United States General Accounting Office 131839

GAO

Report to the Chairman, Subcommittee on Intergovernmental Relations and Human Resources, Committee on Government Operations
House of Representatives

November 1986

COMPUTER MATCHING

Assessing Its Costs and Benefits

RESTRICTED——Not to be released outside the General Accounting Office except on the basis of specific approval by the Office of Congressional Relations.

RELEASED

GAO/PEMD-87-2

537323

GAO

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

Program Evaluation and
Methodology Division

B-221025

November 10, 1986

The Honorable Ted Weiss
Chairman, Subcommittee on Intergovernmental
Relations and Human Resources
Committee on Government Operations
House of Representatives

Dear Mr. Chairman:

This report responds to your March 15, 1984, letter asking us to develop a methodology for evaluating the costs and benefits of a computer match, both prospectively and retrospectively. In agreement with your office, we developed guidelines rather than a formal methodology because of the immaturity of research in this field, because of the considerable diversity of views that exist, and because of the informality that still characterizes the varied methods being practiced. The report synthesizes current approaches to computer-match cost-benefit analysis. We have integrated information on agency practices with a review of the relevant computer-match cost-benefit literature and interviews with experts in the subject. We discuss areas in which the performance of computer-match cost-benefit analyses could be improved.

At your request, we did not seek comments on the report from the agencies in which we conducted interviews. Unless you publicly announce the contents of this report earlier, we will make no further distribution of it for 30 days. Thirty days after the date of the report, copies will be available to those who request them.

If you have any questions or would like additional technical information, please call me (202-275-1854) or Dr. Lois-ellin Datta (202-275-1370).

Sincerely,

Eleanor Chelimsky
Director

Executive Summary

Purpose

The chairman of the House Subcommittee on Intergovernmental Relations and Human Resources asked GAO to study the methods used for assessing the costs and benefits of computer-matching projects. This report synthesizes current approaches to computer-match cost-benefit assessment based on the existing literature, a review of agency practices, and interviews with experts.

Background

Concern over fraud, waste, abuse, and error in government programs has stimulated the development of computer matching to detect and prevent these problems. Defined as the identification of similarities or dissimilarities in the data of two or more computer files, computer matching has often been used to identify individuals who may not be in compliance with program rules and regulations. Proponents of computer matching consider it an effective tool for improving the integrity of payment programs. It has been applied in support of audits and to strengthen internal program controls. Matching has identified inaccurate program data and millions of dollars of estimated savings from actual or projected overpayments.

The use of computer matching has raised several concerns. One is whether particular matches actually achieve the cost-savings or cost-avoidance anticipated. Some critics contend that when all the monetary costs of performing a match are included and all the monetary benefits are accurately measured, the costs may be greater than the benefits.

Government agencies are not now formally required to analyze computer-matching costs and benefits. However, the Office of Management and Budget has issued a computer-match checklist asking agencies for, among other things, cost-benefit information about the matches they perform. The Long-Term Computer-Matching Project of the President's Council on Integrity and Efficiency has produced materials that advocate conducting computer-match cost-benefit analysis and provide some general guidance on how these analyses should be performed.

Results in Brief

In this report, GAO identifies, in detail, the costs and benefits to be considered in computer-match cost-benefit analyses. (See page 25, table 3-1.)

GAO notes that methodological problems place the measurement of certain types of costs and benefits beyond the resources of routine analysis. However, as computer-match operations grow in frequency, scope, and

sensitivity, measurement development should be undertaken, and these costs and benefits should be considered and assessed in special studies of selected match operations. GAO's report does not address whether specific computer matches have been cost-beneficial or not or which nonmethodological conditions or circumstances may promote or preclude the performance of cost-benefit analyses.

GAO's Analysis

Computer-match cost-benefit analyses can be performed prospectively, assessing whether a match should be initiated, or concurrently, to decide whether to continue a match, or retrospectively, to determine whether a match was successful. Accuracy in measuring the costs and benefits of a computer match depends on identifying all the activities associated with its performance, from its initial phase of planning and development to the completion of verification and follow-up activities on the individuals or organizations identified as positive matches. (See pages 26-30.)

Since cost-benefit analysis involves a comparison of total costs with total benefits, all costs and benefits must be recognized and, if possible, measured. To the extent that it is feasible, they should be assigned a monetary value, and then they should be aggregated and compared. It is very important to consider the more qualitative costs and benefits in an analysis, even when it is not possible to assign them monetary values. (See pages 25-26.)

Entities

In developing a synthesis of approaches to computer-match cost-benefit assessment, GAO takes a broad perspective appropriate to the government, whereby the costs and benefits to all the entities that may be affected or involved in a match are addressed, rather than just costs and benefits to the matching agency.

GAO identified six key entities to be included in cost-benefit analyses: the matching agency, the source agencies, the justice system, the agencies' clients, third parties, and the general public. (See pages 24-25.)

Costs

The major costs to matching and source agencies are the salary and fringe benefits of personnel involved in all phases of the match process. Costs to the justice system include the time and resources of personnel involved with any investigation and prosecution activities. (See pages 32-43.)

The costs to program clients and third parties can include the time and resources they spend responding to agency requests for verification information. Theoretically, the costs to the general public include the potential encroachment on constitutional and legal rights and privileges. It is often not feasible to measure these costs in monetary terms or to integrate them into a cost-benefit analysis, but their likely magnitude should be considered along with the more quantifiable costs and benefits. (See pages 47-53.)

Benefits

The major benefits to the matching and source agencies are the recovery of overpayments and debts, the avoidance of future overpayments, improvements in program operations, deterrence, and law enforcement. The latter two are also benefits to the justice system. (See pages 55-68.)

Measuring the recovery of overpayments and debts requires tracking systems that can monitor repayment activity. Actual, not merely projected or potential, repayments should be reported in postmatch analyses. Estimating overpayments that are to be avoided requires determining the length of time a program's clients would continue to be overpaid if not detected by a match. The benefits from deterrence, law enforcement, and improvements in program management may be quite difficult to quantify and should be the subject of special studies.

The benefits that a program's clients may gain include improvements in the delivery of services and the correction of underpayments. The benefits to the general public, those who pay for programs, are related to improvements in program efficiency. While it is conceptually useful to distinguish between benefits to an agency and benefits to the public, their measurement is the same, and care should be taken to count them only once in an analysis. (See pages 69-71.)

Analyses Agencies Performed

GAO reviewed 17 computer-match operations in 9 agencies (see chapter 6) and found considerable variation in the timing and manner in which the cost-benefit analyses were performed. Overall, GAO's review indicated that when agencies performed an analysis, they did not provide a full accounting of costs and benefits. In some of the agencies' analyses, not all the costs and benefits to matching agencies that could reasonably be measured or estimated were included. Further, none of the analyses of matches in which the recovery of overpayment or debts extended over a lengthy period of time used discounting in evaluating them. (See pages 72-79.)

For the match operations GAO reviewed, the information available was not adequate to support a sound decision about whether specific matches were or might be monetarily cost-beneficial. If computer matches are to be justified on the grounds of cost-avoidance or cost-savings, it will be necessary to have adequate information about all significant match costs and benefits. The performance of more rigorous cost-benefit analyses could more firmly establish and monitor the magnitude of the benefits obtained for the resources expended.

Recommendations

GAO does not present recommendations in this report.

Agency Comments

At the request of the House Subcommittee on Intergovernmental Relations and Human Resources, GAO did not ask for official agency comments.

Contents

Abbreviations

AFDC	Aid to Families With Dependent Children
ARIMA	Autoregressive integrated moving average
DOD	U.S. Department of Defense
DOL	U.S. Department of Labor
GAO	U.S. General Accounting Office
HHS	U.S. Department of Health and Human Services
HUD	U.S. Department of Housing and Urban Development
IRS	Internal Revenue Service
OMB	Office of Management and Budget
OPM	Office of Personnel Management
SSA	Social Security Administration
USDA	U.S. Department of Agriculture
VA	Veterans Administration

Introduction

Increasing concern over fraud, waste, abuse, and error in government programs has stimulated the development of techniques that use information technology to detect and prevent these problems. The techniques include error-prone profiling, computer screening, front-end matching, and computer matching. "Computer matching," as we use the term in this document, refers to the identification of similarities or dissimilarities in data found in two or more computer files. One reason often used for making such identifications is to determine whether individuals are inappropriately receiving public benefits.

Computer matching is one response to an increasing awareness among program managers, inspectors general, and legislators of the potential for fraud and abuse in government programs. The government use of computer matching has increased considerably over the last decade. We have reported on the benefits of matching as well as on the need to balance the potential benefits of computer matches with the protection of individual privacy. [1]

The benefits addressed in our reports and testimony include the identification and cessation of erroneous payments, the improvement of eligibility verification procedures, and the correction of program deficiencies. Other potential benefits include the collection of overpayments and debts, the deterrence of fraud and abuse, and increased public support for programs.

Controversy about computer matching has centered on the legal and constitutional issues of unreasonable search and the right to privacy and due process. Critics of computer matching have also charged, however, that some matches may not be cost-beneficial. They argue that when all the costs of performing a match are included and the benefits are realistically determined, the costs of the match may be greater than the money to be saved or recovered. Proper cost-benefit analyses can be useful in determining whether matches are indeed cost-beneficial.

The value of cost-benefit analyses has been recognized but has not been made a consistent requirement in the performance of computer matches.

[1] See Brian P. Crowley, Statement Before the Subcommittee on Domestic Marketing, Consumer Relations, and Nutrition of the House Committee on Agriculture, "General Accounting Office Review of Efforts to Reduce Food Stamp Program Losses," U.S. General Accounting Office, Washington, D.C., April 20, 1983, and U.S. General Accounting Office, GAO Observations on the Use of Tax Return Information for Verification in Entitlement Programs, GAO/HRD-84-72 (Washington, D.C.: June 5, 1984); Eligibility Verification and Privacy in Federal Benefit Programs: A Delicate Balance, GAO/HRD-85-22 (Washington, D.C.: March 1, 1985); and A Central Wage File for Use by Federal Agencies: Benefits and Concerns, GAO/HRD-85-31 (Washington, D.C.: May 21, 1985).

In 1979, the Office of Management and Budget (OMB) issued "Privacy Act of 1974; Supplemental Guidance for Matching Programs."[2] The guidelines were designed to help agencies relate the procedural requirements of the Privacy Act of 1974 with computer-matching programs that involved the disclosure of personal records subject to the act. In addition to their extensive reporting requirements, the guidelines required the performance of a cost-benefit analysis prior to conducting a match.

Because inspectors general and others who perform matches subject to OMB's guidelines argued that the guidelines were overly burdensome and unrealistic, OMB revised the guidelines in 1982. OMB eliminated the cost-benefit analysis as a prerequisite to a match and streamlined the reporting requirements, reasoning that it was appropriate for agencies to conduct cost-benefit analyses of matches but that the analyses should be viewed as only one of several components of decision-making. In 1983, OMB developed a computer-match checklist for agencies initiating matches subject to the Privacy Act. The primary purpose of the checklist was to ensure compliance with the procedural regulations of the act. It included an item requesting an estimate of the likely costs and benefits of a match. The checklist did not specify the costs or benefits the analysis should include or how it should be performed.

These controversies and uncertainties have increased the concern of some members of the Congress about the practice of computer matching. In 1982, the Senate Subcommittee on Oversight of Government Management held hearings on the use of computer matching. In 1984, the same subcommittee heard testimony on the disclosure of Internal Revenue Service (IRS) data to states to support computer matching. In March 1984, the chair of the Subcommittee on Intergovernmental Relations and Human Resources of the House Committee on Government Operations asked us to develop a methodology for evaluating the costs and benefits of computer matching, both prospectively and retrospectively, and to identify and evaluate the criteria that inspectors general have used in deciding whether a particular computer match should be conducted (the request letter is in appendix I).

[2]Office of Management and Budget, "Privacy Act of 1974; Supplemental Guidance for Matching Programs," 44 Fed. Reg. 23,138 (1979).

Objective, Scope, and Methodology

Our objective in this report was to respond to the first part of the congressional request; we respond to the second part in a separate report. To meet this objective, we first performed a broad, preliminary review of federal agency computer-matching activities. This review showed that the methods that have been used for assessing the costs and benefits of matching projects were not well developed, well described, or standardized. Agencies had spent some effort to assess the costs and benefits of computer-matching projects, and further development efforts for estimating costs and benefits were under way. However, the preliminary and fragmented nature of this work, and the diversity of match situations that would have to be addressed, indicated that an extensive and lengthy effort would be required if we were to develop a comprehensive methodology.

Given this situation, we thought that an appropriate first step toward the improvement of computer-match cost-benefit analyses would be an assessment of current methods and practices. Therefore, with the concurrence of the subcommittee's office, we sought to review the practices that have been followed, with the intent of using this information to develop a general framework for conducting or reviewing computer-match cost-benefit analyses.

The methodology we employed was the information synthesis, which integrates data from several sources that have been critically assessed. In particular, the synthesis enabled us to summarize current practices, provide useful guidance in assessing computer-match costs and benefits, identify shortcomings in current methodological practice, and suggest methodologically sounder techniques and procedures. Moreover, it provided a base from which to question or analyze claims about the costs and benefits of specific matches.

We began the synthesis with a review of the literature on computer-match cost-benefit analysis and the computer-matching process.[3] We also reviewed the general cost-benefit literature that appeared to be applicable to computer matching. From the information we obtained from this literature review, we developed a discussion package that identified the basic issues associated with computer-match cost-benefit analyses. We organized this material around the topics of cost-

[3]General discussions of both computer-match cost-benefit analyses and the computer-match process are in U.S. Department of Health and Human Services, Office of Inspector General, Computer Matching in State Administered Benefit Programs: A Manager's Guide to Decision Making (Washington, D.C.: 1983), and President's Council on Integrity and Efficiency, Long-Term Computer-Matching Project, Reference Paper on Computer Matching (Washington, D.C.: 1982).

estimating methods, benefit-estimating methods, crosscutting issues, and considerations for planning and performing a cost-benefit analysis. After identifying an advisory panel of experts from universities, private business, and government who were familiar with computer matching or cost-benefit analysis, we met with them to present the discussion package and obtain their insights into the various topics in our study.

Using their comments in conjunction with the literature we had reviewed, we developed some general, working criteria that could be used as a framework for assessing individual cost-benefit methods and analyses. (An example of the criteria is in appendix II.) We also developed an interview guide and checklist for interviewing agency officials on computer-match cost-benefit analyses and for obtaining additional relevant material on specific matches. The interview guide was designed to serve two purposes: to identify and query practitioners about their methods in performing cost-benefit analyses for specific, selected computer matches and to identify other matches that would provide us with more information on methodology. The checklist was an aid for reminding interviewers of the various cost-benefit topics to be covered.

We used the Inventory of Federal Computer Applications to Prevent/Detect Fraud, Waste, and Mismanagement, prepared by the U.S. Department of Labor (DOL), and OMB's matching reports in order to identify specific computer-matching projects to which to apply our interview guide. From these lists, we selected matches that indicated some reporting of costs or benefits. We did not use a rigorous sampling approach to select matches, because we could not for sampling purposes confidently delineate a population of matches for which some form of cost-benefit analysis had been performed. Descriptions of the matches in the Inventory were sketchy, and the set of OMB matching reports was not complete. We could not identify other comprehensive sources or compilations of agency matches. Moreover, since we were interested in identifying better practices as well as current practices, we wanted to be able to include matches agency officials had identified for their cost-benefit analyses.

We interviewed the match-contact person for each match and collected descriptive information on the match and any related cost-benefit analysis. We asked these persons to identify other matches they knew about that might have involved an assessment of match costs and benefits. Our goal was to identify and follow up on matches that were likely to provide insight into the processes of conducting a cost-benefit analysis of a match and interpreting its results. We spoke with officials from the

following agencies: U.S. Department of Agriculture (USDA), U.S. Department of Defense (DOD), U.S. Department of Education, U.S. Department of Health and Human Services (HHS), including the Social Security Administration (SSA), U.S. Department of Housing and Urban Development (HUD), Internal Revenue Service, U.S. Department of Labor, Office of Personnel Management (OPM), Selective Service Commission, and Veterans Administration (VA).

From the information we obtained from these agencies, we identified matches for which there appeared to be information available on costs and benefits for more detailed review, and we conducted further interviews about these matches with pertinent individuals at these agencies. Because computer matching usually involves different groups at different steps in the match process, it was sometimes necessary to interview two or three persons associated with a particular match. These interviews were focused on the questions that were still outstanding after our initial interviews. We obtained from these further interviews more detailed material on 17 computer-match operations.[4]

We examined this information on the costs and benefits of actual matches in relation to our working criteria and extracted pertinent methodological details. We integrated these methodological details with the previously developed material to prepare this report. In summary, the information we present in the remainder of this report represents a synthesis of the relevant computer-match literature, discussions with experts, and reviews of the cost and benefit measurement techniques used in a set of matches. At the request of the Subcommittee on Intergovernmental Relations and Human Resources, we did not obtain official agency comments on a draft of this report.

The Structure of This Report

Chapters 2 and 3 of this report provide the general information on computer matching and related cost-benefit issues that is needed for the more detailed discussion of measurement issues in the subsequent chapters. Chapter 2 is a general description of computer matching, and chapter 3 is a general introduction to cost-benefit analysis and the several issues related to how cost-benefit analyses of computer matches are

[4]Several of the 17 match operations encompassed more than one distinctly identifiable agency match. In some instances, the match operation involved performing the same type of match in several geographical locations. Other multiple match operations reflected a recurring series of matches or the performance of a pilot match followed by the implementation of a full-scale match. From the information and documents we obtained from agency officials, we identified 28 specific matches associated with the 17 match operations (they are described in appendix III).

conducted. Chapter 4 provides a detailed discussion of measurement considerations for estimating costs; chapter 5 does the same for benefits. Chapter 6 presents our review of the cost-benefit analyses of selected agency match operations. In chapter 7, we identify some criteria or objectives that should be considered when conducting and evaluating a cost-benefit analysis of computer matches.

Computer Matching

Computer matching is one of several computer-aided techniques that have been developed to screen, edit, and scan data to identify irregularities that may indicate fraud, abuse, and error. Because the term "computer matching" has sometimes been applied to encompass such closely related techniques as front-end matching, error-prone profiling, and computer screening, we make some distinctions between these terms.

"Front-end computer matching" refers to the comparison of information provided by a program applicant to information in one or more data bases before the applicant is enrolled in the program, or shortly thereafter, in order to verify the accuracy of the applicant's information, especially as it pertains to eligibility or entitlement. A variant of front-end matching is employed by law enforcement officials when they check individuals, vehicles, or property they suspect against centralized data bases containing information on criminals and stolen property.

"Error-prone profiling," or "computer profiling," is a process in which average data are examined in order to identify the most frequent sources of or occasions for error. The data on a particular applicant are matched or compared to the average profile. IRS uses error-prone profiling techniques to identify taxpayers' returns that have a high probability of being inaccurate. Returns that approximate the profiles are identified for further checking and possible audit.

"Computer screening" is closely related to profiling, in that it identifies abnormal patterns in one or more computer records, usually within the same file. Computer screens scan automated files to detect cases with unusual, unlikely, or inaccurate attributes that may indicate a problem. For example, a file may be screened for Social Security numbers that are known never to have been assigned or that are incongruent with other information in the file (as when two different individuals have the same number).

Computer Matching: A Definition

Our definition of computer matching does not include the techniques described above or the computer verification of information on one or a few individuals against a set of files. "Computer matching" as we use the term is restricted to comparisons of two or more files containing information on persons or organizations of interest to the government. The output of the type of computer match is a list of persons or organizations appearing to have violated program rules and regulations. They are referred to as "computer match hits." For example, in a welfare program match, a file containing wage information reported by employers

to a state may be searched for persons whose names are on the welfare rolls. The purpose of the match would be to detect welfare clients who have not accurately reported their income to the program and who are receiving inappropriate benefit payments.

