinion, to affect the validity of the data base. We then used the data base and SSA's criteria for reportable error to recalculate the error statistics reported internally and externally by SSA for fiscal years 1981-86.

To check the accuracy of the quality assurance reviews done by the regional assessment offices, we considered the reasonableness of the determinations and verified the accuracy of the benefit calculation for the 2,242 cases sampled in 1984; generally, we used the data in the review file. We did not attempt to independently verify the findings through a check of the claims folders in the PSCs.

To check if errors might be going undetected during regional office reviews, we analyzed 5 years (1981-85) of payment accuracy data. Since the 10 regional assessment offices reviewed cases from each of the 6 PSCs, we constructed a matrix showing (1) the number of cases reviewed and (2) the number of error cases found for each PSC by each reviewing regional office. As mentioned earlier, some regional offices detected more errors than others because they were assigned more error cases for review. In order to evaluate whether the differences in error detection rates could be attributed solely to chance, rather than some other reason

(such as analysts in some regions being better at detecting errors than analysts in other regions), we used the logistic (LOGIT) statistical model. We selected the LOGIT model because of its ability to deal with certain characteristics of the sample data, which we needed to consider before evaluating the statistical significance of detection rate differences. Specifically, the model (1) permits use of a dichotomous dependent variable (an error is either detected and valued at 1 or not detected and valued at 0), (2) permits use of nominal independent variables (regions are annotated 1 through 10 and PSCs are annotated 1 through 6), (3) allows PSC comparisons even with small numbers of sample cases in some regions, and (4) controls for interaction effects, that is, regions' receiving different numbers of cases with errors from different PSCs.

Generation of Alternative Error Rates

To generate alternative accuracy statistics, including the errors SSA deemed "not reportable," we used the data base to produce case and dollar error rates that included all detected and recorded errors for each of the 6 years. This procedure was possible because SSA quality review analysts coded all errors noted in sample cases regardless of whether or not the errors were considered reportable. We used statistical techniques similar to those used by SSA and projected the alternative error statistics to estimate underpaid and overpaid dollars and numbers of error cases.

Impact of Errors on Beneficiaries

To evaluate the impact of accumulated errors on beneficiaries, we obtained all available quality assurance review folders for the 1984 Payment Accuracy Study sample from the 10 regional assessment offices responsible for carrying out the review. This year was chosen because the data were the most current available when our review was started. Although the January payment was sampled, not until the following fall were all cases completed, findings entered into the data base, and the data base verified. The 1984 Payment Accuracy Study sample, 2,242 RSI cases from across the country, was reviewed by analysts in the region in which the sampled beneficiary resided.

SSA considers an error to exist only when the sample month payment is incorrect. For the purpose of measuring the impact of errors on beneficiaries, we included all cases where the beneficiary was either due money or owed money because of an error, regardless of whether the sample month payment was correct or incorrect. As a measure of the impact of errors on beneficiaries, the amount of the error for a specific case was calculated and expressed as a total of all money due the beneficiary or due the government. We calculated other cumulative statistics.

Appendix I
Scope and Methodology

such as total benefits paid, months paid, and months incorrectly paid; we also made mathematical comparisons, such as the percentage of monthly payments in error. To compare the impact of accumulated error dollars with the 1984 monthly benefit payment, we calculated a new total dollar error (using cost-of-living increases granted to Social Security beneficiaries), expressed in constant 1984 dollars. To calculate the number of benefit months represented by the error amount, we divided the new total dollar error by the amount of the January 1984 monthly benefit amount. This constant dollar technique was necessary because the value of error dollars over the years was affected by inflation.

Comments From the Department of Health and Human Services



DEPARTMENT OF HEALTH & HUMAN SERVICES

Office of Inspector General

SEP 25 1987

Mr. Richard L. Fogel
Assistant Comptroller General
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Fogel:

The Secretary asked that I respond to your request for the Department's comments on your draft report, "Social Security: Payment Accuracy Rates Are Overstated." The enclosed comments represent the tentative position of the Department and are subject to reevaluation when the final version of this report is received.

We appreciate the opportunity to comment on this draft report before its publication.

