

ADMINISTRATIVE DATA FOR
POLICY-RELEVANT RESEARCH:
ASSESSMENT OF CURRENT UTILITY
AND RECOMMENDATIONS
FOR DEVELOPMENT

A Report of the
Advisory Panel on Research Uses of Administrative Data
of the
Northwestern University/University of Chicago
Joint Center for Poverty Research



EXECUTIVE SUMMARY
SUMMARY OF FINDINGS AND RECOMMENDATIONS

V. JOSEPH HOTZ, ROBERT GOERGE,
JULIE BALZEKAS, AND FRANCIS MARGOLIN,
Editors

Members of Advisory Panel on Research Uses of Administrative Data:

V. Joseph Hotz*

Chair, UCLA

Julie D. Balzekas

*Executive Director;
Communications Director,
Joint Center for Poverty Research*

Norman Bradburn†

*University of Chicago and
Committee on National Statistics*

Henry E. Brady

University of California and UC-DATA

Gerald Gates

U.S. Bureau of the Census

Robert Goerge

*Chapin Hall Center for Children
University of Chicago*

Carol Luttrell

Massachusetts Department of Revenue

Frances Margolin**

American Hospital Association

Bruce Meyer

Northwestern University

Deanna Schexnayder

*University of Texas
Center for the Study of Human Resources*

□ Werner Schink‡

*Department of Social Services
State of California*

Michael Wiseman

University of Wisconsin-Madison

*Research Affiliate of the Joint Center for Poverty Research

†Faculty Affiliate of the Joint Center for Poverty Research

‡Member of External Advisory Board of the Joint Center for Poverty Research

**Original Executive Director of the Advisory Panel

EXECUTIVE SUMMARY

The report of the Advisory Panel on the Research Uses of Administrative Data¹ is concerned with administrative data collected at the state and local levels in the operation of government programs for the poor, such as AFDC/TANF, Food Stamps, Medicaid, and foster care. In addition to their “record-keeping” function, administrative data increasingly are used to monitor and evaluate program performance and ensure agency accountability.

The panel undertook this study at a time when American public assistance policies—particularly those aimed at families with dependent children—are changing. Governing authority and financial responsibility for public assistance have always been shared by federal, state, and local governments. However, as a consequence of the Personal Responsibility and Work Opportunities Reconciliation Act of 1996 (PRWORA), 104-193, control of the policies and programs that affect poor families with children has largely devolved from Washington to the states.

At the same time, the aim of cash assistance to poor families with children has also changed: from a federal entitlement to income support that sometimes provided education and job training to a time-limited benefits program principally focused on moving able-bodied adults into the labor force and their families off the welfare rolls. To achieve these new goals, state-designed programs must now develop and implement large-scale activities and services to help individuals and families become self-sufficient.

These changes are occurring in the presence of considerable uncertainty. Under the old law, the federal government had the means to collect relatively comparable state-generated program data. With the recent and profound devolution of family-related policies and programs to states and localities, the federal government no longer has a reliable means for monitoring what states are doing and how recipients are making out. Moreover, state and local governments have mixed experiences in producing reliable intrastate information on the effectiveness of alternative policies, much less reliable and valid data that permit interstate comparisons. Indeed, the latter, absent federal guidelines, is nearly impossible to achieve.

Consequently, in this report, the Advisory Panel seeks to:

- a. Describe the key practical and political considerations of transforming the information in these programmatic records into research-ready databases;
- b. Identify the strengths and weaknesses of administrative data, relative to that gathered in national surveys, for use in descriptive and evaluative research and in accountability-based monitoring of program performance;
- c. Describe examples of several states’ efforts to develop an ongoing capacity to use administrative data for both programmatic and policy evaluations; and
- d. Make initial recommendations that will improve the quality and usefulness of administrative data for policymakers and program administrators.

The panel examined several states’ efforts to develop intrastate databases from administrative records and reviewed how these databases are used to monitor and evaluate

¹ The Advisory Panel on the Research Uses of Administrative Data was formed under the auspices of the Northwestern University/University of Chicago Joint Center for Poverty Research in the fall of 1996. The Center sought funding from the Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services, to form an Advisory Panel to assess the development of research-ready data from state administrative sources in the areas of public assistance, public health and welfare for use in policy and academic research. The Advisory Panel, of researchers, state and federal officials, and experts in the area of data protection and archiving, first met in the fall of 1996 and continued to correspond and meet through 1997.

key public assistance programs and the disadvantaged populations they serve. In addition, we consulted many additional experts on the production and use of administrative data. We sought to obtain their assessments of the potential value that these sources of data may have in the future, and what issues need to be addressed if these nascent local efforts are to be replicated in other states.