The Purposes of Computer Matching

Computer matching has served a variety of purposes. It is most commonly viewed as an audit tool for assessing particular aspects of a program or an agency's operations. Program administrators and inspectors general have cited computer matching as a means of assessing the adequacy of program procedures and internal controls to ensure that only legitimately entitled individuals receive the correct level of benefits or services. One of the benefits expected from many computer matches is the identification and collection of overpayments, and another is grant adjustments and discontinuations to avoid future erroneous payments. Estimates of the magnitude of these benefits have been reported and range from $4 to $54 for each $1 spent on a match.

Because computer matching can detect persons or organizations that may be intentionally defrauding the government of benefits and services, it is also used by investigators as an enforcement mechanism, because it enables them to identify suspects for further investigation. Knowing that computer matches are being performed may deter some individuals from attempting to defraud or abuse a program for fear of being caught in the match.

Computer matching is also viewed as a way for program administrators to compensate for weak program controls, whether the controls are potentially correctable or inherently weak, by catching errors and inappropriate eligibility determinations that have already been made. Follow-up matches may serve the additional purpose of assessing the effectiveness of program changes implemented as a result of a first match. For example, when program rules and regulations concerning an eligibility determination are changed, computer matching may be used for reviewing compliance with the new regulations and identifying individuals who may not be in compliance with them.

Matches may be undertaken to address only one of the purposes discussed above, but in practice the results of a computer match may support several of these purposes to varying degrees. In addition to revealing problems in program operations, matches that are conducted as a type of program audit can support enforcement efforts by identifying instances that are strong indications of fraud and by referring

them to investigators. Deterrence would be enhanced, according to theory, if all potential fraud cases could be followed up systematically, if program participants were well aware that this was being done, and if program participants perceived that the consequences of being caught were sufficiently severe. Further, the quality and reliability of the data in the files may be improved by the correction of errors or outdated data. Generally, the managers of the match decide the specific use or uses of the match, although a variety of program and technical factors may limit the range of uses to which a specific match can be applied.

The Development of Computer Matching

"Project Match," initiated in 1977 by the U.S. Department of Health, Education, and Welfare, is commonly cited as the federal government's first major effort to combat fraud, waste, abuse, and error by computer matching. Project Match compared the records of roughly 78 percent of all recipients of Aid to Families with Dependent Children (AFDC) with the payroll records of about 3 million federal employees, in order to detect those who might be illegally drawing welfare payments. Prior to this effort, the few computer matches that had been federally conducted were small-scale audits using files within a single agency or state.

Subsequent to Project Match, inspectors general in a number of agencies, especially USDA, DOL, and HHS, adopted computer matching as a regular audit or management tool. It was not uncommon for the office of an inspector general to perform a one-time match as an audit project. If the final, refined hit list was fairly large, the office would verify and follow up on a statistically derived sample of hits to estimate the potential results of the match overall. The match list would then be given to program administrators for verification and a follow-up on all the hits. Depending on the types of problems found in the sample, the inspector general might recommend in the audit report that the matching operation be performed regularly by the program. With this approach, matching would become institutionalized within routine program operations.

In March 1981, the president established the Council on Integrity and Efficiency, composed of the inspectors general and other representatives of federal executive agencies. One of the council's major efforts is the interagency Long-Term Computer-Matching Project, which is intended to facilitate and improve the use of computer matching and related techniques in federal and state government as a broad management tool (for checking on internal program controls, for example) rather than exclusively for combating fraud and abuse. A newsletter is

being published under this project, and several surveys of federal and state computer matches have been completed. The project's current effort is to develop standardized file formats to facilitate the computer-match data exchanges. A field study of the proposed standard formats is being conducted in eight sites.

The surveys of computer matches conducted by the President's Council on Integrity and Efficiency indicate the growth of the use of computer matching. Early in 1982, federal agencies had performed about 85 matches; early in 1984, the Inventory of Federal Computer Applications to Prevent/Detect Fraud, Waste, and Mismanagement identified 186 matches. State use of computer matching is considerably greater. One estimate of the number of matches performed early in 1982 was 170; the Inventory of State Computer Match Technology, published in March 1983 by HHS, identified more than 500 state matches. The majority were routine and continuing matches comparing benefit program files (AFDC, Food Stamps, Medicaid, and the like) with state wage and unemployment compensation files.

The states have increased their use of computer matching, largely because of federal legislation mandating or encouraging its use. In 1977, the Congress enacted Public Law 95-216, requiring the states to match AFDC cases against state wage data. Later laws (Public Laws 96-249 and 97-98) required the states to disclose wage and unemployment compensation data to Food Stamp agencies for use in computer matching. It is likely that state use of computer matching will increase even further because of the Deficit Reduction Act of 1984. This act requires the states to develop an income-and-eligibility verification system to permit the association of records in AFDC, Food Stamp, Medicaid, and other related programs. Wage, asset, and other information from SSA and IRS are to be made available, and the states that do not have a quarterly wage-reporting system are required to implement one, so that they can provide this information. Standardized formats are to be developed for data exchange.

Steps in the Computer-Matching Process

A computer match can be divided into three major operational phases: match definition and development, match initiation and production, and hit verification and follow-up. The activities or steps associated with each phase can be identified, but not all the activities or steps will necessarily have to be performed for any given match, and their sequence can vary considerably.

In match definition and development, the general idea for the match is developed and its feasibility is assessed. If the match is considered viable, formal planning is begun and initial contact is made with other agencies whose participation is required. Planning becomes more detailed as a consequence of responses from agencies and the collection of information about the files to be matched. Requirements for compliance with the Privacy Act, OMB's matching guidelines, and individual agency rules are reviewed in order to identify necessary activities such as the publication of notices in the Federal Register.[1] If no major problems or obstacles challenge the viability of the match, the match may then be initiated.

In match initiation and production, written agreements with participating agencies are developed, notices (such as the Federal Register announcement) are published, and the data files to be matched are obtained. Activities during the initial stages of this phase include the development, testing, and documentation of software necessary to produce and verify the hits of the match. Some preliminary processing of the files may occur in order to verify the format and quality of the data and prepare them for the matching run. Problems that emerged in the first phase or that are emerging in this one may lower expectations about the ability of the match operation to achieve its goals. Resources such as staff or software to overcome these problems may not be available, and a decision not to proceed with the match may be made. (It should be noted that many of the activities described up to this point are performed only for the first match and are not necessary in second or subsequent matches. For example, the development of software will require little modification in subsequent matches.)

Next, records of the source files are matched for one or more common data elements such as Social Security number, name, date of birth, and address. Many of the positive matches that result are further screened and selected for other criteria, such as the amount of the discrepancy between self-reported income and income reported by other sources, and are then printed. Each of these matches is known as a "raw hit" and indicates that some potential problem (such as erroneous benefit payments) or extraordinary characteristic is associated with the individual who is the subject of the match. Usually, a sample of raw hits is examined in detail in order to verify the existence of a problem and to

[1]OMB's matching guidelines provide that a notice be published in the Federal Register as close as possible to the date matches subject to the Privacy Act are initiated.

characterize it. This usually involves a review of case files or other documentation that may indicate either that inaccurate information was provided to one of the source files, a data entry problem occurred, a program requirement was misinterpreted, or an intention to defraud a program was present or that all the information in the files is correct and the individuals are in complete compliance with all rules and regulations.

Depending on the number of raw hits and the results of the sample verification, the match parameters may be altered and the match rerun. A new raw hit list is generated and another sample is verified. The general goal of this iterative process is to produce the maximum number of raw hits that, when verified, reflect actual problems such as error, abuse, and fraud rather than mismatches of different people or legitimate program enrollments. Once this refinement has ended, a final raw hit list is generated, and the matching process enters the third and final phase.

In hit verification and follow-up, the refined list of raw hits is verified. Hits that are verified as actual data element matches are called "solid hits." Depending on the types of problems identified in the verification, a further validation of information may be made, as in checking with an employer to get the exact dates of a person's employment.

As verification is completed, a variety of actions may be taken. Clients or program participants who are verified solid hits may be asked to respond to the findings of the verification process and the consequent program actions, such as the redetermination of benefit amounts or termination from the program. Instances of program abuse may be handled administratively. Verifications that indicate fraud may be referred to investigators for follow-up and possible prosecution. A review of the magnitude and types of problems reflected in the solid hits may result in recommendations for changes in program operations or regulations. A report may be produced that details the follow-up actions, such as collections, case rebudgeting, program terminations, or prosecutions, and the potential benefits to be realized after all follow-up actions have been completed. The report may also contain figures on benefits actually obtained prior to the issuance of the report. When deterrence is the primary purpose of the match, the report of the match results may be widely publicized.

Overview of Approach and Issues in Making Cost-benefit Analyses of Computer Matches

One purpose of cost-benefit analysis is to provide decision-makers with information that will help them determine whether to implement or continue a program. The most apparent value of the cost-benefit analysis is that it summarizes a variety of information in a single number that gives a clear message, as long as the measurement assumptions underlying the analysis are satisfactory. It also has some secondary benefits, one of which is that it provides information on the magnitude of individual cost and benefit elements that can, in turn, provide insights concerning correctable process inefficiencies.

Cost-benefit analyses, if conducted properly, can determine the value of match operations for achieving efficiency improvements and cost savings in programs whose beneficiaries are being matched. Information about the magnitude of match benefits may be especially relevant in a consideration of the costs that matching might pose to individual privacy and the right to due process. Also, the very process of examining match activities carefully and measuring their costs and benefits may indicate areas in which changes should be considered in match operations.

In the simplest terms, a cost-benefit analysis gives a total of the costs and a total of the benefits of a program and compares these two totals. All costs and benefits should be recognized, measured, and aggregated in comparable terms. The usual yardstick for aggregation is dollars. That is, the monetary value of each cost and each benefit is determined and they are subsequently aggregated as a measure of total costs and benefits.

Although this is the objective of cost-benefit analysis, the process can, and in many cases is likely to, fall short, because of the variety of conceptual issues and practical problems in identifying, measuring, and aggregating costs and benefits. The purpose of this chapter is to identify the issues and problems associated with computer-match cost-benefit analyses. In support of this purpose, we present a broad framework with which to assess the full range of potential costs and benefits associated with a computer match.

Whose Costs and Whose Benefits

A cost-benefit analysis involves several steps. The first is to develop an understanding of the program or process to be assessed: the computer-match operation. Without this understanding, the next step—developing a complete list of potential costs and benefits—cannot be taken. Once the costs and benefits have been enumerated, each one must be

considered separately in relation to strategies that can assign them an appropriate dollar value. Finally, total costs are compared to total benefits.

One of the most basic issues in initiating a cost-benefit analysis concerns the scope or perspective of the analysis, or the question, "Costs and benefits to whom?" Assessments of the costs and benefits of computer matching have often focused narrowly on the agency initiating a match. The narrower the scope of the assessment, the fewer are the costs and benefits that have to be included. However, an analysis that has a narrower scope will provide incomplete information on the overall effect of the match.

In the private sector, it is common practice to limit the perspective of the cost-benefit analysis to the firm or organization conducting the activity. The emphasis is on examining whether the internal goals of the firm are being achieved. Thus, effects on external factors are considered only to the extent that they have repercussions for the firm. It is often argued, however, that public sector cost-benefit analysis requires a broader perspective because of the broad scope of public sector goals. Private firms are expected to pursue goals related to internal benefits, while public entities are expected to fulfill their own internal objectives and to pursue the general goals of the public.

In discussing the cost-benefit analysis of computer matching, we emphasize a very broad perspective: that of the nation as a whole. In other words, we assume that the goal of most computer matches is, or should be, to maximize benefits for or minimize costs to the general public. However, matching is not intended to improve the welfare of those who are discovered abusing government programs, and, consequently, the costs imposed on them as a result of a match should not be counted in a cost-benefit study of the match. However, the costs of matching imposed on program participants who are innocent of wrongdoing should be counted.

Given this perspective, we identified six different entities potentially involved in assessing the costs and benefits of a computer match. We classified each entity in one of three separate groups: (1) government or the matching agency, the source agency, and the justice system; (2) the clients, or individuals with a specific relationship to the program; and (3) the general population, made up of the general public and subgroups with specific relationships that tie them closely to the match. The six entities in these three groups are as follows:

1. government agencies

a. the agency, federal or state, that initiates the match, referred to as "matching agency"

b. the source organization providing the data to be matched, whether another federal agency, a state agency, or a private organization, referred to as "source agency"

c. the justice system, authorized to process and prosecute match-identified cases that cannot be resolved administratively

2. clients

d. the client population, or the participants in a program or members of a population of interest in relation to the match goals

3. general population

e. third parties, who may be requested to provide materials in the process of match-related verification and follow-up activities

f. the general public, or citizens who are not affected or are only indirectly affected by the match; citizens affected indirectly include individuals whose names appear on matched files but who have no involvement with the program.

Not all these entities are relevant to every match. For example, a match that does not result in the referral of hits for prosecution is not likely to have a cost-beneficial effect on the justice system. However, these entities should be considered in assessing the scope of a match.

Specific cost and benefit elements are likely to be associated with each of these entities. We discuss their measurement in detail in chapters 4 and 5 but enumerate them here. The cost and benefit elements that we ascribe to computer match operations were identified in our review of the computer-match cost-benefit literature and an analysis of the match-related activities that the entities perform. Our listing is extensive but not exhaustive; not all the cost and benefit elements are necessarily incurred in any particular match. As with the entities, not all cost-benefit elements are relevant to every match. In some instances, we have incorporated several related elements that other sources identified separately. It should also be noted that specific elements, such as staff

morale, may be costs or benefits, depending on the circumstances of the match. See table 3.1.

Table 3.1: Potential Computer-Matching Cost and Benefit Elements

Entity	Cost	Benefits
Agency		
Matching	Salaries, fringe benefits, travel, materials, and facilities; lowered staff morale; reduced service delivery; degraded client relationships	Avoid overpayments; recover overpayments and debt; improved law enforcement; increased deterrence; improved management; increased public confidence and program support; improved staff morale
Source	Same as costs to matching agency	Similar to matching agency benefits if match is intended to be mutually beneficial
Justice	Salaries and fringe benefits, materials, facilities	Improved law enforcement; increased deterrence
Client	Time, materials, professional services, erroneous termination from program rolls, invasion of privacy	Improved service delivery; increased resources; less participation stigma; identification of underpayments
General population		
Third party	Salaries and fringe benefits, materials, facilities	
General public	Invasion of privacy, discouragement of legitimate clients	Improved program efficiency

Measuring the Cost-benefit Elements

Computer-match cost and benefit elements range on a measurement continuum from quantitative to qualitative. In principle, many of the common cost-benefit elements are quantifiable in monetary terms. The cost of personnel time spent in making a computer match and the benefit from recovering overpayments are examples of elements easily measured in dollars. Overpayments are unlikely to be estimated or counted in terms other than dollars.

Other cost-benefit elements are quantifiable, but the units of measurement are not easily transformable into dollars. Survey techniques can be used to measure the level, and changes in the level, of public support for a program, but transforming this into dollar values is difficult. It might be possible to assign quantitative magnitudes to such elements, but it seems unlikely that a dollar value could be measured for them in a straightforward manner. Although, in principle, individual citizens could be asked hypothetical questions about how many dollars they

would be willing to pay to see a program continue, the validity of using answers to such questions is doubtful.

Other potential cost and benefit elements can be quite difficult to value in either monetary or other quantitative terms. Elements such as the value of law enforcement are inherently qualitative, and any quantification of them will be controversial. With law enforcement, for example, it is possible to measure the relative number of arrests or prosecutions resulting from the use of different techniques for detecting fraud and abuse. In other words, the quantitative measurement of some cost and benefit elements that at first appear entirely qualitative may be possible but not in dollar terms. However, the use of some of these quantification techniques may be costly and of questionable validity.

The presence of qualitative cost-benefit elements for which measurement is conceptually either not feasible or impractical does not mean that cost-benefit analysis should be abandoned, nor does it imply that these factors should be stricken from consideration. Such elements should be identified in the analysis along with the elements that have quantified measures.

Identifying Activities Incurring Costs and Yielding Benefits

In performing a computer-match cost-benefit analysis, it is necessary to identify the activities and events that are considered to be part of the match operation, so that their associated costs and benefits can be assessed. The issues in this identification concern the conceptualization of the activities to be ascribed to the match and how they should be handled in the analysis. Counting costs and benefits should start when activities become sufficiently defined as related to the match.

When to stop counting match activities is not as clear. The question is whether the cost of following up on match hits should be considered one of the match activities. It is sometimes argued that since many of the activities undertaken in relation to hit follow-up are identical to activities that would have been performed regardless of the source of information, they should not be identified as specific to the match. We believe that if the costs of follow-up activities—client contacts, benefit redeterminations, investigations, and the like—are not counted, then the benefits that result from these activities should not be attributed to the match. The question of whether to count them is important, because follow-up activities can represent a large portion of total match costs. Similarly, the benefits that result from follow-up activities—including avoiding overpayments, recovering overpayments, and deterrence—are

likely to account for a large portion of total match benefits. Leaving out the costs of follow-up can distort cost-benefit ratios.

Matches can have both intended and unintended consequences. Failure to consider activities associated with unintended consequences may produce an incomplete picture of the actual costs or benefits of the match. Activities that stem from the match and would not have occurred otherwise should be considered. For example, an unexpectedly high rate of match-related prosecutions or appeals from clients may incur greater administrative costs than anticipated.

Another issue is opportunity cost. This cost can be defined as the benefit forgone from not having done something else with the resources devoted to the match. Other activities could have been directed toward the same goals as the match, as in curtailing fraud and abuse, or unrelated to it, such as processing clients' applications. For example, a member of an inspector general's staff may devote time to planning a computer match instead of improving an error-prone profiling program. Similarly, a program caseworker may work on the match verifying hits rather than other routine activities of processing clients.

The concept of opportunity cost is important in conducting cost-benefit analyses of computer matching, because the opportunity cost of personnel, materials, facilities, and the other resources used in matching represents the true social cost of these resources—that is, the benefits society forgoes because of a match. A perfect cost-benefit analysis of a match would measure forgone benefits directly. Unfortunately, this is rarely possible. For example, benefits forgone when a caseworker does not process a client's application but, instead, investigates hits from a match probably cannot be put into monetary terms. Although the opportunity costs associated with using resources for matching can rarely be measured directly, the budgetary expenditures of government agencies in purchasing these resources (a caseworker's wages, for example) are usually readily available. Consequently, budgetary outlays are typically used as a surrogate for opportunity costs.

Budgetary expenditures may provide a reasonable approximation of opportunity costs, if several assumptions or conditions are met about the full and efficient use of the match's resources in alternative activities. The notion of "efficiency" implies that the last unit of a particular type of resource used in performing other productive activities produces output that is equal in value to the cost of the unit, and it implies that the employees involved in the match activities would not have been idle

or underemployed in the absence of the match. For example, if the last caseworker who was hired receives $15,000 in wages, the value of this worker's output should also equal $15,000. If the worker is diverted from routine client-processing to computer-match activities, the opportunity cost (that is, the value of the output forgone because the worker did not process applications) will equal the $15,000 budgetary expenditure for the worker's wages.

In some instances, the conditions outlined above may not be met. Personnel, materials, and facilities can be underused or used inefficiently. This may be partly because of the difficulty of measuring the value of output produced in a nonmarket setting. To the extent that resources would be underused or inefficiently used in the absence of matching, budgetary expenditures for them will fail to reflect accurately the opportunity costs (the benefits forgone) of using them in matching. For example, if the caseworker in the paragraph above were to produce output worth $14,000 or $16,000 by continuing to process clients' applications, the actual cost of using this worker to investigate raw hits is not the worker's $15,000 salary but either $1,000 less or $1,000 more. Nonetheless, in the absence of information with which to make an adjustment in the relationship between the value of expended resources and the value of subsequent output, they are assumed to be equal.

That individual matches can be part of a series (for example, a match that is performed quarterly and uses wage and unemployment benefit records) or a single occurrence (for example, an experimental match of two files in order to determine their potential for the production of hits) underlies another issue related to the inclusion of match activities in a cost-benefit analysis. It is comparatively easy to define a realm of activities for the performance of a single match, because relatively fewer activities are involved than in a series of matches. Defining the range of activities for matches in a series is more problematic.

A series of matches entails the question of whether it is more appropriate to consider the costs and benefits of each match in the series separately or to treat all the matches in the series together. This is an important question, because of the likely disparities in costs and benefits between at least some of the matches in a series.