Sincerely yours,

A handwritten signature in dark ink, appearing to read "Richard P. Kusserow".

Richard P. Kusserow
Inspector General

Enclosure

Appendix II
Comments From the Department of Health
and Human Services

COMMENTS OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES ON THE
GENERAL ACCOUNTING OFFICE DRAFT REPORT, "SOCIAL SECURITY
PAYMENT ACCURACY RATES ARE OVERSTATED"

General Comments

This report presents the General Accounting Office's (GAO) analysis of the Social Security Administration's (SSA) measurement and reporting of payment error rates in the retirement and survivors insurance (RSI) program. It also presents data developed by the auditors on the amount and duration of erroneous payments to Social Security beneficiaries. In essence, the report concludes that SSA's sampling methodology and procedures for measuring errors and determining annual payment accuracy rates are adequate. In fact, the auditors used the review data produced by SSA as the basis of all of their findings. They did not see the need to perform an additional audit of the quality assurance review process. However, the report notes that the data could be better presented to fully disclose the amount of error in the RSI program, and to allow for better program oversight by Congress and SSA. Specifically, the report faults SSA for not reporting case error rates, for understating the extent of errors by not also reporting on underpayments, errors from \$1 to \$5 per month, and errors barred to correction, and for not reporting whether errors SSA detects are caused by it or others. The report also has findings related to the performance of SSA's regional assessment offices in detecting errors.

We are in basic agreement with the thrust of the report that more comprehensive disclosure of data is both desirable and feasible. We plan to review our overall procedure in the area of reporting payment accuracy data, and we will in the future report on case as well as dollar accuracy for both overpayments and underpayments.

GAO Recommendation

The Secretary of Health and Human Services should direct the Commissioner of Social Security to report all categories of detected RSI program errors. Such reporting should include:

- case accuracy rates as well as dollar accuracy rates,
- all errors detected during the sample period, and
- the incidence of errors attributed to SSA.

Appendix II
Comments From the Department of Health
and Human Services

Department Comments

We agree that SSA's disclosure of payment accuracy data, both internally and externally, could be more uniform and comprehensive. The House Appropriations Committee in its report (number 100-256) makes specific mention of the payment accuracy data that SSA provides, and indicates that the data reported to Congress should "disclose the incidence of errors and the extent of error attributed to SSA as well as the percentage of dollars paid erroneously." As we take action to comply with the Committee's instructions, we plan to reexamine our historical definitions and practices regarding error data with a view towards determining what specific changes should be made and the correlating effects on trust fund accounting, program oversight and resource management. We will comply with the Committee's instructions beginning with the data for the period ending December 31, 1987.

GAO Recommendation

The Secretary should also direct the Commissioner of Social Security to determine:

- the extent to which regional assessment offices may not be identifying errors during the payment accuracy study and the effect such undetected errors have on the payment accuracy rate, and
- what additional verification or changes in the monitoring of regional quality review analysts' determinations may be warranted.

Department Comments

This recommendation stems from GAO's efforts to determine whether some regional assessment offices perform better than others in detecting errors. Using a statistical model, GAO compared the differences in error detection rates between all possible pairs of SSA's 10 regional assessment staffs and found statistically significant differences in the error detection rates for the Chicago and Denver staffs. These two regions showed consistently lower error rates. GAO does not believe that the lower error rates are due to sampling variability, chance, or differences in actual error rates. (According to the report, GAO's "statistical model controlled for other influences by mathematically removing the effect of some offices receiving more cases with errors for review than others.") GAO concludes that this finding suggests that these two review offices are not finding errors, resulting in

Appendix II
Comments From the Department of Health
and Human Services

national accuracy rates that are higher than they should be. We do not agree with this conclusion that the disparity in error detection rates is necessarily due to performance deficiencies in these two regions. In our view, the variances could be due to other factors, such as regional variations in error rates. Additionally, we do not believe that we can rule out demographic and economic causes for a lower error rate in these regions.

These concerns notwithstanding, we find merit in the thrust of GAO's finding and will pursue efforts to determine the reasons for these apparent disparities in regional detection rates. We anticipate preliminary results by the end of calendar year 1987.

END

DATE

FILMED

FEB.

1988