Our report synthesizes the findings from our investigation and recommends ways that various groups—including policymakers, administrators, researchers, and foundations—can develop better administrative data for monitoring and evaluating the implemented welfare reforms. The remainder of this section summarizes our findings and recommendations.

Findings

1. The Value of Administrative Data

The Personal Responsibility and Work Opportunities Reconciliation Act of 1996 (PRWORA), 104-193, has set in motion an array of new policies and programs, including Temporary Assistance for Needy Families (TANF), which replaces Aid to Families with Dependent Children (AFDC).

To answer the critical policy and program questions of “what works,” “for whom,” and “at what cost,” the Advisory Panel finds that:

- **Policymakers and program administrators will require more and better data sources than they now have if they are to adequately monitor program operations and evaluate program outcomes.** Program administrators and policymakers will need reliable state and local data if, among other things, they are to: summarize program operations; determine who is being served by which programs and determine who is being underserved; who is not being served but should be; how services can best be targeted to those in greatest need; determine which strategies and services are most cost-effective; track individual work histories; track individual and family earnings and income; and describe the conditions of poor children and their families, relative to the conditions of other households.
- **Current national survey research data cannot adequately monitor the diverse, local programs now being established by state and local governments.** In the context of devolution, none of the national cross-sectional and longitudinal data sets is large enough to support separate analyses of poverty-related issues for any but the few largest states.
- **To obtain reliable information over time and across programs and agencies, it will be necessary to augment current administrative databases and to link them together.** Administrative data provide detailed and accurate program information, large sample sizes that allow for more types of analyses, and state-specific data that reflect variations in state and local programs.
 - For example, *administrative data offer the advantage of allowing for sub-state analyses*, thereby allowing the many AFDC/TANF waiver programs operating in a limited number of counties to be better studied.
 - *Administrative data can also provide information on the same individual or case over long periods of time.* Such capabilities are increasingly important if we are to understand how, for example, recipient behavior and well-being change in response to both time-limited benefits and varying economic and labor force conditions.

- *Due largely to advances in computer technology, linking administrative databases is easier, less expensive, and more reliable than ever before.* Such efforts can provide richer, more comprehensive information on how the poor, the working poor, and others are faring and how, and to what extent, they contribute to or consume public tax and transfer benefits.

2. Key Operational Issues for Developing Research-Relevant Administrative Databases

A clear case can be made for greater reliance on state and local administrative data systems for monitoring and evaluating public assistance programs in the future. To date, administrative databases have mostly been used in one-time evaluations based on random-assignment designs. While that experience offers many valuable lessons for improving administrative data, the structural changes in the welfare system under PRWORA may mean that states will be less likely to use experimental designs to evaluate their public assistance programs, if they conduct any impact evaluations at all. In this context, states and local governments are likely to make even greater use of administrative data for whatever evaluations they conduct.

Based on our review of past evaluations and research using administrative data and our investigation of present efforts to develop an ongoing capacity to provide research-ready administrative data, the Advisory Panel finds that:

- **There are three operational issues that can “make or break” the development of administrative databases: (1) negotiating appropriate interagency agreements; (2) negotiating agreements in which agencies retain adequate control over any new demands researchers impose on agency employees and the nature of the information researchers may disclose about agency operations; and (3) developing protocols that protect the privacy of clients and the confidentiality of data.**

EXAMPLES OF WORKING MODELS

The Advisory Panel found that when the issues of interagency cooperation, researcher-agency agreements, and client privacy are worked out in a mutually satisfactory way for all parties, solid, ongoing working relationships developed. In turn, over time, and across discrete projects, such relationships remove obstacles to using administrative data for research. Indeed, in some instances, researchers and agencies were able to agree to take additional steps to improve and use administrative data for research on an ongoing basis. The Advisory Panel describes five successful collaborative efforts in this report:

California, where the collaboration of state and university researchers on data collection and evaluation for California’s federal AFDC waiver, the 1992 California Work Pays Demonstration Project, has led to the creation of five ongoing databases and the construction of analytic datasets for continued program evaluation and research.

Illinois, where a collaborative effort between university researchers and a single state agency in the early 1980s has evolved into a multiservice, integrated research database, constructed out of administrative data gathered by numerous public agencies serving children and families in Illinois, and now used to track the impact of reforms on caseloads and across agencies.