The first match in a series is likely to incur the greatest costs for the initiating agency and the source agency, because of the newness of the process. Specific software will probably have to be developed, and

processing raw hits can be and often is relatively inefficient. On the benefit side, the number of solid hits may be greatest in the first match, because a backlog of fraudulent cases or cases of error may be identified. In subsequent matches, error or fraudulence that occurred since the prior match will probably constitute the majority of identified cases. Similarly, overpayments will have been made mainly during the interval between one match and the previous match, but the overpayments identified in the initial match are likely to be those that have gone undetected for years. The point is that costs and benefits are distributed unequally across a series of matches. Only after a match series has "stabilized" are costs and benefits likely to be consistent from one match to the next.

Whether or not the "start up" costs associated with a first match, as well as the relatively large benefit from the match, should be counted in a cost-benefit analysis depends on whether the purpose of the analysis is to determine whether a new series of matches should be initiated or an existing series should be continued. When the second or a later match in a series is being considered in terms of costs and benefits, the "start up" costs of the first match are most appropriately viewed in relation to the match in question as "sunk" costs. Sunk costs are resources spent before making the decision to undertake the second or a later match in a series, and they are unaffected by decisions related to later matches and, consequently, should be ignored in cost-benefit evaluations of the later matches.

Time Factors in Cost-benefit Analysis

Cost-benefit analysis can be undertaken anytime during a match operation. It can be conducted prospectively as a prematch analysis, in order to assess whether a match should be initiated. It can be performed concurrently, in order to evaluate an ongoing operation, or retrospectively, in order to determine whether a completed operation was successful from a cost-benefit standpoint. A concurrent analysis is performed at the beginning or in the middle of verification and follow-up activities, when only a small portion of the match hits have been processed. The postmatch, or retrospective, analysis is conducted after the majority of all hits have been processed.

The difficulty with prematch analyses is that they must, of necessity, rely on estimates rather than actual cost and benefit figures. Their strength as predictions is directly related to the accuracy of the data sources from which they are drawn and their applicability to the current match. Concurrent and postmatch analyses, in contrast, are likely

to draw on better information, but they too depend on the extent and accuracy of records related to the match.

Postmatch analysis can also look forward, in the sense that the lessons from past operations can be applied to similar matches being considered for implementation. In fact, in prematch assessments, the inherent problem of measuring something that has not yet occurred requires substantial reliance on information from previously conducted concurrent or postmatch analyses of similar operations. Concurrent and postmatch analyses of matches are of considerable importance, because they provide information that can be used in prematch analyses that follow them.

The analysis of computer-match costs and benefits should consider their timing and whether their duration warrants discounting. Some of the costs and benefits associated with a computer match may occur at different points in time. Costs, in particular, are likely to be incurred before many of the benefits are received. For example, follow-up costs are typically incurred months after a computer match has been run, but overpayments are typically recovered over an extended period, perhaps years, after the match. In instances like these, the various cost-benefit elements cannot be directly compared.

Money that must be spent over the next few months is valued more highly than an equal sum of money that will not be received until several years from now. It is important to consider the effects of time on the value of money, so that all elements in the analysis are comparable. Discounting accomplishes this. "Discounting" is defined as translating the value of money at one time into the value of money at a different time. Of course, in a cost-benefit analysis, the goal is to make sure that all costs and benefits are considered in relation to the same time and are, thus, comparable with one another.

Understanding the steps in a given computer match and how they will unfold provides the information needed to decide whether discounting is appropriate. If relatively short spans of time are anticipated (or have passed, as in retrospective assessments), discounting may make no significant difference in the analysis. If all match costs and benefits are incurred in a relatively short time, and interest rates are low, discounting may not be needed. Because match benefits are generally not realized until after most of the costs have been incurred, failure to discount when long periods of time are involved tends to yield overestimated benefits compared to costs.

The effects of discounting may be important also when several matches are being assessed as a unit. A series of matches could extend for years, making the value of money over time essential to a sound analysis.

Alternative Approaches

Our discussion of computer-match costs and benefits addresses the broad concern of assessing the net benefits of a match when all costs and all benefits are considered. Other approaches may provide a different and somewhat more limited perspective on the fiscal effect of a computer match. Earlier in this chapter, we briefly discussed an approach that takes into account only costs and benefits to the matching agency. This approach can provide information useful to the agency for internal management decisions but does not take into consideration many potentially important costs and benefits to other entities. It assesses the match as an agency policy rather than as a public policy in the broader sense.

Another approach assesses only the marginal costs of and benefits from a match. That is, only match-related costs and benefits over and above existing costs and benefits of nonmatching activities are counted. The interest in this approach is in determining the additional costs or additional benefits that accrue from matching, not the total costs and benefits of the match. In this approach, it is important to limit both costs and benefits to their marginal increase in amount. It is inappropriate to balance marginal increases in costs against total benefits.

In some cases, agencies may want to decide between computer matching and an alternative activity that has the same goal. They may ask whether computer matching or the alternative is the most cost-effective. Answering this question requires a focus on the difference between the marginal costs and benefits associated with alternative activities. This approach rests on establishing or assuming that the value of either the cost or the benefit elements of computer matching and the alternative are not significantly different from each other. This assumption of equivalence between the cost or benefit elements of matching and some alternative may not be appropriate if only limited empirical data are available on the costs and benefits of matching.

Estimating or Tracking the Costs of a Computer Match

The complexity of operating a computer match, the variety of settings in which matches are performed, and their diversity of purposes make it exceedingly difficult to detail the exact procedures that should be followed for a particular match. However, it is possible to identify broad cost categories for which data should be developed and to describe some methods for obtaining these data. In this chapter, we address the full range of costs that we have identified for computer matches. The significant sources of match costs and their associated cost categories, identified in chapter 3, are the framework for this chapter.

Costs to Government Agencies

Collecting cost data from matching, source, and justice agencies requires identifying the persons, materials, and facilities involved in all the major activities associated with a match. This includes not only data processing costs but also the costs of planning activities and following up on hits produced by the match. Costs associated with travel necessary for the match activities should be included, and so should costs associated with the performance of cost-benefit analyses of the match. In addition, consideration should be given to such qualitative costs as potential changes in staff morale, program efficiency, and client relationships caused by the matching or source agencies. As we noted in chapter 3, the resources spent in performing computer-match activities are typically used to represent the opportunity costs of the match.

Although the agency with major responsibility for conducting a match usually incurs greater costs than other participating agencies (such as those supplying the data), the costs to these agencies should be included. In general, the approaches to assessing the costs of a matching agency are applicable to other agencies.

Salary and Fringe Benefit Costs

What to Measure

A variety of personnel can be involved in the operation of a match from its conceptualization to the final processing of the match hit list. In the match definition and development stage, personnel from within the agency initiating the match are engaged in research and analysis on the feasibility of the match and, in many cases, the development of a formal proposal. In addition, they meet with staff in programs whose files are to be matched. Furthermore, staff lawyers and privacy officers may be

consulted on legal matters related to the match and the publication of notices. Agency cost analysts and economists may be involved in budgetary aspects of the match. Individuals from other agencies who may have relevant information or data also shape the development of the match. Finally, some matches may require the advice of persons who have pertinent information from commercial or other nongovernment institutions.

During the match initiation and production phase, the combination of personnel changes. It includes some of the individuals who participated in the match definition and development phase as well as such data processing personnel as programmers, operators, and systems analysts.

In the final phase, hit verification and follow-up, the commitment of personnel resources may be the highest, especially if only limited automated refinement of the hit list was possible. Staff time may be spent in such activities as locating and reviewing clients' records, contacting and interviewing clients, contacting third parties, computing and collecting overpayments, processing appeals, and preparing cases for hearings. Although match operators or auditors may verify and follow up on a sample of hits to assess the viability of processing the full list, line personnel usually perform most of these activities.

In addition to the match personnel and program personnel who perform the verification and follow-up activities, personnel are needed to deal with specific hits that will be referred for more intense follow-up. Indications of fraud can lead to referral of a hit to investigators for additional validation. The results of an investigation may lead to the referral of the case to prosecutors for legal action. Program personnel and investigators may refer a case for an administrative hearing. In brief, personnel from a variety of different units in one or more programs and from the legal system may be involved in this phase of activity.

Approaches to Measurement

Identifying and obtaining data on the costs incurred by all the individuals involved in a match is a formidable task. However, the costs are quantitative, and it is possible to assign monetary values to them, even if an analysis is performed before the match. The basic objectives of cost analysis for all phases of match activities are to (1) identify the persons who participate in the match and (2) obtain values for (a) their hourly salary rates, (b) their fringe benefits as a fraction of their salaries, and (c) the number of hours they spend in match-related activities.

To establish cost, generally the hourly salary rate for each person is multiplied by the number of hours each one spent in match activities and then the total is found. This sum is multiplied by the fringe benefit rate, and the resulting figure is combined with the sum.[1] When individuals are not employed by federal agencies, established fringe benefit rates associated with their salaries should be used.

Given the objective of analyzing the cost of salary and fringe benefits, several general approaches can be considered for developing the necessary data. Adopting a particular approach depends on whether the match is being performed for the first time or is a recurring match and on the point in the match process at which the analysis is to be performed. For concurrent and postmatch analyses, the tracking or reconstruction procedures to be discussed can be applied to either a first or a subsequent match. In a prematch analysis of a match that is being performed for the first time, there is less information on which to base the analysis. Prematch analyses, because of their timing, require predictions of most of the costs of the match. Table 4.1 gives an overview of approaches for predicting and assessing salary and fringe benefit costs.

[1]OMB circular A-76 contains guidance on fringe benefit costs to federal agencies.

Table 4.1: Approaches to the Cost Analysis of Salary and Fringe Benefits

Approach	Prematch	Concurrent	Postmatch	Description
Review proposal	X			Estimate personnel costs from a match plan if it is adequately detailed by proposed activities, schedules, and staff
Extrapolate from previous matches	X			Estimate costs for the next match in a series from data from one or more previous matches, preferably a second or later match (costs may be higher for the first match; estimate costs for a first-time match on cost data from similar first-time matches if they exist)
Use expert opinion	X			When no previous cost data are available or data from related or previous matches must be refined, ask persons experienced in matches to review planned match activities for probable costs
Use routine job billing		X	X	Identify time matching and source agency personnel spend on match from management records or job billing system
Track and reconstruct time		X	X	When management records or job billing system do not exist or are inadequate, identify time personnel spend on match from records maintained for this purpose or reconstruct it by reviewing documents or interviewing persons involved in match or both
Combine approaches	X	X	X	Any appropriate combination of the other approaches, depending on the type of match (first-time or recurring) and when the analysis is performed (prematch, concurrent, or postmatch)

Routine Job Billing

In agencies that use management information or job billing systems to track employees' time on various projects, it may be possible to obtain salary and fringe benefit information without increasing staff reporting requirements. Doing this, however, would require giving the computer match a unique identification code in the system. Plans are under way at SSA, for example, to add a code to the work-reporting system in order to track the time to be spent on a new match. The office of inspector general in several agencies maintains an information system that tracks the time staff spend on various audit projects. Computer-match projects are tracked as audits on these systems and are assigned their own codes. In one agency, the tracking system disaggregates staff time by type of activity.

In agencies with such tracking systems, reporting methods should be reviewed to ensure that they accurately reflect the time personnel spend on the match. A system for tracking the personnel costs of the match will be more accurate and require less modification if it (1) includes all individuals who are involved in the entire match operation, (2) initiates tracking early enough in the match to include most of the early planning

activities, and (3) does not include staff time spent on activities not related to the match or spent on planning future matches. (Staff time spent developing a new match may be assigned to an ongoing match for convenience, until the new match is formally assigned a project number.)

Given the diversity of units and types of personnel in a match operation, it is unlikely that any one tracking system will be sufficient for tracking all staff costs. For example, the tracking systems used by inspectors general do not include the time of the personnel who are outside the offices of the inspector general and may be involved in follow-up activities. Even with a comprehensive tracking system, persons from other agencies are not included; time information has to be obtained for them by other means.

Time Tracking and Time Reconstruction

For individuals outside the agency, or where an information system does not exist or is inadequate, another basic approach is to track the hours they spend from the initial discussions about implementing the match. Implementing this tracking process requires assigning responsibility for documenting and monitoring the planning activities to someone who is closely involved with the planning. Creating a file that identifies the major planning activities during the match definition and development phase, the persons involved, and the time they spend would provide sufficient information to reasonably identify the cost of resources.

If a tracking file has not been created or is incomplete, a modification of this approach is to try to construct the information that a tracking file would have contained. This requires a review of available planning documents in order to identify persons and activities. The persons or their supervisors are interviewed in order to obtain their estimates of time and to check that all relevant individuals and activities have been identified. Reconstructed estimates of this kind are inferior to figures obtained through routine tracking systems, but they provide some approximation of the actual time spent and provide a reasonable basis for developing the personnel costs of match activities. Care should be taken with time reconstruction and time tracking procedures that the costs of implementing these techniques are reasonable for the scope of the match and that paperwork is not a burden.

Combined Approaches

Whether the cost analysis is performed prior to, during, or after the match, some combination of approaches may be necessary. One measurement technique can be used as a supplement to or as a substitute for another technique. For example, in a postmatch analysis, a management information system may provide data on activities that occurred as the direction of the match became more clearly defined. A file of planning papers, meeting notes, or other papers may provide information on some of the earlier activities and identify the persons to contact for additional information and estimates of time spent on earlier activities not tracked by the management information system. In data processing units where some type of intra-agency or interagency service billing occurs, recordkeeping may provide a fairly complete account of the personnel and the time they spent on developing software related to the match project. In one SSA match, for example, automated data processing is performed by a state agency and costs are reimbursed by the matching agency. In this instance, the data processing costs have already been specified.

Prematch Cost Analysis

For prematch estimation, some basis for developing projections of the costs of starting up the match and producing hits has to be established. Several possible approaches could be used individually in order to estimate different cost components or in combination in order to help strengthen the development of estimates. In principle, prematch analyses of salary and fringe benefit costs can take approaches associated with concurrent and postmatch cost analyses. Since prematch estimation typically does not occur until the end of the match definition and development activities, personnel costs for this phase should need to be not predicted but tracked as they occur. When a prematch estimation of costs and benefits becomes necessary, most match development activities should have been completed and their actual tracked costs included in the analysis.

In the following sections, we discuss three approaches to developing prematch estimations: proposal review, extrapolation from previous matches, and expert opinion. It should be noted that expert opinion uses extrapolations indirectly from previous matches. Proposal review may do this too, depending on how the match plan is derived. However, with these two approaches, the use of previous match data is more informal and implicit and may not be based on specific matches.

Estimates Based on Proposal Review

Basing estimates on proposal review is most feasible if the match definition and development phase concludes with the preparation of a

detailed match plan. If it does, the proposed activities, time schedule, and staffing can serve as a basis for the personnel cost estimation. Salary data can be used to compute anticipated salary and fringe benefit costs.

This approach, however, is only as good as the plan. If the plan does not accurately reflect what the actual match will involve, and reflect it in sufficient detail, the estimates may be neither sound nor useful. Therefore, it is necessary to make sure that the plan includes a complete description of the activities that will be performed by all the source agencies and the staff required to perform them. Plans that lack all or some of this information will place an additional burden on the cost-benefit analysis by requiring its development before the cost estimates can be made.

Extrapolation From Previous Matches

Another possible approach involves the search for and identification of similar matches for which salary and fringe benefit costs have been tracked. Their data can be used for developing estimates, if there is some evidence to support their use and if the procedures that were followed in generating the data were clearly and completely described. For example, if a match has been proposed because of an audit report establishing the cost of validating individual cases in the audit process, this figure could be used, in conjunction with a projected number of hits, to compute follow-up costs. However, this approach depends on the completeness of the audit validation effort. If it has not approximated the follow-up procedures, there is no sound basis for estimating these costs.

Data from one or more matches in a recurring series can provide a basis for developing a prematch cost estimate of the next match in the series. This technique requires some consideration of the position of the match in the series. In general, subsequent matches should be less expensive than the first match. The cost of verification and follow-up activities for later matches should be less, since fewer persons are matched, and problems in the development and production phase of the first match will have been resolved. For these reasons, cost data from the first match probably provide a less reliable basis than cost data from later matches for projecting a prematch cost estimate of the next match in the series.

Moreover, subsequent matches can be refined and made more efficient from experience with previous matches. For example, the costs OPM reported for performing one of its match operations a second time were roughly 50-percent lower than the costs reported the first time. The

reduction was partly attributable to the increased use of program personnel instead of investigators to perform various verification and follow-up activities in the second match. OPM used this cost data on this second match as a basis for projecting annual staff costs for the continued operation of the match series. Match operators in the VA office of inspector general reported that performing a match with one state's wage records allowed them to develop and use automated screening techniques to generate more reliable hits when the match was repeated with another state.

These observations suggest that when a match has become routine and when cost data from several of the matches are available, a trend line should be established and, barring major revisions in the match operation, used as a basis for projecting the costs of future matches. IRS has used this technique in conjunction with proposal review as a means of estimating the cost of various cyclical audit projects.

The approach described above obviously cannot be used for first matches, because they have no direct predecessors. However, it may be possible to develop cost estimates from similar first matches for which cost data have been collected. The similarity may be based on the programs involved in the match. For example, the costs of matching a federal program file with a data file from one state might be used as a base for estimating the costs of a match of the same federal file with a data file from another state. The usefulness of this approach is limited by the extent and nature of the dissimilarities between the proposed match and the match or matches used for cost data. If the proposed match is expected to have different charges for automated data processing, a different hit rate, a different emphasis on benefit types, or variations in follow-up procedures, cost data from other matches will be less useful unless adjustments can be made to compensate for the differences.

One of the problems in making a prematch adjustment of estimates is in the development of an expected hit rate upon which to base verification and follow-up costs, especially when no relevant previous matches can be identified. One technique is to use information from case quality-control review data. Some programs independently verify a random sample of cases for the accuracy of information and the appropriateness of program actions. The data on error rates can provide a figure from which to develop an adjusted hit rate. The appropriate application of this technique depends on using only the case error rates associated with the types of errors the match is expected to detect. Since a match

cannot detect all case problems, an overall case error rate should not be used for projecting a match hit rate.

Expert Opinion

When no previous cost data are available from a match resembling the proposed match, or from any other first match, a speculative approach might be considered. Persons with experience in first matches might be asked to consider the planned match activities and provide their expert opinion about the costs that will be incurred. Experienced match operators might also be consulted for help in adjusting estimates derived from data on previous or related matches.

Major Sources of Error in Prematch Cost Estimates

Errors in prematch cost estimates arise when the match, for whatever reason, is not conducted according to the intentions of the match when the estimates were made. This may occur as a result of either deliberate management decisions to change the match procedures from the original plan or unexpected events or both. For example, if plans to set priorities for hits were not followed during verification and follow-up, the underlying assumptions of the prematch estimates will not be valid, and it is not likely that the prematch estimates will correspond closely to the actual results. First matches are especially prone to unexpected events that require changes in the match operation.

For first matches, unanticipated events and resultant management decisions can lead to estimation error from problems associated with developing and determining the number of solid hits the match is to produce. During match initiation and production, unanticipated events in the iterative process of refining raw hits by the automated procedures that are used to develop a final list of solid hits can make prematch cost estimates of first matches speculative and unreliable. Data of poor quality, or the lack of critical data elements in the original match files, may limit the number of possible hits or require unexpectedly greater effort to produce a set of solid hits than originally planned. The usefulness of basing hit rate estimates on such techniques as case quality-control review for error rates may be undermined by unanticipated technical problems.

Even if match planners have some indication that specific problems may arise, it may be quite difficult to correctly anticipate the effort that will be required to overcome them. For example, the lack of Social Security numbers, usually a critical match element, for program clients at HUD reduced the number of potential hits that could be obtained. The

agency's attempt to acquire this information would have incurred additional costs. Similarly, in a Department of Education match to locate student-loan defaulters, employment addresses provided by one of the matched files were not current, because a number of individuals identified by the match had changed jobs and their forwarding addresses were not available, so that additional, not always successful efforts had to be made to track down current addresses for them.

The limitations of software and initial verification procedures may result in a large proportion of raw hits that are not substantiated after complete verification. It may be that little can be done, in any practical way, to incorporate such possibilities into prematch cost estimates, but a sensitivity to these factors may be helpful in developing a cost range, and they underscore the need for monitoring costs as a match proceeds.

Another complication in developing prematch costs is that in the verification and follow-up phase, some costs depend on the number of solid hits from the match. Costs associated with such activities as obtaining responses from third parties, recomputing benefits and overpayments, collecting debts and handling terminations, conducting appeals hearings, and proceeding with prosecutions will be a partial function of how many hits require that such actions be taken. In some instances, the decisions of clients at various points in the follow-up process and the responsiveness of a third party to a verification request influence the cost of processing a hit.

In one match, the lack of adjusted estimates or insensitivity to the distribution of hit dispositions resulted in a prematch cost-benefit analysis in which the number of projected hits, if multiplied by the estimated follow-up cost per hit, would yield highly inflated total match costs. Included in the per-hit cost analysis was an estimate for investigative costs that constituted more than two thirds of the total cost per hit. Not all hits require investigation, however. The analysis could have been improved with some projection of the number of cases expected to require investigation or with an acknowledgment that the cost-per-hit analysis identified the maximum possible follow-up cost.

Concurrent Cost Analysis

Some match operators have noted that performing a pilot match encompasses many activities that must be performed for a full-scale match. The costs associated with defining and developing the match are essentially the same. In the match initiation and production phase, automated data processing costs are only marginally higher for a full-scale match

than for a pilot match. Given such circumstances, some match operators prefer to perform a full match and follow it with a pilot sampling of hits for verification and follow-up. This practice is frequent when a match is being performed for the first time.

If the sampling procedures are properly performed and the verification and follow-up procedures closely approximate the procedures to be followed in the actual match, the estimates obtained from a sampling of hits should provide a reasonable basis for determining not only the potential benefits of processing the entire hit list but also the major costs of the match before they are incurred. SSA plans to use this type of concurrent analysis to assess the costs of verification and follow-up activities for its match of IRS nonwage-income records with Supplemental Security Income program rolls.

Analyzing costs based on a pilot sampling of hits concurrently with tracking previous match-development and automated data processing costs can yield cost estimates that are methodologically more defensible than prematch cost estimates. Concurrent analysis provides knowledge of what was spent and a better assessment of what is still to be incurred if a full-scale verification of the hit list is to be implemented. Moreover, sample verification provides some insight into the nature and importance of the more qualitative costs to clients, third parties, and participating agencies. In addition, unintended costs may be revealed during the sample follow-up. For a concurrent cost analysis, the basic approaches to measurement we discussed earlier, especially time tracking, can be applied.

Postmatch Cost Analysis

Postmatch, or retrospective, cost analyses are performed to provide a final accounting of match costs. Because the verification and follow-up processes on all hits may extend over a fairly long time, it is not possible to specify precisely when a postmatch analysis should be performed. A significant portion of hits should be resolved before initiating the analysis, in order to ensure a basis for determining actual costs that is stronger than that used in a concurrent analysis. However, waiting until all hits are processed may result in an untimely analysis. One approach to this problem is to perform periodic interim assessments, monitoring the rate of expenditure during the verification and follow-up phase so that future costs can be extrapolated from the time trend.

Approaches to cost measurement for postmatch analysis are essentially the same as approaches for a concurrent analysis. However, the considerably greater amount of data at the postmatch point may provide a broader basis for developing estimates. Cost data can be obtained from hit tracking systems instituted at the start of hit production and refinement. Feedback sheets, logs, and periodic reports could be used to collect information on the persons involved in the match and the time they spent in processing match hits. At OPM, for example, the unit responsible for performing hit verifications of several different matches maintained daily records of the time spent on hits from each match. Collecting additional information on the number of hits processed in the time period worked could provide a basis for calculating an estimate of personnel costs per hit.

For many matches, it is unlikely that extensive tracking systems have to be implemented, but a combination of approaches may have to be employed in a postmatch analysis. A sample of hits could be tracked along the lines discussed for concurrent analyses. If this is not feasible, activity reconstruction or consultation with experts could be used to develop estimates. However, these are second-best techniques and should not be used when it is possible to identify individuals and collect time data by tracking or monitoring.

Travel

What to Measure

For some matches, it may be necessary for some personnel to travel in order to perform match-related activities. Because airfare, lodging, and other travel expenses can be costly, they should be tracked and included in the total match costs. Assessing travel costs should not be limited to the costs of the matching agency but should include the travel costs of source agencies and other participating organizations.

Approaches to Measurement

In most cases, a review of travel vouchers submitted by personnel involved in a match provides the information necessary to compute travel costs. In some instances, a cost tracking system may identify travel expenditures. If travel vouchers do not contain adequate information, or if a tracking system does not clearly identify match-related travel costs, it may be necessary to retrace events by interviewing the personnel who traveled and reviewing the documentation, in order to

identify the agency vouchers that pertain to the activities of the match. One problem in this may be differentiating match-related costs on trips taken for several different purposes in addition to the match. Reviewing travel vouchers and interviewing the travelers can help determine the amounts that should be charged solely to the match.

Materials

What to Measure

Costs for computer-match materials include expenses for purchasing office supplies, computer tapes, and disks; obtaining documents for planning the match; publishing notices and arranging for publicity; producing special notices; and tracking forms and response forms for clients and third parties. Material costs also include reimbursements that a matching agency makes to a source agency or third party for data. Fully reimbursed activities and materials should not be included in assessments of source agency or third-party costs, since the matching agency reimburses these. When the reimbursement covers only part of the costs, it will be necessary to exclude the reimbursed portion in order to avoid counting the same costs twice.

Approaches to Measurement

The objective of the analysis of material cost is to identify the type, quantity, and price of materials used as part of or in support of match activities. Like personnel costs, material costs are quantitative and can be tracked. Tracking them requires the maintenance of inventory lists, supply bills, or purchase orders that identify the materials individuals used in match activities. Care should be taken that tracking these costs does not make reporting practices burdensome or more expensive than the materials being tracked. Standard cost-accounting procedures, if they are available for prorating the cost of materials within the agency, may be useful in developing material costs without establishing special reporting or tracking procedures. The discussion in the following section on facilities illustrates a type of procedure that might be used.

Facilities

What to Measure

Facilities costs include expenses associated with an agency's maintenance and use of office space, telephones, computers, and furniture.

Sometimes called "overhead" or "indirect" costs, they support a variety of activities rather than a single project such as a computer match. That is, these costs are shared and must be apportioned across the range of activities that incur them. The objective of the cost analysis for the match is to apportion to the match the costs that are attributable to it.

Approaches to Measurement

The basic measurement approach consists of expressing an agency's total facilities cost as a fraction of the agency's total salary expenditures. This figure may be readily available in some agencies; some agencies compute this rate for other uses. The rate can be multiplied by the total salary spent on the match to yield an estimate of the facilities costs for the match. Care should be taken to avoid counting twice items such as office supplies and forms, among others, included in material cost. The estimate will be biased to the extent that units or persons associated with the match used a disproportionately larger or smaller amount of the facilities than others not involved. At SSA, for example, this bias can be avoided, because SSA uses differential overhead rates in computing reimbursable costs for units within the agency.

For some phases of a match, notably initiation and hit production, computer costs may be obtained from automated data processing billing routines. If computer account numbers for a match operation are unique, or if match-related jobs can be distinguished, computer costs can be determined directly. However, if automated data processing billing routines are relied on to establish cost, it should be determined that the charging algorithm and resulting costs are reasonable. Automated facilities that typically serve only internal users may apply billing procedures that were established for other purposes and that do not reflect the real cost of these resources. To avoid counting twice, it should also be determined that separately computed data processing costs are not inadvertently included in the facilities costs.

Qualitative Costs to Government Agencies

While the quantitative costs of staff salary and fringe benefits, travel, materials, and facilities are important components of a government agency's computer match costs, qualitative costs may also be associated with a match. Qualitative costs are difficult to assess and integrate with monetary measures, but they should be given careful consideration, because they may involve major issues with serious consequences. The difficulty in measuring these costs, however, means that it is not possible to identify or specify measurement approaches to the extent it is possible to do with quantitative costs. In this section, we suggest some

general approaches to quantifying the costs that illustrate a few potential ways of obtaining information about these costs.

Staff Morale

The possible change in a government agency's staff morale, as it is associated with situations in which program personnel are required to perform hit verification and follow-up activities, has been described as a qualitative cost or benefit of computer matching. Change in staff morale is associated not only with the matching agency but also with source agencies, if their staff participate in verification and follow-up activities. Its being a cost or a benefit appears to depend on whether hit verification is thought to have a good probability of detecting inappropriate payments or eligibility and on whether the staff who perform hit verifications believe that restitution and prosecution are likely to follow. Morale is likely to suffer if the staff feel already overloaded with other work, verification is time consuming, the chance of finding a true instance of a solid hit is low, or little action occurs when a solid hit is found.

Staff attitudes about match-related activities are the primary indicators of morale. Level of productivity and behavior such as absenteeism might also reflect the status of, or changes in, morale. However, a variety of other factors that may affect performance or behavior make it quite difficult to demonstrate a relationship between low productivity, for example, and staff morale. Surveys of staff attitudes and interviews about conducting verification and follow-up activities can provide some information on the effect match activities have on morale.

Service Delivery

When a large number of hits must be verified or the time needed to perform verification is long, a decrease in the performance of other assigned tasks may be a qualitative cost. As with staff morale, this cost can apply to source agencies as well as the matching agency, if staff in the source agency are involved in the match extensively. For example, administrators in two agencies that perform match verification and follow-up activities expressed concern that scarce administrative resources would be consumed if more matches were initiated. They noted that matches may save program benefit dollars but performing a match requires the expenditure of a program's administrative resources. If administrative resources are at a fixed level, placing an emphasis on matching might result in other activities in the administration of the program being given a lower priority.

Changes in the nature of service delivery are a qualitative aspect of a match's opportunity costs. Resources spent on the match are no longer available for these other activities. As we noted in chapter 3, these costs can be measured in terms of the time, travel, materials, and facilities expenditures for the match. From a program-operation perspective, it may be useful to monitor match workload measures together with such nonmatch workload measures as case processing time, response time for clients' inquiries, and number of completed case actions. Monitoring can help establish the relationship, if any, between an increase in match activities and the performance of other program tasks, and it can help assess whether any loss in program efficiency, if it occurs, seems to exceed what was expected.

Change in Client Relationships

Another qualitative cost to agencies performing and participating in a computer match is related to an agency's interaction with clients. Program participants, especially those who are affected by a match, may subsequently not respond to the program or refrain from making further contact with it. They may alter or withhold information or otherwise change the nature of their interaction with the agency. For example, when follow-up actions have included the offset of tax refunds, some of the persons whose tax returns are offset have not filed tax returns in the subsequent year or may otherwise have altered their tax payments.

Developing useful measures to assess change in clients' relationships because of a match requires identifying the possible ways in which clients may alter their behavior and collecting data that will reflect the change in behavior. Reviewing previous or similar matches may provide some indication of the changes in client relationships that might be expected. To the extent that these changes can be identified and quantified, some assessment of the cost may be possible. Measurement would require the design of special studies using control groups or the collection of prematch baseline data similar to the data used for assessing deterrence benefits (discussed in chapter 5). Given their complexity and cost, such studies would not be part of a routine analysis; they should be conducted when a potential change in client relationships would be expected to have a significant influence on the operation of the program.

Costs to Clients

When an agency believes that evidence from a computer match shows that some clients receive or have received benefits to which they are not

entitled, the agency may ask the clients to provide evidence of the legitimacy of their claims. This might be the verification of health status, income, or assets or the submission of demographic data (such as a birth certificate) or a variety of other evidence.

What to Measure

To respond to these requests, clients with legitimate claims must spend time and other resources to submit the required documentation. Their use of time and its cost are similar, in many respects, to the costs associated with the time an agency's personnel use. This time should be valued as a match cost, but it is methodologically difficult to establish a basis for determining it. Other costs that clients may incur include the use of materials and facilities and, in some cases, the services of outside experts who are called on for help in responding to a verification request.

Approaches to Measurement

Salary is a reasonable way of valuing agency employees' time, but there is no convenient yardstick for the value of clients' time. Their salaries could be used as a surrogate indicator, but it might be controversial, since the verification activity is separate and distinct from their regular employment. For example, if an attorney who charges clients $100 an hour performs the task of photocopying and mailing her own tax return to satisfy an agency's verification request, should the time she spends on this task be valued at $100 per hour or at a lower rate based on its clerical nature? It could be argued that an attorney who valued time for this task at less than $100 an hour might look for more work at her rate and pay someone else to do the copying and mailing.

Is the value of time for unemployed clients, who have no observable salary on which to base cost, equal to zero? Clearly not, because the clients must forgo other activities in order to submit the requested documentation, so that the time they spent was of some cost to them. One might distinguish between clients who are involuntarily unemployed and those who are voluntarily out of the labor force. For the former, the value of time could be evaluated at their previous market wage; for the latter, a potential market wage could provide an estimate.

A reasonable evaluation of time in dollars per hour is elusive, although the minimum wage could be assigned. This approach would provide some basis for cost estimation but would tend to understate the value of time for persons who earn more than the minimum wage and overstate it for persons employers are unwilling to hire at the minimum wage.

Information about the client population should be used in developing an upper limit and a middle range for the value of clients' time. An analysis could then address the range of possible client costs.

The practical problems of collecting an estimate of the time clients spend on verification can be difficult. Although querying them about their verification activities can be subject to a variety of problems, it may be the only feasible way of obtaining an estimate of their time. In a match for which clients are not involved in any significant documentation tasks but simply report for an interview, their time could be reasonably estimated by noting how long the interview took. Travel time could be estimated by noting where the clients live and their mode of transportation to the interview.

Obtaining actual data of this type on all clients is too burdensome for an agency to incorporate into routine match operations. It may be more reasonable to estimate the time clients are involved in verification by collecting data from a sample of clients, examining similar matching activities, or asking experts what they think might be required by the verification process.

Although the major cost for clients is probably time, they pay, on a smaller scale, for materials and facilities—for example, documents and travel—just as the matching agencies do. A special form of expense for clients might be the services of others, such as physicians and lawyers, to respond to a verification request. When outside services are purchased, their cost can be substantial. Whether these data should be collected depends on both the incidence of the use of outside services and their cost. As with time estimates, data on these costs can be collected by using survey techniques, but a client's bias in reporting time and out-of-pocket expenses may lead to an overestimation of these items.

Qualitative Costs

For clients, as with government agencies, qualitative costs may be associated with a computer-matching operation. They should be given careful consideration, despite the difficulty of measuring and integrating them with monetary measures. One such cost may be the invasion or loss of privacy. This cost applies not only to the general public (as we discuss in a later section of this chapter) but also to the legitimately entitled program clients whose names appear on a match hit list.

In verification and follow-up, an agency may seek more recent or additional information about legitimate clients. It may request this information directly from the clients or obtain it from other sources. In either case, clients may feel that personal or confidential information must be divulged in order to satisfy agency requests and demonstrate the legitimacy of their participation in the program. For these clients, the price for continued participation may include not only the quantitative costs we discussed above but also the feeling of an erosion in or the loss of control over personal information.

A more extreme cost may occur when a client is terminated in error from a program's rolls as a result of a computer match. A client's cost in this situation can include not only the elements discussed above but also mental anguish and physical distress. Obviously, one way of reducing these costs is to reduce the incidence of the erroneous terminations of clients from the rolls. Federal guidelines on the use and accuracy of client information have been established to help avoid this problem.

Costs to Third Parties

When a computer match identifies cases with possible problems, a request may be made to a third party for information. A client's employer or former employer might be asked to provide data on the client's earnings and term of work. A school might be asked to confirm an enrollment, or a physician might be asked to send copies of medical records.

What to Measure

In responding to verification requests, third parties incur many of the same types of cost that government agencies do. Their personnel must spend time preparing the information and must use facilities and materials in the process. A request for an employment history, for example, might require the transmission of the request through an organization to the right staff member, a computer run to search for data, the transcription of the data onto an appropriate form, and mailing the response.

The efficiency and speed with which third parties process these requests has, of course, substantial bearing on overall cost. Since the efficiency and speed of third parties are likely to differ, what it costs them to fulfill identical requests is likely to vary. In one of OPM's match operations, for example, it was necessary to request death certificates from state bureaus of vital statistics. The fees the states charged for this service varied from nothing to around $5 per verification.

Approaches to Measurement

Recognizing that it costs third parties to provide data, some agencies pay them a fee to defray the expense for each request. The relationship between the amount of the fee and the cost of providing information should be assessed, in order to determine that the fee represents actual costs. As with other reimbursed costs, care must be exercised to avoid counting third-party costs twice when agencies reimburse all or part of these expenses. If such fees are not paid, determining third-party costs would entail a consideration of the same cost categories used for government agencies—salary and fringe benefits, materials, facilities, and travel—would be estimated for the activities a third party would have to perform to respond to the verification request.

Costs to the General Public

The potential encroachment on the constitutional and other legal rights and privileges of citizens is perhaps the most controversial aspect of computer matching. Automating and using personal data in a computer match have been viewed as an invasion of privacy, especially when individuals are likely to be unaware that files containing identifying information about them are being used. OMB's 1982 supplemental guidance for conducting matching programs provides that agencies must consider whether disclosing personal records for the purpose of matching would violate their responsibilities under the Privacy Act.[2]

Concern has also been expressed that potential clients for program benefits might be discouraged from seeking benefits to which they are legitimately entitled. The anxieties of the potential match verification process could inhibit them. Assessing this cost may be difficult, since the population is potential rather than current clients, but it should nonetheless be recognized.

Persons opposed to matching are concerned, on a broader scale, that the growth of matching may lead to the creation of large-scale data bases that may further erode such individual rights as freedom from unreasonable search, the presumption of innocence, and the due process of law. Moreover, the use of these data for purposes other than those for which they were originally collected is viewed as entailing a significant potential cost. The public response to loss of control over personal data and erosion of rights may be reluctance to provide personal information to others or even a tendency to provide it inaccurately. This type of cost may be too abstract to measure, particularly for a single computer

[2]Office of Management and Budget, "Privacy Act of 1974: Revised Supplemental Guidance for Conducting Matching Programs," 47 Fed. Reg. 21,656 (1982).

match, but an analysis of what this cost may entail is important. The net monetary value of matching should be weighed in relation to the concern about and the qualitative value of privacy and due process.

Assessing the public cost of a specific computer match would be extremely difficult and beyond the scope of most routine cost-benefit analyses. One approach—surveying the public about a particular computer match—is probably not feasible. Even with well-directed questions, responses may reflect views on computer matching in general, not the specific match in question. Moreover, the survey might influence the phenomenon it is attempting to measure, since it might make respondents more aware of concerns about matching.

Nonetheless, assessing general opinion about the value of privacy relative to the public benefits of matching may provide useful data about when and how matches are viewed as threatening or not threatening constitutional rights. The public may believe the cost of matching is greater when matches use data provided by individuals without their knowledge, individuals are inadequately notified that their records will be subjected to a match, raw hits are inadequately verified, and the requirements of demonstrating innocence are burdensome. Survey techniques that enable respondents to rank the relative merits of the costs and benefits of a match could be used to evaluate the importance of privacy and due process under varying match situations.

As we mentioned above, potential clients may be discouraged from applying for benefits to which they are legitimately entitled. In surveys of the public, respondents should be queried about this possibility. A sample of new applicants and current clients could also be queried, in order to see if they are reluctant or have second thoughts about participating in a program because of a matching operation. Some indirect indicators of the potential public cost might be based on data collected from the segments of the population to which a match has more personal relevance. The responses of innocent parties affected by a match and of individuals whose records are in matched files should be considered.

Since the cost to the public involves constitutional issues, another assessment approach might be to draw on the expert opinion of jurists about the implications a specific match has or matches in general have for actual or potential violations of the constitutional rights and privileges. Their opinions and analysis could help delineate situations that

pose greater or lesser threats to a person's constitutional and other rights.

As the scope of matching grows, information about larger numbers of citizens is likely to be processed in match operations. Even though the broad social costs of matching may not be easily quantified, a better understanding of them is needed. These approaches to the assessment of public cost represent some preliminary steps that could provide useful information in the development of this understanding.

Summary

In this chapter, we have identified and discussed the major cost categories associated with entities that may be involved in a computer-match operation. A complete accounting of computer-match costs should include a valuation of all resources spent on its activities. This requires that all personnel involved in the match be identified, as well as the time they spend on match activities. Facilities, materials, and travel costs spent on the match should be included. Many government agency costs are quantitative and can be put into dollar terms. Qualitative costs may also be incurred in the performance of a match. Some suggestions of possible ways of assessing these costs are provided, but an adequate approach to measurement needs considerable development.

Clients and third parties have some costs that are analogous to those of government agencies. However, the methods of collecting data on costs are usually different for clients and third parties, and it is difficult to establish a basis for valuing a client's time. There may also be a general cost to the public for computer matching, although this too is extremely difficult to measure. The cost has been characterized as an erosion of the right to privacy and other constitutional safeguards.

Estimating or Tracking the Benefits of a Computer Match

As in chapter 4, on costs, in this chapter we address the full range of benefits that have been credited to computer matching. We describe general methodological approaches to estimating benefits, since specific applications depend on the characteristics of a particular match.

The first step in assessing benefits is to identify their individual components. Not all the benefits in table 3.1 are likely to be obtained in a specific match, and some benefits may be weaker (and more difficult to measure) in some matches than in others. It seems extremely unlikely that all benefits from a match are of one kind; it also seems unlikely that all potential types of benefits are relevant to any one match. Therefore, identifying the benefits to include in an analysis is a key preliminary task in evaluating a computer match. Determining how to measure these benefits is another.

Identifying the benefits to include requires a consideration of such things as the properties of the program being matched, the characteristics of the population being served by the program, and the purpose and operation of the match. For instance, matches involving programs that provide no payments (a program to verify compliance with selective service registration, for example) cannot produce direct monetary benefits, nor can matches that identify past overpayments but do not lead to attempts at collection. Similarly, a match may have the potential for producing several different benefits, but its planners may not be interested in or may be constrained from pursuing them all.

Some benefits may be realized from a given match, even though they were not specifically intended or considered by its planners. For example, a match conducted to detect and recover retirement payments to deceased beneficiaries also identified inaccurate data in the files and demonstrated the need to crosscheck them. This led to improvements in record-keeping and the adoption of matching procedures for checking entries in the data system. Taken as a whole, these considerations reflect the need to review carefully the entire match operation and specify clearly the benefits that are appropriate for the analysis and the reasons they were included in it.

The second step in assessing benefits is to measure them. This requires an assessment of the identified benefits in terms of their quantifiability, the availability of relevant data, and the potential for collecting new data that may be needed for estimating them. Avoiding overpayments and recovering overpayments and debts are examples of benefits that

are readily quantified in monetary terms. Other benefits such as law enforcement and program integrity are inherently qualitative.

As with costs, it is important not to confuse the quantifiability of a benefit with its overall worth. Benefits that cannot be quantified for computing a net value may be important in other ways and should be considered along with net benefit amounts that have been quantified. Similarly, it is important not to confuse the potential for benefits with benefits that are actually realized. The former may be substantial, but only the latter should be included in computing net benefit, since some errors or cases in a match will never be resolved.

Benefits to Matching and Source Agencies and the Justice System

Many of the benefits commonly attributed to computer matches are associated with the matching agency, but the source agency may share the potential for them. The benefits from law enforcement and deterrence can be realized by agencies in the justice system as well. In the sections that follow, we discuss the benefits we believe accrue to government agencies from computer matches.

Avoiding Overpayments

What to Measure

One of the major benefits claimed for many computer-matching projects is the avoidance of future overpayments to program beneficiaries. In discussing this benefit, it is important to distinguish between the closely related savings from discontinuing and reducing grants and savings from disqualifying grant recipients. When it is found that a program's recipients receive more than is due them, the program can achieve savings by appropriately reducing or discontinuing their payments. Disqualification is a punitive procedure imposed for violating program rules, can occur regardless of whether a recalculation shows what would be due a recipient who would have otherwise been eligible for benefits, and reduces payments to zero for some fixed period of time.

The distinction between recipients who are disqualified and recipients who are not disqualified but whose payments are discontinued can be important in measuring the benefit amounts. Recipients whose payments are discontinued can rejoin the program rolls whenever they can demonstrate that they meet the program's requirements. The agency saves the amount of overpayment that these persons would have

received in the absence of the match. Recipients who are disqualified are eligible to return to the program rolls after the specified disqualification period has ended. The agency's savings can include the elimination of legitimate benefit payments for the duration of the disqualification.

Approaches to Measurement

Avoided overpayments can be difficult to measure. Clients identified in a hit have already been in the program for some time, and in the absence of the match, most of them would probably remain on its rolls for varying periods. Determining the length of this period is central to the accurate estimation of overpayment avoidance.

The first step in estimating the amount of avoided overpayments is to identify cases that are subject to benefit adjustments, discontinuances, and disqualifications and cases that should be excluded or deleted from these categories. For example, adjustments may not be made for cases receiving very small overpayments. Some cases may be appealed successfully, and others may be dropped from consideration for other reasons. For example, clients may be found eligible for the same level of benefits because of changes in family composition or income that offset the match-generated adjustment. Appropriate exclusions leave the cases whose payments are altered and for which the estimate of avoided overpayments should be computed. In a postmatch analysis, available records may allow a simple enumeration of these cases. In a prematch analysis, the number will have to be predicted.

The next step is to determine the overpayment that would have been received per month or some other interval in the absence of an adjustment. The amount may be the total benefit payment or some portion of it. However, it should be noted that the amount of overpayment being received at the time of the match may differ from future overpayments that would have been received in the absence of the match. Just as a correct payment can vary over time with changes in income, family composition, and assets, so can the amounts of an overpayment. One reasonable and expedient approach to addressing this problem is to set the amount of the monthly overpayment for each case at the monthly level of the discovered overpayment. Since fluctuations in overpayments over time will undoubtedly result in overstatements for some cases and understatements for others, the use of this procedure depends on the assumption that there is a tendency for errors to be offset when all cases are considered together. Using tolerance levels in the selection

of hits, however, may violate this assumption and introduce bias into the estimates.

Determining the length of time in which overpayments would be avoided is the next task. If payments are monthly, the main question is how many months overpayments would have continued without the computer match. Recipients might leave the program voluntarily or eliminate the reason for the overpayment by quitting a job or selling an asset, for example. Or an overpayment may be detected by some other means. One common practice is to apply some standard period, such as 1 month or 1 year, to all cases. Another practice is to use the average length of time that the recipients of a program's benefits remain on the rolls.

Neither method is likely to be adequate. The first is not based on experience and, depending on the time period selected, is as likely to underestimate benefits as to overestimate them. For example, a time period may be selected to coincide with the period for which related cost data are reported, such as a fiscal year, so that match costs and match savings will be projected for a comparable period. The second method may be based on actual program experience but may be inadequate, if the rate at which typical recipients enter and leave the program does not apply to the special subgroup of recipients of concern here—those who, in one way or another, have received overpayments from the program.

A combination of data on a standard time period and average case length could be used to derive a "conservative" estimate of the average length of time recipients receive overpayments. For example, 1 year may be viewed as conservative if the average recipient remains on the rolls for 18 months. One agency acknowledged the difficulty of determining this period and based cost-avoidance figures on three time periods—6 months, 1 year, and 2 years. This technique helped establish a potential range of time in which overpayments might be avoided but offers little guidance for selecting the most reasonable estimate.

A more sophisticated approach was developed in a recent postmatch study of the costs and benefits of matching wages in the AFDC and Food Stamp programs. The focus was on the discovery of periods of overpayment that were in process and halted earlier than they would otherwise have been. Data from the match showed the duration of these "interrupted" spells of overpayment but not, of course, how long they would

have lasted in the absence of the match. The study's authors used statistical techniques that had been developed to infer distributions of completed periods from data on interrupted durations in other contexts (such as in estimating the duration of unemployment and job tenure) in order to estimate the distribution of completed spells of overpayment for the first wage match conducted by a northeastern state. This information was used to infer the overpayments avoided because of this match and other matches assessed in the study.

Recovering Overpayments and Debts

What to Measure

The recovery of overpayments and the recovery of debts are benefits that are obtained in different situations, but many of the issues related to their measurement are the same. Overpayments can be recovered through matching when it is determined that individuals have been paid more than they should have. Debts can be recovered when it is determined that a loan recipient has defaulted on payments and the defaulted amount is subsequently collected. In the recovery of debts, what is in question is not the original benefit amount but that the terms of repayment have not been met.

Another difference is that the total amount to be recovered from loan repayment can frequently be determined without any investigative effort. The figure is simply the total amount in default and is likely to be readily available from payment records. In contrast, the total amount to be recovered from overpayments is more problematic. The computer match is likely to spot what seems to be an overpayment at a given time and identify the amount, but contact with employers, banks, or other private institutions will probably be required to verify the amount and determine the full period during which it has been paid. These data must then be used to calculate the total amount overpaid. It should be noted that when recoveries are tracked, the dollar amount of benefits in this category can be established with a high degree of certainty.

Approaches to Measurement

In measuring benefits from the recovery of overpayments and debts, it is important to distinguish between amounts that are identified in matching and amounts that are actually recovered. Often, only the identified amounts are reported, but it is the latter figure that constitutes

the actual monetary return from matching. The amount identified is only the potential upper limit of the benefit. This distinction was acknowledged to some extent in the report of a match of SSA beneficiaries with records of investment income from California financial institutions. The document reported overpayments in three categories: amount refunded or expected to be refunded by agreement, amount awaiting resolution, and amount waived as uncollectable.

Regardless of repayment agreements, it is unlikely that the full amount to be repaid will be repaid. Actual recoveries are likely to be a function of how arduous and extensive the effort is to collect from the clients in the program. The success of a collection and, thus, the extent to which actual recoveries reach the level of potential recoveries depends on factors such as those listed in table 5.1.

Table 5.1: Factors Potentially Affecting Overpayment and Debt Recovery

Factor	Description	Example
Client		
Ability to repay	May increase likelihood of repayment	Welfare recipients have limited funds and are less able to repay than most physicians or employed student-loan recipients
Ties with program	Repayment should be easier to obtain from persons receiving benefits than from those no longer associated with the program since they should be easier to locate and payment offsets can be applied	Student-loan recipients no longer in school and former Food Stamp recipients are often difficult to locate and benefits cannot be easily withheld
Ability to appeal overpayment or default	Persons with resources and inclination to appeal might be more likely than others to delay or avoid repayment partially or completely	Physicians who receive Medicaid overpayments are likely to have greater resources available for appeal than most recipients of need-based assistance
Program		
Administrative and legal	Stringent notification and appeals requirements are likely to increase costs, time, and difficulty of collection	Financial institutions were not required to return funds deposited to a deceased person's account more than 45 days after the death
Collection method	Ability to garnishee wages, offset tax refunds, or attach assets can increase probability of recovery, depending on cooperation of employers, other agencies, etc.	Wages can be garnisheed or tax refunds can be offset for civil servants and child-support evaders
Tolerance level	Tolerance levels function as dollar amounts above which a hit will be pursued; low tolerances should result in more hits, and potentially greater recovery amounts, than high tolerances; given limited follow-up resources, however, it may be more prudent to concentrate efforts on hits likely to result in the greatest recovery	Hits are prioritized in several state Food Stamp and AFDC income-verification matches, based on level of discrepancy between earnings reported by client and employer; hits with an earnings discrepancy above the tolerance level receive follow-up

Because many variables can affect the success of a collection, accurate prematch estimates of benefits in this category are difficult to make. A pilot match in which collection is undertaken on a representative sample

of all hits (or on a sample of all hits above a given dollar value) could provide a reasonably sound basis for estimating recovery benefits. A straightline projection of the collection rate for the total population would be appropriate, if collection efforts were sustained at the same level in the full match.

When prematch estimates are needed but cannot be obtained from a pilot match, experience with similar collections may be used as a guide. A previous, similar match may give a useful indication of the actual amounts likely to be recovered from a proposed match. Estimates of potential recoveries are reasonable to the extent that the client population to be matched and the level of collection effort to be implemented are like those of the previous match, on which the new estimates are being based.

One practice in the prematch estimation of overpayment and debt recovery uses the agency's current overall recovery rates. In a match of SSA beneficiaries with IRS wage data, for example, SSA used its overall recovery rate of about 60 percent. This practice can provide a reasonable basis for developing an estimate but only to the extent that debt collection following the match uses the agency's usual procedures and the characteristics of the population identified in the match are similar to those of other debtors to the agency.

In a postmatch assessment of recovery, the amounts of money actually collected can be determined, at least in principle, but doing this requires tracking repayments as they are received. Many programs have provisions for recording repayments, but most do not analyze match-related recoveries, because they cannot determine exactly which payments are attributable to a specific computer match. For example, the payments received from a match of loan defaulters in one agency were not known because the computerized tracking system does not currently include an indicator showing that repayment was initiated as a result of the match. The agency also could not identify whether repayment was voluntary after the match-initiated contact or the result of the imposition of a salary offset.

An indicator was placed on an individual's repayment record in another agency if the match was performed by the office of the inspector general. We found no indication, however, that these indicators are used to compute total repayments for individual matches. It appears that only aggregated repayment amounts for all audit-related projects are reported. Another agency identified repayments from individuals as the

result of a match but has not developed software to summarize the total amount received from all the individuals identified by the match.

When recovery files are not identified by source, one way of measuring match-related recoveries is to select a random sample of processed hits and search (manually or by computer) for their records in the files accumulating data on their repayments. If the sampling is appropriate, these data can be used to project a reasonable estimate of payments recovered in the entire match.

In postmatch assessments, the timing of the analysis is important. Many repayment situations extend over years rather than a few months. Processing hits and initiating collection procedures can considerably precede the time when collections are actually made. One match that identified more than 500 cases of overpayment yielded full recovery for only 50 cases a year after the match was run. Thus, an assessment of recovery benefits at the end of the first fiscal year of a match may understate its total recovery.

As the quality and quantity of information on overpayment and debt recoveries increase, appropriate analyses of postmatch data from previous matches may improve the ability to predict recovery benefits in prematch planning. For example, tracking payment patterns for various subgroups (by age, ability to pay, the length of time money has been owed, and so forth) could be useful in projecting repayment patterns for specific groups to be subjected to new matches that are comparable in purpose and scope. Additionally, assessing repayment patterns over time (how much is repaid in the first 6 months, the first year, and so forth) could result in greater precision in estimating the flow of repayment recovery in future matches. It should be noted that discounting, which we discussed in chapter 3 on costs, may be a critical consideration, because amounts owed may be recovered over a period of years. However, no matches we examined used discounting in evaluating recovery benefits.

Law Enforcement

Matches can detect various types of noncompliance. For example, those we discussed in the preceding section focused on payment errors from individuals' not being in compliance with rules, regulations, procedures, or laws governing the operation of loan and other benefit programs. Matches that do not involve money may be performed for the primary purpose of enforcing the law by detecting individuals who may have violated some aspect of a legal code. Examples are a match to detect

nonregistrants for the selective service system and a match to detect physicians unqualified to practice medicine.

While law enforcement entails some quantifiable, and even monetary, benefits, a potentially more important benefit is nonmonetary: the intrinsic value of law enforcement itself. When society enacts a statute or implements a regulation, it presumably places a positive value on compliance with the law or regulation. Therefore, bringing individuals into compliance has a value to society beyond any notion of monetary savings. Like the more qualitative costs discussed in chapter 4, the nonmonetary benefits of law enforcement should not be overlooked or considered less important than monetary benefits such as avoiding overpayments and recovering them. Indeed, the primary benefit of matches designed to detect nonfinancial noncompliance is compliance with the law.

Although the law-enforcement benefit can be characterized as primarily qualitative, some quantifiable measures may be used to assess or monitor various aspects of this benefit. One approach is to count the match hits that are referred for criminal investigation, the referrals that are subsequently presented to prosecutors, the referrals that are accepted by prosecutors, and the disposition of the cases accepted. Part of the law-enforcement benefit may be quantified in the form of monetary amounts collected from fines or penalties for noncompliance detected in a match, keeping in mind that the fines actually collected are likely to be somewhat less than the fines imposed.

In some cases, the law-enforcement benefit can be assessed by counting the number of cases or individuals brought into compliance by a match. For instance, the number of persons eligible for military duty who are brought into compliance with the registration requirements because they received a match-generated notice to register could be counted. This approach is particularly useful when comparing the effectiveness of various methods of detecting noncompliance in equivalent cases. In such situations, cases can be counted for each method, and the estimates of the costs associated with each method can be divided by the totals. The resulting ratios, which are really measures of the dollar cost of each success, can be directly compared.

Deterrence

A frequently touted benefit of computer matching is its ability to deter noncompliance. The concept of deterrence is based on an expectation that as the probability of detection is perceived to be high or the

severity of the penalty for wrongdoing is perceived to be sufficiently great, the option of compliance becomes increasingly attractive. If the computer match detects noncompliance and if the subsequent sanctions for this behavior are perceived to be substantial, then future noncompliant behavior may be curtailed. Deterrence because of computer matching is generally treated as a qualitative benefit. Anecdotes of changes in the behavior of clients have been presented to support this, but efforts to demonstrate the existence and magnitude of such changes have been limited.

Several approaches can be used to measure deterrence from a computer match. There are difficulties, however, in using these approaches, either individually or collectively, to establish a clear relationship between computer matching and a change in compliance behavior. In addition to reflecting possible changes in compliance behavior from the deterrent effect of the match, some potential measures may change in response to other factors not related to the match's deterrent effect. This increases the difficulty of differentiating the contribution of computer matching from other deterrence-oriented activities such as the implementation of error-prone profiling or the influence of extraneous program or environmental events. Because of these problems, the assessment of the deterrence should be undertaken in special studies, so that adequate research designs and multiple measurement approaches can be applied to the variety of methodological issues associated with the measurement of deterrence.

Detailed discussion of the methodological problems and study designs associated with these approaches is beyond the scope of this report. We can provide only a general description of them and identify some of the measurement concerns associated with their use. One approach to finding out whether a specific match has deterred noncompliance is to ask the population of interest questions about their behavior and what affects it. Although questioning respondents about noncompliant behavior is obviously problematic, using survey techniques such as the randomized response, which maintains the confidentiality of the answers, may identify cases of noncompliance. Tax compliance, for example, has been probed with some success by direct question and randomized response techniques as measured against nonsurvey indicators of actual noncompliance rates.

If appropriately selected survey respondents are generally unaware of the computer-matching activities, however, then these approaches cannot be used to establish the presence or absence of a deterrent. If the

respondents are aware of the match and report that it changed their behavior, then deterrence can be validated to some extent, if the survey structure, content, and procedures have been carefully developed to control for various types of response bias and recall errors.

A second approach is to measure changes in noncompliance over time. The basic notion is to collect data on the number of documented periods of noncompliance that begin during a specified interval such as each month. Conducted at monthly intervals, in this case, the match itself could be used to identify new instances of noncompliance beginning each month. Collecting this information at the beginning of the match and continuing to collect it for some time thereafter could lead to a monthly time series of newly begun instances of noncompliance. If the match were acting as a deterrent, the number of noncompliance periods that began each month might be observed to decline over time.

One important limitation of the approach outlined above is that even if instances of noncompliance appear to decline over time, one cannot be certain of the degree to which the match caused the decline. For example, changes in the program or the environment after the match was initiated might change the number of instances of noncompliance. If so, they may make it less clear that any changes in the instances of noncompliance are the result of the match's acting as a deterrent.

The problems described above cannot be eliminated, but methodological refinements might help control for some of them. One possible method is to expand the analysis in order to include a comparison population, or a control group, that should be subject to the same program or environmental influences as the client population. The two groups would have to be as comparable as possible, and the time-series information on new periods of noncompliance described above would have to be developed for both the "treatment" population and the "comparison" or control population. Developing this information would require subjecting both populations to the computer-match series but subjecting only the treatment group to the aspects of computer matching—the follow-up activities—that are likely to deter noncompliance, such as reducing or discontinuing program benefits and instituting prosecutions for noncompliance with program rules.

Verification for the "comparison" group would be conducted only to the point of reliably identifying new periods of noncompliance. Assuming both groups were subjected to the same program and environmental factors, deterrence could then be assessed by comparing the time-series

information for the two groups. If the match were deterring noncompliance, the time series should indicate that, over time, there were greater reductions in the instances of new cases of noncompliance within the treatment population than within the comparison population.

The usefulness of adding a comparison group to improve the measurement of deterrence depends not only on how similar the two groups are but also on avoiding, or overcoming, other problems that may compromise the utility of this group as a control. For example, if clients in the comparison group become aware of how the match is affecting the clients going through full verification and follow-up in the treatment group, they may change their own compliance behavior. In other words, deterrence may spill over into the control group.

Even when a control group is not possible, the basic approach can be refined by using the computer-match population as its own control. This would entail conducting the limited match operation suggested for the control group above enough times to obtain a baseline series on the incidence of new cases of noncompliance. Then the full match operation, with complete verification and follow-up activities (payment adjustments, investigations, prosecutions, and so on) would be implemented. Given a sufficient number of data points for the entire series, both baseline and full match operations, an ARIMA interrupted time-series analysis could be performed. One advantage of having more than one baseline data point and conducting this type of time-series analysis is the ability to adjust for trends or seasonal characteristics of noncompliance.

Many circumstances may not permit a deterrence study, either with or without a control group or baseline time series. However, a special study of deterrence might be feasible for some computer matches. For example, a new match that is anticipated to have a deterrent effect and that is being implemented in different locations over an extended period of time might allow the identification of a comparison group of clients or provide sufficient lead time to develop a baseline series of data points on the rate of noncompliance before the match is fully implemented.

The methodological difficulties associated with assessing deterrence make it unlikely that any one study will adequately demonstrate or estimate the existence and magnitude of deterrence from a computer match, but several methodologies producing a set of studies that show similar results might provide convincing evidence of the ability of computer matches to produce a deterrence benefit.

Management Improvements

Several types of management improvements can stem from a computer match. One is related to the streamlining of program operations that can result from the match process and can take several forms. First, the match can replace the less efficient manual collection and updating of data. Before computer matches were common for banks, welfare programs attempting to identify unreported bank accounts did so largely through the manual use of clearance forms used by banks. Computer matching can allow an agency to collect data from computer files, eliminating the need for staff to make contacts with clients. It is probably more efficient to collect data on taxpayers' interest income from computerized IRS records than from interviews with individual taxpayers, for example.

Another type of management benefit is the several ways a match may strengthen a program's internal controls. Matched data on the characteristics of noncompliant clients or high-risk populations can be used to develop error-prone profiling. For example, a Food Stamp research project used computer matching to determine whether it were possible to develop a reliable profile of household characteristics in order to identify groups of program participants more likely than others to report wages improperly. Similarly, in a match of civil service annuitants, data were analyzed to identify the age and other characteristics of persons whose deaths had not been reported to the agency. A match between marriage records and surviving spouses who were beneficiaries of Supplemental Security Income was conducted by the office of the inspector general at HHS, and this match data is to be used to develop profiles that would allow SSA to focus on ineligible recipients. Profiles like these can help screen incoming applicants in addition to detecting noncompliance among those whose names are already on the rolls.

Match results can be used for developing other matches. In one agency, verification activities performed for one match indicated a need to conduct recurring matches between the files the agency maintained internally, on two different programs, in order to ensure that information on clients was consistent across both programs.

Data from matches can lead to the inclusion of questions on applications or other data forms that may elicit information from applicants that would prevent the enrollment of those who do not qualify for program participation. One recommendation among several from a match of civil service beneficiaries was a proposal to require everyone holding certain types of joint accounts with annuitants to sign a statement saying that they would inform the agency of an annuitant's death. This would

improve the agency's legal position in prosecuting persons who used a deceased annuitant's payments without being entitled to them.

Matching can also improve the quality of program data. Matches can lead to the identification and correction of data management problems in the computer files of the matching and source agencies. In one match, an agency's lack of current information on individuals was revealed when the agency attempted to use its information on them for verification activities. Erroneous Social Security numbers have been identified in several matches. In one instance, a special project was set up to correct Social Security numbers as a result of the match.

An increase in staff morale might be considered a management benefit. Improvements in staff morale, if they occur, can lead to improvements in productivity. However, the relationship between morale and productivity can be extremely difficult to establish and measure empirically, as we discussed in chapter 4.

Collectively, various management improvements should be related to decreases in program error rates and general improvements in a program's integrity. Some approaches to measuring these essentially qualitative benefits may be possible. Depending on the nature of the specific improvements associated with a match, it may be possible to quantify some of them. For example, a match expected to streamline program operations might be assessed by measuring changes in the level of productivity or resources required to perform the tasks that the match addresses. This would require available data relevant to the administrative goal of the match, such as routine workload statistics on the number of client contacts the staff make. For a productivity measure, it would also be necessary to adjust for changes in the allocation of resources to perform these tasks. If the extent to which matching has made operations more efficient is to be measured, data must be available that document productivity or the resources required to do the tasks before the match or in its absence.

When matches result in changes to a program's internal controls, such as the adoption of error-prone profiles, or new application procedures, a measure of the match's contribution might be derived by assessing the number of ineligible enrollees or applicants detected because of the match. Counting the number of corrections brought about by the match can help quantify the management benefit of improvement in data files. An alternative, more global approach would be first to assume that the interrelated improvements in program operations, internal controls,

data quality, and staff morale should decrease program and administrative error rates and, then, to assess data on program error rates before and after the match.

Almost every measurement approach, however, will have a number of methodological problems associated with it. Almost all measures will be influenced by factors such as other changes in program operations that may help improve productivity or staff morale or other sources of information and analysis that help develop procedures for screening applicants or the like. Ineligible clients may have been identified for reasons that are not related solely to the match.

As in the measurement of the benefits from deterrence, it may be possible to overcome some of these problems with various research designs that control or adjust for the other factors that can influence the measures. Control groups and ARIMA time-series analysis discussed previously for assessing deterrence might be usefully applied in the measurement of management benefits. Measuring management improvements is generally beyond the scope of the routine cost-benefit assessment of computer matches and should be the topic of special study, in order to establish and quantify their magnitude.

Greater Public Confidence and Program Support

A computer match may increase public confidence in a program if the match is perceived as promoting the program's integrity. The public may disagree with the tenets of the program but not necessarily find fault with the integrity of its operation. To affect public confidence, however, a computer match must be known to the public. Many matches are announced in the Federal Register, but we found only a few instances in which details about matches and their results were widely publicized. If a specific match or its results are not known to the public, no direct association can be made between the match and changes in public confidence. For these reasons, it is probably not feasible to assess changes in public confidence as a result of a specific match.

However, the general question of whether computer matching can build public confidence in program operations could be probed with survey techniques. For example, if a survey's respondents indicated that they thought computer matching had a substantial part in ensuring a program's integrity, matching as a factor that improves public confidence might be confirmed.

Benefits to Clients

Computer-matching benefits are probably thought of as associated generally with government agencies, but clients can experience some benefits as well. We found that some matches had client benefits as a goal, but in most matches, client benefits were secondary to agency benefits. In determining whether to expect client benefits as a result of a specific match, it is particularly important to give careful consideration to such factors as the nature of the program, the purpose of the match, and the manner in which the match is operated.

We identified three types of client benefits from computer matches: the identification of underpayments, improvements in service delivery and resources, and a reduction of the stigma of participation. The identification and correction of underpayments can be quantified in straightforward monetary terms. The measurement of the two other benefits is more problematic. Because of the difficulty of assessing service-delivery improvements and the reduction of participation stigma, measurement approaches suggested for them should be viewed as not definitive but illustrative of some potential ways in which information about these benefits can be obtained.

Identifying Underpayments

Some matches have uncovered and led to the correction of underpayments to program clients. Detecting underpayments is seldom the direct goal of a match but may be an indirect outcome that could benefit some clients. It should be noted, however, that matching software programs that do not generate hits when a client's actual income is less than claimed (and, therefore, indicating possible underpayment) will not detect underpayments. The benefit is the dollar amount of the underpayment. It must be remembered, however, that the discovery of this underpayment is not sufficient. To become a benefit, it must be paid to the clients.

Improvements in Service Delivery and Resources

Some matches may produce an improvement in services and their delivery to clients. This may be a primary reason for conducting a match. For example, one agency that provides health services to clients conducted a match to confirm the credentials of physicians engaged in the program. Two key purposes of the match were to protect clients and to ensure the quality of services provided to them. For programs whose benefits are fixed, as in a student loan program, stopping incorrect payments may make existing funds available to a larger number of legitimate applicants. For programs whose benefits are determined by formula, terminating ineligible clients from the rolls may increase the

number of clients who can legitimately be served within the funding limits of a program.

Measures of the service-delivery benefit should be related to the purpose of a match. For example, a nonmonetary measure of the benefit of assessing the credentials of physicians associated with a health program could be to count cases in which verification and follow-up led to the removal of unlicensed physicians from the program. A reduction in the incidence of malpractice cases that could be attributed to the match might also be an indicator of the benefit.

When the purpose of a match is to streamline program operations that bear directly on client contacts, a benefit may be viewed as accruing to either the agency, as a management improvement, or the client, as an improvement in service delivery. For measuring cases such as these, some of the approaches we discussed for measuring management improvements might be applicable. However, care should be taken that a benefit that can be applied to either the agency or its clients is counted as a benefit only once.

Reduction of the Stigma of Participation

Stigma may be associated with receiving benefits from a program that has acquired a poor reputation for its recipients. Similarly, some clients may feel reluctant to participate in a program that is not relatively free of fraud and abuse. The benefit related to reducing stigma from program participation may entail removing from its rolls those who are not entitled to participate. This benefit is qualitative and involves feelings about a program. Although its clients could be queried about their feelings, it seems highly unlikely that these feelings could be traced to a specific computer match or to computer matching in general, because feelings about a program are probably the result of a multitude of factors.

Benefits to the General Public

The general public benefits when computer matching leads to improvements in the operation and administration of government programs. Their tax dollars buy more, and a program's efficiency is, at least conceptually, a benefit to them. From a measurement standpoint, one approach is to consider the benefits to the general public analogous to the agency benefits that affect efficiency. Avoiding and recovering overpayments are the most direct realization of program efficiency in the management of public funds, but other benefits such as management improvements, law enforcement, and deterrence can also be beneficial.

From the perspective of the general public, an assessment of improved program efficiency rests on the net effect of these benefits. Therefore, measures of the directly quantifiable benefits to government agencies can also be used as specific indicators of the benefit to the public. In using these measures (discussed in the preceding sections of this chapter), care must be taken not to count benefits twice, once in relation to the agency and again in relation to the public.

Summary

The benefits of computer matching can accrue to agencies, their clients, and the general public, and they range from actual dollar savings for a program to improvements in attitudes about program participation. Their measurement ranges from what is readily quantifiable to what is quite qualitative in nature. In terms of dollars, the two major benefits are the avoidance of overpayments and the recovery of overpayments and debts. Measuring overpayments avoided always requires an estimation, predicting what would have been erroneously paid in the absence of computer matching. Estimates based on information about clients, especially those who were overpaid, are better than estimates that rely on assumptions about fixed time periods and overpayment amounts. The recovery of overpayments and debt can be measured by tracking the amounts as they are received. This monetary benefit of computer matching seems to offer the best measurement possibility.

Measuring deterrence as a benefit requires special effort but may be possible under some circumstances with a carefully planned study. Some aspects of agency benefits from management improvements are also quantifiable. When a computer match supplants a manual activity, for example, then measures of the resources required to perform the activity by the two methods might allow a monetary estimate of the management improvement. Similar to overpayment and debt recovery, the benefit to clients of detecting underpayments can, if tracked, be measured precisely in monetary terms.

Agency Cost-benefit Analyses of Computer Matching

The scope of our work included a review of the cost-benefit information available from the 17 agency computer-match operations listed in appendix III. We selected most of these match operations from preliminary indications that either match costs or match benefits or both had been assessed. Approximately two thirds of the selected operations encompassed more than one match and had additional matches planned for the future. Some of the single-match operations had not reached the point at which concurrent or postmatch analyses could be made. We included single and multiple match series in progress in order to obtain information on current practices and procedures for estimating and tracking costs and benefits.

We requested from the agency officials we interviewed any available documentation on match costs and benefits. We did not attempt to assess the costs and benefits of the 17 match operations independently but, rather, used our assessment criteria as a guide for reviewing whatever analyses and information that the agencies had developed and were available (see appendix II). We did not restrict our review to formal cost-benefit analyses but included whatever information had been provided to us on costs and benefits. We did, however, restrict our review to available information and did not pursue cost-benefit information in storage.[1]

Our selection of matches should not be considered a representative sample of agency computer-match operations, because we intentionally included matches that would offer cost-benefit methods for examination. It should be noted also that agencies were not formally required to perform cost-benefit analyses for these matches. Our review is not intended to be a critique of agency analyses; rather, it indicates the extent to which the analyses we reviewed considered the entities and cost-benefit elements we discussed in the previous chapters.

From our discussions with agency officials and our review of the materials provided to us, it appears that the consideration or assessment of costs and benefits for these matches varied considerably in nature and timing. The assessments ranged from informal, undocumented reports of matches for which costs and benefits had been considered at various points in decision-making to special studies of match costs and benefits performed by contractors. Formal cost-benefit analyses were available

[1]We did not encounter any situation in which stored papers were definitely known to contain additional, detailed information on match costs and benefits (that is, no one we interviewed was certain that stored papers would provide information not otherwise available).

for only a few of the operations. Reports of informal analyses for which no supporting documentation could be obtained are not included in our discussion.

Reporting Cost-benefit Analyses

We were able to obtain information on some costs or benefits for all the match operations we reviewed, but its availability varied considerably. We obtained prematch cost and benefit estimates for 12 of the 17 operations; for 3 operations, no information was available on either costs or benefits. The 2 remaining operations were divided: only costs were available for one and only benefits were available for the other. (See table 6.1.)

Table 6.1: The Cost-benefit Information Available for 17 Match Operations

	Type of analysis	
Available information	**Prematch**	**Concurrent[a]**
Costs	1	0
Benefits	1	4
Costs and benefits	12	9
Neither costs nor benefits	3	0

[a]Excludes 4 match operations that had not progressed to the point at which concurrent analysis had been produced.

More than half the operations had both cost and benefit information available for concurrent analyses. Except for one cost analysis (discussed later), we found no comparative analysis of prematch estimates with cost or benefit figures from concurrent analyses.

Three operations did not report both costs and benefits in a single document. One of these had information on costs and benefits, but it was not integrated into a single report. Of the 14 match operations that did produce a document containing both cost and benefit information, 5 had both prematch and concurrent analyses that presented both costs and benefits, 6 had only prematch analyses, and 3 had only concurrent analyses.

All the prematch analyses had one or more of the following problems: incomplete cost accounting, staff resources estimated in terms of time rather than money, and benefits estimated in terms of maximum potential return unadjusted for actual recovery. Five of the 8 concurrent analyses provided figures on major verification and follow-up costs, and 2 of the 5 also included overpayment collection costs. Four of the 8 concurrent analyses either adjusted for or provided some indication of actual

recovery rates when reporting benefits in terms of total overpayments. One analysis that reported actual recoveries and one of the prematch analyses were formal cost-benefit studies performed by contractors. The documents for these two studies provided the most detailed accounts of methodology and data sources.

Agency Cost Analyses

Taking the broad perspective to cost-benefit analysis discussed in the previous chapters, we found no instance of a complete cost analysis for match operations. When estimated or actual cost figures for a match were provided, they were usually limited to costs for automated data processing and a partial accounting, expressed in terms of staff time, of the staff resources needed to initiate or conduct a match. They seldom included staff resources from all the units (such as a collection unit or an appeals unit) participating in match activities. Costs incurred by source agencies, third parties, and clients were seldom reported except in some cases for charges to the matching agency by a source agency for data or automated data processing.

Table 6.2 presents an overview of the cost reports available for the 17 match operations we reviewed. Combining categories 2, 4, and 5 in the table shows that prematch estimates of some of the resources needed for a match were available for 13 of the 17 operations. Combining categories 3 and 4 shows that there were some cost figures for 8 of the 13 match operations that had progressed to the point at which a sample of hits could be assessed for a concurrent cost analysis. Three match operations were old enough for some type of final report to have been developed; however, only 1 of these operations provided figures on the actual versus estimated costs to the agency. The postmatch figures for this one operation were developed as part of an independent audit within the agency, not as part of the match assessment. We found no other reports in which prematch estimates of cost were compared with later assessments of actual costs. However, 2 ongoing match operations in our sample have formal plans for performing this type of analysis.

Table 6.2: Cost Reporting for 17 Computer Matches[a]

Category	Cost reporting	Number of match operations	Maximum number of match operations in category[b]
1	None	2	17
2	Prematch only	6	17
3	Concurrent only	2	13
4	Prematch and concurrent	6	13
5	Prematch, concurrent, and postmatch	1	3

[a]Multiple match operations were coded as having a prematch or concurrent analysis even if the analysis applied to only one of the matches in the series. Because postmatch analyses are cumulative, multiple match operations were coded as having postmatch analysis only if it covered the entire match series.

[b]Figures in this column represent the number of match operations that have been performed long enough for each type of analysis to have been completed.

No prematch report provided a range of estimated costs, although one analysis did provide various cost figures associated with different procedures for verification and follow-up activities. One concurrent report provided a range of high, low, and "best" estimates of match costs for the period covered by the analysis.

For operations consisting of a series of matches, routine updates of costs were generally not prepared. Tracking systems for some of these operations regularly produced printouts showing staff time charged to a match, but the charges are not translated into monetary figures and combined with other match costs to develop an overall expenditure rate. Further, not all agency staff involved in matching activities were always included in the tracking systems. In some instances, reports were prepared on the costs of specific matches in a series; agencies identified these as final reports. They provided some information about specific matches in a series but did little to illuminate costs for the whole series, since costs may vary from match to match. These reports are more indicative of a concurrent analysis when viewed in the context of a continuing match series.

For all the cost analyses we reviewed, regardless of when they were performed, figures were usually limited to automated data processing costs and staff days needed to perform match activities. Less than half of the match operations included some estimation of staff requirements for verification and follow-up activities by program field personnel. One analysis included the costs of administrative hearings and prosecutions associated with match hits.

Assessment or discussion of actual or potential costs to third parties or clients was not provided in any concurrent or final cost report we reviewed. In a few instances, reports sent to OMB on planned matches did briefly address potential risks to clients in terms of exercising their rights and privileges.

Agency Benefit Analyses

Benefits from the match operations were assessed more often than costs. Thirteen of the 17 match operations (categories 1 and 3 in table 6.3) had prematch reports addressing the potential benefits of the match. All the match operations (categories 2 and 3) that had progressed to the point at which a concurrent benefit analysis was possible had some figures of match benefits available. None of the 3 matches that had been in operation long enough to produce some form of final report had done so. These 3 matches had concurrent analyses of match benefits, but updated reports of the benefits of actual recoveries were not available; benefit updates were not routine.

Table 6.3: Benefit Reporting for 17 Computer Matches[a]

Category	Benefit reporting	Number of match operations	Maximum number of match operations in category[b]
1	Prematch only	4	17
2	Concurrent only	4	13
3	Prematch and concurrent	9	13
4	Postmatch reporting	0	3

[a]Multiple match operations were coded as having a prematch or concurrent analysis even if the analysis applied to only one of the matches in the series. Because postmatch analyses are cumulative, multiple match operations were coded as having postmatch analysis only if it covered the entire match series.

[b]Figures in this column represent the number of match operations that have been performed long enough for each type of analysis to have been completed.

In general, only benefits to the matching agency were addressed. Benefits to source agencies participating in the match, the justice system, third parties, and the general public were not identified or discussed in most of the analyses we reviewed. In a few instances, benefits to one of these groups were identified, but actual measures of the benefit were not provided.

Nonmonetary Benefits

The approaches to addressing benefits such as management improvement, law enforcement, and deterrence were fairly consistent across all

match operations. For most prematch analyses, nonmonetary benefits were identified simply as a potential outcome of a match. For most concurrent and postmatch analyses, the benefits of management improvement were most often described in terms of program vulnerability or a lack of internal controls contributing to the problems detected by the match. Recommendations on how to correct these problems were also included. In some instances, data were included on the magnitude of the problems a match identified—for example, the number of file records with incorrect Social Security numbers—but there were no instances in which an analysis attempted to quantify the nature or extent of the management improvements associated with the match.

The benefits of law enforcement were indicated by noting that cases had been referred for investigation or prosecution. However, only a few benefit analyses included actual figures on the number of cases referred to investigators. One analysis contained figures on the referral of cases to prosecutors and the disposition of these cases.

Several analyses identified deterrence as a potential benefit of a match, but few gave any information on its nature or magnitude. However, one analysis provided an estimate of the magnitude of the effect of deterrence, and another reported an increase in the number of persons who voluntarily reported changes in income after the match. This was seen as an indicator of deterrence.

Monetary Benefits: Recovering Overpayments and Debt

Only 3 operations had prematch analyses with figures on the expected magnitude of the recovery of overpayments and debts generated specifically for a proposed match. (In 1 match, the benefit of recovering overpayments or debts was not expected; therefore, our discussion of this benefit is based on 16 match operations.) A few match operations provided an estimate of the total overpayments or debts that might be identified, but they performed no adjustments to reflect what the actual recovery rate might be. In a few instances, factors that would affect the realization of an estimated benefit were discussed. However, the majority did not attempt prematch estimates. Of the 3 prematch analyses that provided figures adjusted for actual recovery rates, 1 specifically identified the rate that was used for the adjustment and its rationale.

Concurrent analyses of the recovery benefit were made for 13 match operations. Eight of these provided figures on total overpayment or debt identified without adjusting for the actual recovery rates. For the 5

remaining operations, maximum potential benefits were estimated, and actual rates of recovery were indicated or adjusted for. Two of these 5 regularly updated their reports of money recovered in their monitoring processes. Another analysis tracked the level of recovery for the period covered in the assessment. Of the 2 remaining analyses, 1 based an estimate on a sample of recovery actions and the other provided a recovery rate but gave no indication of how it was determined. Two analyses provided a range in which the estimated maximum and actual recovery amounts were expected to fall. No analysis applied a discount rate to situations in which repayments would be made over several years.

Monetary Benefits: Avoiding Overpayments

The benefit of avoiding overpayments was not expected for 5 match operations (for example, matches to identify loan defaulters; therefore, our discussion of this benefit is based on 12 operations, 3 of which had not progressed to the point at which a concurrent analysis could be performed). Three prematch and 4 concurrent analyses of match operations included estimates of the overpayments that could be avoided by a match. One prematch analysis used data on rate of annual case turnover and on payment error from other sources to estimate the annual overpayment that would be avoided by the match and reported the rationale and computation method for this figure.

The 2 other prematch analyses and 2 of the 4 concurrent analyses adopted the fixed period of 1 year and based a projection of the amount of overpayments to be avoided from the number of verified hits detected by the match and, presumably, the average annual program payment. Both prematch assessments used data from previous matches to estimate the number of hits. For all 4 analyses, the figures on overpayment avoidance were accompanied by little detail on how they were computed, although 1 prematch analysis described some of the assumptions that were used in developing the figures.

The 2 remaining concurrent analyses used information about the actual distribution of previous overpayments as a basis for estimating the future overpayments that would be avoided. One of these analyses developed and used a sophisticated statistical modeling procedure to estimate the overpayment amount to be avoided in each case. The other analysis computed average overpayment amounts and length of overpayment periods detected by the match and used the overpayment figures to produce a range of projections for 6 months, 1 year, and 2 years. The average length of time in which overpayments were made

was bracketed by the 1-year and 2-year projections. Both analyses included caveats on interpreting these figures.

Summary

We attempted to identify computer matches that provided information on their costs and benefits and on how they were measured. We reviewed cost-benefit information on 17 computer-match operations. The assessments varied considerably in when and how they were performed. Most of the reports were incomplete or not reported in monetary terms or both. Costs and benefits to clients and third parties were seldom addressed.

Analyses of match benefits were more frequent than analyses of costs, but measures of some benefits were lacking and other benefits were not adequately presented in several reports. Some reports did not contain both cost and benefit information, or the information was not presented in a way that allowed a comparison of costs and benefits. The measurement of such nonmonetary benefits as deterrence was generally not attempted. The benefit of recovering overpayments and debt was often presented in terms of the maximum potential amount that might be collected but without an acknowledgment of or adjustment for money that might actually be recovered. With one or two exceptions, estimates of the overpayment-avoidance benefit were presented with little or no description of the computation or its rationale or underlying assumptions. None of the match analyses utilized discounting procedures. Descriptions of methodology are quite limited in most of the reports.

These limitations indicate that improvements are needed in the assessment of computer-match costs and benefits. A need to assess computer-match costs and benefits had been expressed, but we found little evidence suggesting that a cumulative, systematic, and accessible body of cost-benefit data has been developed on the match operations we reviewed.

Improving Cost-benefit Analyses of Computer Matches

We developed and used a set of working criteria to serve as a conceptual framework for our assessments of computer-match cost-benefit analyses and the related literature (see appendix II). Our experience in applying the criteria to the 17 match operations discussed in chapter 6 led us to modify four of these criteria—completeness, technical adequacy, validation, and full reporting—in order to provide a structure for addressing problems in current practices for assessing computer-match costs and benefits. We have expressed the criteria as general objectives that should be sought in conducting and evaluating cost-benefit analyses of computer-match operations. Each criterion is discussed in relation to the areas in which analyses could be improved.

Completeness, technical adequacy, and validity are interdependent and represent a progressively better methodological base for performing cost-benefit analyses. The fourth criterion, full reporting, helps establish a useable data base on computer-match costs and benefits. In discussing these criteria, we make general suggestions for practices, procedures, and studies. Many of these suggestions have been noted in previous chapters, some represent practices already in use, and the others are new or would refine current practices.

Completeness

Cost-benefit analyses of computer matches require planning and organization. In most cases, the planning for an analysis should be an integral part of the planning for the match itself and should remain integral to the match as it progresses. This requires careful consideration of the match process and the individuals and organizations that it is likely to affect. Conceptual issues and the formulation of data collection and analysis techniques deserve attention. It should be remembered that costs are incurred during all phases of a match and that benefits occur only after the match has been conducted. After the match has been conducted, the analysis should be concluded with an assessment of actual costs and benefits and their implications for future matches.

All costs and all benefits of a computer match should be considered, both those that affect the agency conducting the match and those that affect other individuals and organizations. In table 3.1, we listed the key entities that should be considered in identifying the costs and benefits and gave examples of the cost-benefit elements for each of these entities. Not all costs and benefits are quantifiable, and even those that are may not be measurable in monetary terms. When costs and benefits are quantifiable, it may be unreasonable to incur the costs associated with measurement and analysis for some elements, given the scope and

importance of the match operation being assessed. Nevertheless, such costs and benefits should be not ignored but considered along with those that can be reasonably quantified. Attempts to quantify some of them should be encouraged.

In postmatch analyses, unintended costs and benefits should be considered along with those that were intended. Matches conducted in order to achieve qualitative benefits such as improvements in program controls should still track the costs and benefits associated with the hits that receive verification and follow-up.

Technical Adequacy

In addition to being complete, the measurement of costs and benefits should be technically adequate. This criterion addresses the need to improve agency procedures and practices so that sources of error and bias can be reduced or at least recognized and so that results can be interpreted with an understanding of their limitations. Selecting inappropriately constructed measures and poor data collection practices are two areas in which problems of technical inadequacy can arise.

Measures of costs and benefits should rely on sources of data and data collection procedures that are reliable and amenable to independent corroboration. To the extent possible, sources of bias and error should be recognized and controlled. Lack of adequate tracking systems for obtaining measures of such costs as the time personnel spend in verification and follow-up activities and such benefits as debt recovery increase the possibility that estimates will be unreliable and biased. Although some cost and benefit elements are not amenable to precise measurement, others are. Elements that cannot be measured and included in a cost-benefit ratio still require appropriate consideration.

Collecting data on the costs and benefits associated with each match hit may not be required if appropriate sampling procedures are employed. Collecting data on both the personnel costs related to a match and the recovery of overpayments—two elements that are likely to influence the cost-benefit ratios significantly—may be accomplished through sampling procedures that provide reasonable estimates while reducing the data collection and analysis efforts. For instance, an estimate of overpayment recovery that is derived from a random sample of hits should provide an adequate estimate of recovery for the match as a whole, if the collection procedures for the sample and the total match are consistent. However, if the sample is not random, drawn from a subpopulation of hits that are felt likely to be most fruitful, then extrapolating results

to the entire population would be unfounded and technically inappropriate.

When the stream of costs and benefits of a computer match spans years, discounting should be employed to keep the various cost and benefit elements in comparable dollars. Since costs precede benefits, the failure to discount the benefits may result in their overestimation relative to the costs of the match.

Technical adequacy is related to the point in a match at which match analysis is performed. For example, basing estimates on similar previous match experience may be considered technically adequate before the match is conducted, but after the match, when actual data should be available, basing estimates on other matches would not be appropriate. Even in the prematch situation, considerations of technical adequacy should guide the development of estimates. For example, an overall agency recovery rate may be less appropriate as a basis for projecting repayments if data on actual recovery rates from previous matches are available, especially if the matching agency performed the matches with the same client population as the proposed match.

Given the difficulties of estimating computer-match costs and benefits, especially in prematch analysis, a single value for a projected cost or benefit element may not convey the imprecision associated with their measurement. Some analyses might be improved in technical adequacy by development of a best estimate for a particular cost or benefit and a range into which the actual value is expected to fall. The usefulness of this approach rests on the existence of some reasonable basis for determining the two ends of the range.

Validation

The general issue addressed by this objective is the concern that the cost-benefit analysis accurately reflects the true costs and benefits of the match. Validity is compromised to the extent that measures of costs and benefits rest on inappropriate or incorrect assumptions. "Validation," in this context, refers to an examination of whether the measures employed in the methodology of a cost-benefit analysis are meaningful—whether they measure what they purport to measure and whether the methodology is based on correct assumptions.

We have identified two areas of concern in the validation of cost-benefit analyses. The first is the lack of information on the relationship between prematch analyses and concurrent or postmatch analyses. Knowing this

relationship helps test the performance of the estimation procedures and the validity of their assumptions and measurements.

For example, basing a projected recovery rate on an overall agency rate or on actual recovery rates from previous matches rests on the assumption of similar repayment behavior in the match population to the behavior of other agency debtors. Comparing data from before and after a match should provide the information necessary to assess this assumption. Actual match recovery rates could be analyzed as a function of client characteristics and collection efforts, and doing so could provide a more solid foundation for estimating likely recovery rates for subsequent matches.

Prematch analyses are not likely to improve until sufficient information has been collected and reviewed to determine what techniques and procedures seem to work best. When prematch estimates of costs and benefits are developed, a concurrent or retrospective analysis should also be performed after a match has been implemented. Differences between the two sets of figures should be studied for improvements that might be made in the estimation methods. This prematch validation effort should be documented in order to facilitate the development of a cumulative body of knowledge on estimation methods.

The second area of concern is the lack of evidence and the quality of evidence used to support methods of estimating specific match benefits. For example, techniques for estimating the benefit of avoiding overpayments rest on assumptions whose validity has not been specifically assessed in the documentation we were able to obtain. As we noted in chapter 5, estimating this benefit requires a determination of how long an overpaid client would have remained on the program rolls in the absence of the match. Several different techniques of varying sophistication have been used to project this figure, which, at some point, should be validated against clients' actual behavior.

In the matches we examined, assessments of the deterrence benefit, for example, were for the most part based on anecdotal evidence of its existence. Increases in the incidence of voluntary withdrawal from a program and client-initiated reports of change in income or assets were noted as indicators of the deterrence in a few of the reports we examined. However, figures on the frequency of this behavior were seldom provided. Baseline data from which to gauge the magnitude of an increase in deterrence against normal program fluctuations were also lacking.

The complexity of the methodological problems associated with these validation issues precludes the development of routine procedures and techniques for their assessment in most match operations. However, agencies that engage in computer-matching operations could conduct special studies in these areas. Studies using experimental designs might assess the adequacy of the assumptions and methods used to estimate the benefit of avoiding overpayments. This would require collecting data on such things as how long overpaid clients remain on a program's rolls and the rate at which overpaid clients who are dropped from rolls as a result of a match return to the program.

A variety of experimental and quasiexperimental research designs could be used in a more methodologically rigorous attempt to study deterrence. Despite the considerable methodological problems we noted in chapter 5, collecting time-series data on new cases of noncompliance can help detect and assess changes associated with the implementation of a computer match. Systematically identifying and counting actions cited as anecdotal evidence of deterrence might be considered as a rudimentary means of quantifying this benefit, but analyses using such measures would have to include baseline data. Methodologically difficult as the measurement of deterrence may be, a better understanding of how computer matching can help prevent fraud and abuse depends on attempts to quantify this benefit.

A third area of concern is the cost of the possible invasion of privacy by computer matching. Special studies could usefully address assumptions about the true costs and benefits of computer matching for the public. Their relevance to computer-match operations is controversial and needs to be elucidated. Proponents of matching have stated that privacy safeguards are already adequate. Opponents have argued that they are inadequate.

As we discussed in chapter 4, none of the assessment approaches that have been identified for measuring the cost of the invasion of privacy could be routinely employed in most match cost-benefit analyses, and the results from any one approach could be not be considered definitive in any sense. However, special studies using different approaches could offer a general assessment of the nature and importance of this broad social cost. In general, the performance of the various special studies in all these areas will not be able to provide conclusive answers about the validity of computer-match cost-benefit analyses. They should, however, indicate methods that should be modified or replaced.

Full Reporting

For the match operations we reviewed, the reports of results varied considerably in their provision of cost-benefit information. It was often not clear how extensive the analysis was or how it was performed. Detailed information about a match's costs and benefits was often not readily available, except for summary reports, which were tersely written.

Just as an analysis should be complete, so its reports should be comprehensive enough to meet the following criteria. Descriptions of the objective, design, scope, and perspective of the cost-benefit analysis should be included. The methods used in the analysis (such as sampling procedures) should be described, and the underlying assumptions and rationale should be stated. A summary report should be linked to more detailed appendixes or supplemental reports. The costs of the match should be sufficiently disaggregated to identify the significant sources of match costs. For both costs and benefits, there should be an explanation of how the figures were derived.

Full reporting also entails planning for a report's distribution and obtaining feedback on its use. Improving the performance of a cost-benefit analysis is of little value if the knowledge is not shared. One improvement in match plans would be to include a specification of how and to whom—not only officials within the matching agency but also staff members of other participating agencies—the results of the analysis will be disseminated. Agencies that share client populations might benefit from sharing match results. The Long-Term Computer-Matching Project of the President's Council on Integrity and Efficiency has established several mechanisms that would help notify and inform others about the costs and benefits of a match.

Efforts to follow up on how reports are used may provide useful information on how cost-benefit analyses can be improved. Such efforts might assess the relevance of analysis, including data and methods, to decisionmakers and its effect on current matches or the development of new matches.

One value of full reporting is its usefulness to decisionmakers and match operators for refining and directing current matches and planning new ones. Full reporting of concurrent match analyses and postmatch analyses of a pilot match is necessary for making informed decisions about whether to continue, modify, or curtail a match operation. Insufficient information about its full costs may result in inappropriate decisions that compromise its overall efficacy. Administrators removed from the

day-to-day operations of a match may depend more heavily on reports than on briefings and informal contacts with the staff who run it.

Prematch estimation and analysis will improve only if there is an extensive and accessible body of information on previous matches upon which to base the analysis. It is through drawing analogies with prior matches that more reasonable estimates and analyses of planned matches can be made. Improvements in the ability to predict recovery rates, the cost of investigations, the benefits of deterrence, and the like can lead to more efficient operations. Without this gain in information, facilitated by good reporting, one of the potential benefits of current matches, that of helping to build more efficient operations in the future, is lost.

Summary

The performance and reporting of computer-match cost-benefit analyses can be improved. The criteria of completeness, technical adequacy, and validation reflect a sequential dependence that would constitute increasingly better analyses as each criterion is met. A prerequisite for the technical adequacy of a computer-match cost-benefit analysis is completeness. Similarly, a valid analysis is both complete and technically adequate. It should be added that satisfying a prerequisite does not ensure that the next criterion will be met. A complete analysis is not necessarily technically adequate, and a complete and technically adequate analysis is not necessarily valid.

The techniques and approaches to be used in an analysis can be assessed in relation to each criterion. The methods that appear best able to meet the criteria should be adopted. Given the difficulties of assessing computer-match costs and benefits, full reporting of an analysis is crucial if a cumulative body of information is to be developed on better ways to assess computer-match costs and benefits.

Request Letter

TED WEISS, N.Y., CHAIRMAN
JOHN CONYERS, JR., MICH.
SANDER M. LEVIN, MICH.
BUDDY MacKAY, FLA.
EDOLPHUS TOWNS, N.Y.
BEN ERDREICH, ALA.

ROBERT S. WALKER, PA.
ALFRED A. (AL) McCANDLESS, CALIF.
LARRY E. CRAIG, IDAHO

NINETY-EIGHTH CONGRESS

Congress of the United States

House of Representatives

INTERGOVERNMENTAL RELATIONS AND
HUMAN RESOURCES SUBCOMMITTEE
OF THE
COMMITTEE ON GOVERNMENT OPERATIONS
RAYBURN HOUSE OFFICE BUILDING, ROOM B-372
WASHINGTON, D.C. 20515
(202) 225-3648

March 15, 1984

The Honorable Charles A. Bowsher
Comptroller General of the United States
General Accounting Office
441 G Street, N.W.
Washington, D.C. 20548

Dear Mr. Bowsher:

The use of computer matching by Federal and state agencies to detect fraud, waste, or mismanagement in government programs has increased dramatically in the last few years. Agency Inspector Generals have been very active in promoting computer matching by Federal, as well as by state agencies.

Ever since it began, computer matching has been controversial. Critics have argued that matching constitutes an unreasonable search, violates personal privacy, and may not be cost-effective. The legal and constitutional questions surrounding computer matching are not likely to be resolved without litigation. However, I believe that it is possible to resolve some of the factual questions about the cost-effectiveness of matching. I would like to enlist the aid of the General Accounting Office in this task.

Evidence presented to the Senate Government Affairs Committee by the New York Civil Liberties Union and others suggests that at least some computer matches may not be cost-effective when all costs are taken into account and when realistic evaluations are made of the benefits of matching. Prior to May 1982, the OMB Privacy Act matching guidelines called upon agencies to determine prior to conducting a computer match whether a "demonstrable financial benefit" can be realized. This requirement of a cost-benefit analysis was dropped in 1982.

There are no generally accepted rules by which the costs and benefits of computer matching can be measured. It is possible that those who conduct matching programs overstate the benefits and ignore many of the costs. I request that the General Accounting Office develop a methodology that will permit an evaluation of the cost-effectiveness of a computer match. This will require an identification of all relevant costs, including the cost of preparing records for matching, computer time, investigations of cases identified by the computer, overhead, and other costs. If possible, a method of estimating the cost to individuals who become the subject of investigation should also be developed. Where matching is a cooperative venture among different levels of government, the costs of each participant should be included.

Please note that I am not asking GAO to make a determination of the cost-effectiveness of computer matching. I do not believe that such a determination would be meaningful at this time. What I want is a methodology for evaluating cost-effectiveness that could be used both before and after conducting computer matches.

As a second part of this project, I would like the General Accounting Office to select a sample of computer matches conducted by Federal agency Inspector Generals and determine how the decisions to conduct the matches were made. The goal is to identify and evaluate the criteria used when deciding whether a particular computer match should be conducted. The purpose is to determine if decisions to conduct computer matches are based on appropriate criteria. The focus of this work should be on the reasons that supported the decision to conduct a match rather than on any justifications that were developed after the matching operations were complete.

Because of the importance of this subject, I ask that my request be given the highest priority and that work begin as soon as possible. If you have any questions, please contact Susan Steinmetz of the subcommittee staff.

Thank you.

Sincerely,

TED WEISS
Chairman

Conceptual Criteria for Reviewing Computer-Match Cost-benefit Analyses

Completeness	Does the cost-benefit analysis report adequately describe the objectives, design, scope, and perspective of the study; its assumptions and their rationale; the resources and time needed to perform the study; and the costs and benefits included and not included in the analysis?
Verifiability	Did the cost-benefit analysis have adequate supporting documentation? Can parts of the study be independently corroborated? Is the information provided sufficient to permit a check or recomputation of figures under the same or other assumptions?
Technical Adequacy	Did the analysis address the objectives of the study? Were the study methods and procedures selected and applied appropriately? Are the data that were collected reliable and appropriate? Are the findings and recommendations supported by the analysis? Is the report well organized, logical, and internally consistent? Were the measures or procedures for estimating costs and benefits appropriate? Are significant or quantifiable costs or benefits not reported or not acknowledged?
Validation	To what extent were prematch analyses followed up with postmatch results? To what extent were interim analyses updated? To what extent has the analysis been replicated? Have the results been discussed?
Utilization	Is there a plan for distributing the information in the analysis? How available is the analysis? How relevant is the analysis to computer-match decisionmakers? What effect has the analysis had on current and future computer-match operations?

The 17 Matches We Reviewed for Cost-benefit Analysis

Agriculture	1. Evaluation of Food Stamp and AFDC wage-matching techniques, a research project assessing the costs and benefits of state Food Stamp and AFDC wage matches in Camden and Mercer counties, New Jersey; San Joaquin County, California; and the state of New Hampshire.
Defense	2. DOD retiree files matched with VA, OPM, and state and commercial death records, in order to identify benefits sent mistakenly to deceased military retirees. 3. Dual compensation and pension match to identify retired military officers who were employed as federal civilians and exceeded pay ceilings or failed to initiate payment offsets.
Education	4. Federal employees and federally insured student loans, a series of matches comparing federal employee records with student loan records to identify loan defaulters and apply salary offsets.
Health and Human Services	5. IRS Form 1099 (reporting unearned income) matches with Supplemental Security records, legislatively mandated to detect ineligible or overpaid recipients. 6. Missouri Supplemental Security and title XIX state data exchange match to identify persons erroneously receiving Supplemental Security Income payments as a result of unreported changes in title XIX living arrangements. 7. National "death matches," a series of legislatively mandated matches, developed from a pilot match, comparing federal and state death records with SSA records to identify payments made improperly because of unreported deaths. 8. Public assistance files matched with Massachusetts bank records to identify unreported income and assets of public assistance recipients.
Housing and Urban Development	9. Federal employees and title I debtors, a match to identify federal employees who defaulted on HUD-insured housing loans and apply salary offsets.

10. Tenant income verification matches, a series of matches in Atlanta and Athens, Georgia; Birmingham, Alabama; Kansas City, Kansas; and Kansas City, Missouri, state wage records and federal employment data on comparing public housing authority program tenants to identify unreported income.

Internal Revenue Service

11. Credit-for-the-elderly compliance program, a match to identify taxpayers who claimed a credit as elderly persons and received pensions or annuities from Social Security or the Railroad Retirement program.

Labor

12. Interstate crossmatch, an evaluation of the costs and benefits associated with matches identifying unemployment insurance overpayments to claimants in one state receiving benefits while working in another state.

Office of Personnel Management

13. SSA Death Matches I and II to identify unreported deaths by matching the civil service retirement annuity file and SSA's master beneficiary file of death records.

Veterans Administration

14. Federal employees and VA education loan defaulters, a match between active and retired federal civilian and military employees and defaulters on education loans under the GI bill.

15. Physicians credentials, a match between records on VA physicians and California state medical board records to detect physicians not properly certified to practice. A national match is in progress.

16. State wage data matches in Georgia and Florida comparing records on compensation and pension program beneficiaries with state records of employee earnings to detect unclaimed income. Additional matches are planned.

17. Vital statistics match, a match of compensation, pension, and education dependency records with state vital statistics to detect payments to ineligible surviving spouses and deceased clients.

Bibliography

Brown, Jay. "Collecting New Jersey Taxes." The Bureaucrat, Fall 1983, pp. 28-31.

Chelimsky, Eleanor. "Reducing Fraud and Abuse in Entitlement Programs: An Evaluative Perspective." GAO Review, Summer 1981, pp. 29-33.

Crowley, Brian P. "General Accounting Office Review of Efforts to Reduce Food Stamp Program Losses." Statement Before the Subcommittee on Domestic Marketing, Consumer Relations, and Nutrition of the House Committee on Agriculture. In Review of Fiscal Year 1984 Budget for the Food and Nutrition Service, U.S. Department of Agriculture. Hearings. 98th Cong., 1st sess. Washington, D.C.: U.S. Government Printing Office, April 20, 1983.

Dallaire, Gene. "Computer Matching: Should It Be Banned?" Communications of the ACM, June 1984, p. 537.

Duncan, Dean F. The Use of Wage Crossmatching as a Verification Technique and as a Method to Detect Errors. Prepared for U.S. Department of Agriculture, Food and Nutrition Service. Research Triangle Park, N.C.: Research Triangle Institute, May 1983.

Durnil, Jim. "Detecting Fraud, Waste and Abuse Through Computer Matching." Government Accountants Journal, Summer 1983, pp. 24-27.

Fischel, M., and L. Siegel. Computer-Aided Techniques Against Public Assistance Fraud: A Case Study of the Aid to Families with Dependent Children (AFDC) Program. Prepared for U.S. Law Enforcement Assistance Administration. McLean, Va.: Mitre Corp., 1980.

Glagola, Edward M. "Computer Matching GAO Seeker System." Paper presented at IMTEC Information System Review Methodology Workshop, U.S. General Accounting Office, Alexandria, Va., December 5-16, 1983.

Gramlich, Edward M. Benefit-Cost Analysis of Government Programs. Englewood Cliffs, N.J.: Prentice-Hall, 1981.

Greenberg, David. "Sniffing Out Fraud: Computer Sleuthing Comes to Public Welfare." Public Welfare, Summer 1984, pp. 32-39.

Greenberg, David, and Jennifer Pfeister. Wage Matching Techniques Used in Administering the Food Stamp and AFDC Programs: An Interim Analysis. Prepared for U.S. Department of Agriculture, Food and Nutrition Service. Menlo Park, Calif.: SRI International, 1983.

Greenberg, David, and Douglas Wolf. An Evaluation of Food Stamp and AFDC/Wage Matching Techniques: Final Report. Prepared for U.S. Department of Agriculture, Food and Nutrition Service. Menlo Park, Calif.: SRI Institute, May 1984.

Haendler, Frank. "New Computer Matching Project Aims to Shoot Down Fraud and Overpayment." Government Data Systems, November-December 1981, pp. 24-25.

Hayden, T., and N. Rollins. "An Evaluation of New York State's Wage Reporting System: The Real Cost of Computer Matching." In Oversight of Computer Matching to Detect Fraud and Mismanagement in Government Programs. Hearings Before the Subcommittee on Oversight of Government Management of the Committee on Government Affairs. 97th Cong., 2nd sess., Washington, D.C.: U.S. Government Printing Office, 1982.

Internal Revenue Service. Recapture of Accelerated/Accrued Deductions on HUD Involved Housing Study of the Deferred Adverse Tax Consequences Program. Washington, D.C.: June 1984.

Keisling, Phillip. "The Case Against Privacy." The Washington Monthly, May 1984, pp. 12-28.

Kirchner, Jake. "Privacy: A History of Computer Matching in Federal Government." Computerworld, December 14, 1981, pp. 1-16.

Kusserow, Richard. "Fighting Fraud, Waste, and Abuse." The Bureaucrat, Fall 1983, pp. 19-23.

———. "The Government Needs Computer Matching to Root Out Waste and Fraud." Communications of the ACM, June 1984, pp. 542-45.

Langan, Kenneth J. "Computer Matching Programs: A Threat to Privacy." Columbia Journal of Law and Social Problems, 15 (1979), 143-80.

McBride, Thomas. "Computer Matching: A Potent New Government Weapon Against Fraud and Waste." Stanford Lawyer, 17 (Winter 1982-83), 16-19.

Management Engineers, Inc. Evaluation of Pilot Study of Interstate Crossmatch Program. Prepared for U.S. Department of Labor, Unemployment Insurance Service. Reston, Va.: March 1984.

Marx, Gary T., and Nancy Reichman. "Routinizing the Discovery of Secrets." American Behavioral Scientist, 27 (1984), 423-52.

Novotny, Thomas W. "The IGs—A Random Walk." The Bureaucrat, Fall 1983, pp. 35-39.

Office of Management and Budget. "Agency Computer Match Checklist." Memorandum M-84-6, Washington, D.C, December 29, 1983.

———. "Responsibilities for the Maintenance of Records About Individuals by Federal Agencies." Circular A-108, Washington, D.C, July 1, 1975.

———. "Privacy Act Implementation." 40 Fed. Reg. 28,949 (1975).

———. "Privacy Act of 1974: Guidelines on the Relationship of the Debt Collection Act of 1982 to the Privacy Act of 1974." 48 Fed. Reg. 15,556 (1983).

———. "Privacy Act of 1974: Revised Supplemental Guidance for Conducting Matching Programs." 47 Fed. Reg. 21,656 (1982).

———. "Privacy Act of 1974: Supplemental Guidance for Matching Programs." 44 Fed. Reg. 23,138 (1979).

Office of Technology Assessment. Federal Government Information Technology: Computer Matching. Issue Paper 7. Washington, D.C.: July 1984.

———. Federal Government Information Technology: Electronic Record Systems and Individual Privacy. Washington, D.C.: June 1984.

President's Council on Integrity and Efficiency, Long-Term Computer-Matching Project. Computer Matching Newsletter. Washington, D.C.: 1985. Vols. 1-4.

President's Council on Integrity and Efficiency, Long-Term Computer-Matching Project. Incentives and Disincentives to Computer Matching. Washington, D.C.: 1983.

———. Inventory of Federal Computer Applications to Prevent/Detect Fraud, Waste, and Mismanagement. Washington, D.C.: U.S. Department of Labor, Office of Inspector General, n.d.

———. Model Control System and Resource Document for Conducting Computer Matching Projects Involving Individual Privacy Data. Washington, D.C.: December 29, 1983.

———. The Need for Standard Formats in Computer Matching. Washington, D.C.: U.S. Department of Health and Human Services, Office of Inspector General, n.d.

———. Proposed Standard Computer Matching Formats. Washington, D.C.: U.S. Department of Health and Human Services, Office of Inspector General, n.d.

———. Reference Paper on Computer Matching. Washington, D.C.: 1982. Pp. 1-9.

———. Report Forms for Standardized Formats Project: Monthly Final Implementation of Formats, Match Activities, Match Results. Washington, D.C.: U.S. Department of Health and Human Services, Office of Inspector General, n.d.

———. Reporting Standards for PCIE Report. Washington, D.C.: 1982. Pp. 5-8.

———. Report on the Development of Standardized Computer Matching Formats. Washington, D.C.: 1983.

Rodgers, Allan. Memorandum on Massachusetts Bank Match Program. Boston: Massachusetts Law Reform Institute, April 12, 1984.

Sassone, Peter G., and William A. Schaffer. Cost-Benefit Analysis: A Handbook. New York: Academic Press, 1978.

Shattuck, John. "Computer Matching Is a Serious Threat to Individual Rights." Communications of the ACM, June 1984, pp. 538-41.

Social Security Administration, Office of Family Assistance. Information Memorandum on Welfare Recipient/Banks Match Information Package. Baltimore, Md.: 1983.

Stokey, Edith, and Richard Zeckhauser. A Primer for PolicyAnalysis. New York: W. W. Norton, 1978.

Thompson, Mark. Benefit-Cost Analysis for Program Evaluation. Beverly Hills, Calif.: Sage Publications, 1980.

U.S. Congress, House of Representatives. Use of Internal Revenue Service Resources and Taxpayer Information for Draft Registration and Welfare Restriction Purposes. Hearings Before a Subcommittee of the Committee on Government Operations. 97th Cong., 2nd sess. Washington, D.C.: U.S. Government Printing Office, 1982.

———. VA Inspector General's Activities in Combating Waste, Fraud and Abuse and Review of VA Debt Collection Efforts. Hearings Before the Subcommittee on Oversight and Investigations of the Committee on Veterans' Affairs. 97th Cong., 2nd sess. Washington, D.C.: U.S. Government Printing Office, 1982.

U.S. Congress, Senate. Computer Matching: Taxpayer Records. Hearings Before the Subcommittee on Oversight of Government Management of the Committee on Governmental Affairs. 98th Cong., 2nd sess. Washington, D.C.: U.S. Government Printing Office, 1984.

———. Federal Identification Fraud. Report by Permanent Subcommittee on Investigations of the Committee on Governmental Affairs. 98th Cong., 1st sess. Washington, D.C.: U.S. Government Printing Office, 1983.

———. Fraud in Federal Programs. Hearings Before the Committee on Governmental Affairs. 97th Cong., 1st sess. Washington, D.C.: U.S. Government Printing Office, 1981.

———. Oversight of Computer Matching to Detect Fraud and Mismanagement in Government Programs. Hearings Before the Subcommittee on

Oversight of Government Management of the Committee on Governmental Affairs. 97th Cong., 2nd sess. Washington, D.C.: U.S. Government Printing Office, 1982.

U.S. Department of Labor, Office of Inspector General. Inventory of Computer Matching Activities in State/Labor and Related Agencies. Washington, D.C.: 1982.

———. Reference Paper on Computer Matching. Washington, D.C.: 1982.

U.S. General Accounting Office. "Better Wage Matching System and Procedures Would Enhance Food Stamp Program Integrity." Draft report, Washington, D.C., 1984.

———. A Central Wage File for Use by Federal Agencies: Benefits and Concerns, GAO/HRD-85-31. Washington, D.C.: May 21, 1985.

———. Eligibility Verification and Privacy in Federal Benefit Programs: A Delicate Balance, HRD-85-22. Washington, D.C.: 1985.

———. GAO Observations on the Use of Tax Return Information for Verification in Entitlement Programs, HRD-84-72. Washington, D.C.: 1984.

———. "A Position Paper on GAO's Present and Possible Future Role in Computer Matching Issues Related to Federally Funded Benefit Payment Programs." Manuscript, Philadelphia, 1984.

U.S. Department of Health and Human Services, Office of Inspector General. Computer Matching in State Administered Benefit Programs: A Manager's Guide to Decision Making. Washington, D.C.: 1983.

———. Inventory of State Computer Matching Technology. Washington, D.C.: 1983.

———. Nationwide Impact of Verifying Resources by Matching Recipients of Public Assistance with Bank Records. Washington, D.C.: 1983.

Usher, Charles L., Donna L. Watts, and Dean F. Duncan. The Development and Application of Error-Prone Profiles in the Food Stamp Program. Prepared for U.S. Department of Agriculture, Food and Nutrition Service. Research Triangle Park, N.C.: Triangle Institute, February 1982.

Veterans Administration, Office of Inspector General. Final Project/ Report: Federal Employees Receiving Government Assistance. Washington, D.C.: 1983.

———. Final Report: Review of Federal Employees and Annuitants Receiving VA Unemployability Compensation or Income-Based Benefits. Washington, D.C.: 1983.

Weiss, Laura. "Government Steps Up Use of Computer Matching to Find Fraud in Programs." Congressional Quarterly, February 26, 1983, pp. 431-33.

White, Robin. "Department of Income Maintenance Bank Matches Preliminary Summary." Paper delivered at Computer Matching Advisory Panel Meeting, U.S. General Accounting Office, Washington, D.C., July 17, 1984.

———. "A 1983 Status Report on State Bank Matching Programs and Related Activities." Paper prepared for American Civil Liberties Union, n.p., October 1983.

Glossary

ARIMA Interrupted Time-Series Analysis	Autoregressive integrated moving average (ARIMA) modeling refers to a class of stochastic process models that empirically describe changes in a variable over time as a function of the past behavior of this variable and, if appropriate, as a function of other substantively or empirically related variables. An intervention (for example, a computer match) hypothesized as affecting the variable (for example, monthly spells of program noncompliance) can be introduced into the model. A nonlinear regression is then used to estimate the weighted function of past observations together with the effect of the intervention on the subsequent level of the time series.
Concurrent Cost-benefit Analysis	An analysis performed in conjunction with the final refinement of a match hit list and initial verification and follow-up activities. The purpose of the analysis is to estimate, from a sample of hits, the costs and benefits of conducting verification and follow-up activities on the entire list of hits. The costs of processing the sample and the expected or actual benefits are used to project anticipated costs and benefits for the entire match.
Deterrence	A potential, qualitative benefit of computer matching, either specific or general. Specific deterrence is the inducement of persons who are not complying with program rules and regulations to change their behavior and comply with program requirements because they know that they may be detected by computer matching. General deterrence is the discouraging of potential abusers from trying to take improper advantage of benefits by making a computer match publicly known.
Discounting	Discounting translates the value of money, or other units of value, at one time into the value of money at a different time in order to make time periods comparable in assessing the costs and benefits of a computer match.
Hit	A hit is information on one or more data elements in two or more automated files that appear to be identical or similar (name, Social Security number, address, date of birth, and the like). A raw hit is the initial output of records from the different files that have been matched on selected data elements. It is subjected to some type of preliminary verification process in order to determine the specific reasons for the hit. Hits from coincidences or errors in recording names, Social Security numbers,

	and the like are eliminated. A solid hit requires further verification to determine whether possible fraud, error, or abuse exists; solid hits may be referred to investigators because of some indication of fraud or abuse.
Hit Prioritization	The selection of hits for verification and follow-up because of specific characteristics such as the amount of overpayment or debt to be recovered or the number of data elements that have been matched.
Match Definition and Development	The first phase of a match, in which the idea for the match is developed, its feasibility is assessed, and formal planning is begun. Initial contact is made with source agencies, and necessary compliance with legal requirements is identified.
Match Initiation and Production	The second phase of a match, in which legal regulations are complied with and agreements are made formal with participating agencies. Software is developed and tested, and some preliminary files are processed in preparation for the match. A list of raw hits is produced and sometimes refined by sampling and verification.
Match Verification and Follow-up	The third and final phase of a match that entails an iterative process of refining or processing hits to some final disposition. Starting with raw hits, match information is verified and hits with legitimate reasons for appearing in both files are eliminated from consideration. Additional information is sought on the remaining hits, in order to determine the reasons for them and the need for remedial action or referral. Indications of fraud or abuse may lead to more extensive verification and follow-up activities by investigators. For some hits, these activities may include administrative appeals, trials, and other legal proceedings.
Overpayment and Debt Recovery	A benefit of matching that results from the detection of an overpayment or debt and the collection of the money owed to an agency.
Overpayment Avoidance	A benefit of matching that results from the prevention of overpayments by identifying and correcting an error, thereby eliminating a source of future overpayments.

Pilot Match	The limited implementation of a match in order to assess the feasibility of a full-scale effort and estimate its costs and benefits. A pilot match may be restricted to a limited geographical region or to verification and follow-up activities on a statistically valid sample drawn from all raw hits.
Postmatch Cost-benefit Analysis	An analysis performed at the completion of a match, or in the latter part of verification and follow-up, to assess its effectiveness and efficiency. Figures on costs should reflect actual expenditures. Figures on benefits should include actual recoveries and projected estimates of outstanding collections and costs to be avoided.
Prematch Cost-benefit Analysis	An analysis in which the costs and benefits of performing a proposed match are estimated. The analysis usually extrapolates from previous matches, pilot matches, audit reports, and quality-control data on error rates.
Validation	The examination of the measures employed in a cost-benefit analysis in order to determine their meaningfulness. For example, the measures estimating the amount of overpayments to be avoided by a computer match are assessed in order to determine whether they do measure what they are intended to measure and whether the assumptions on which the methodology is based are correct.

U.S. GOVERNMENT PRINTING OFFICE:1986- 181- 235/ 40120

Requests for copies of GAO reports should be sent to:

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

Telephone 202-275-6241

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

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

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

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

Official Business
Penalty for Private Use $300

Address Correction Requested

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