Massachusetts, where efforts to link tax administrative databases with other agency databases for the purpose of enforcing child support payments demonstrated the potential of administrative data for research and evaluation and earned support for the creation of a longitudinal database for research and evaluation of social service programs, which is now in use.

Texas, where the availability of administrative data for research has been largely facilitated by the establishment of performance measures by the state legislature in the early 1990s, evaluation of the Texas Jobs Program, implemented in 1990, and the multiagency data collection and data sharing that both required, and permit the establishment of key performance measures for workforce development and other programs.

Oregon, where an integrated database project, still under development, was mandated by the state legislature in 1993 in order to provide a database for future evaluations of welfare reform.

In examining these efforts, the Advisory Panel noted the following aspects of successful operations:

- A collaboration between one or more state agencies and outside academic and independent research groups or institutions is key to developing successful, ongoing administrative databases for research.
- Initial development of ongoing administrative databases tends to be idiosyncratic and entrepreneurial, where someone involved in the research enterprise possesses a larger vision of what administrative data might do and is able to implement the vision.
- Ongoing databases are more typically the result of “bottom-up” rather than “top-down” development efforts. That is, they tend to be the result of localized and more idiosyncratic efforts as opposed to mandates from above.
- A key feature of the “entrepreneurial” effort that initiated and sustained the existing databases is the presence of someone or some group that holds a longer-run perspective so that the database is not viewed as a “one-shot effort,” useful only for a single project or contract.

From these observations, the Advisory Panel concludes that a mutual investment by social assistance agencies, policymakers, and researchers in an entrepreneurial effort to create and sustain ongoing capacity will improve the quality of administrative data for research and make possible the monitoring and evaluation demanded by emerging programs and policies.

QUALITY AND INFORMATION CONTENT OF ADMINISTRATIVE DATA: POTENTIAL LIMITATIONS

Another set of challenges has to do with the nature and quality of administrative data. Primarily because administrative data are gathered in the context and for the purpose of administering a program, issues concerning the quality and information content of administrative data pose considerable obstacles to drawing inferences about program trends and impacts. The Advisory panel identified the following potential concerns:

- **Administrative data possess some limitations that diminish their value in certain types of research, including: (1) inability to estimate such things as the rates of program participation; (2) inability to measure all outcomes, such as indicators of well-being that would not be tracked in the program-based data, or to measure**

anything when a person is “off the program”; and, (3) difficulty in comparing programs across states in the absence of standardized information collection.

However, the panel also sees significant opportunity to offset these limitations through linking data. While administrative data from one program seldom contain enough information for a useful evaluation, by linking administrative data from different programs it can become possible to obtain an array of explanatory and outcome variables. Further, linking information from state administrative databases with survey data on individuals and households has considerable promise.

RECOMMENDATIONS TO FOSTER THE DEVELOPMENT OF ADMINISTRATIVE DATA

Based on these findings, the Advisory Panel offers the following recommendations to foster the development of administrative data as an integral data source for public assistance research in the future. The recommendations cover three key areas: (1) *institution building*; (2) *confidentiality and privacy protection*; and (3) *assessing the quality and comparability of interstate administrative data*. They are summarized below.

Fostering institution building

Across the country, opportunities are emerging for the development of ongoing administrative databases for research of social assistance programs and policies. It is the Advisory Panel’s view that states and the nation need to build on these promising efforts and develop permanent, ongoing administrative data capacities. To help realize that goal, the panel offers three sets of recommendations to foster the construction of permanent administrative data “institutions.”

- **A centralized and ongoing repository of information on administrative data efforts should be established (and funded).**
- **States without administrative databases organized for research should be encouraged to establish partnerships with independent research organizations, such as those at universities, to help develop, maintain, and use administrative databases on an ongoing basis for program monitoring and evaluation.**
- **National organizations, such as American Public Welfare Association (APWA) or the Welfare Information Network (WIN), as well as organizations and groups within the academic community, such as Association for Public Policy Analysis and Management (APPAM) and the National Association of Welfare Researchers and Statisticians (NAWARS) need to find ways to recognize and encourage the use of administrative data in research.**

Further assessment of confidentiality and privacy concerns

While the Advisory Panel found that existing principles and recommendations regarding confidentiality and privacy apply to the research uses of administrative data, some new issues need to be addressed, having primarily to do with disclosure limitation techniques and the applicability of federal legislation to particular states. Therefore,

- The Advisory Panel calls on independent organizations (such as the National Research Council’s Committee on National Statistics) as well as professional organizations (such as the American Statistical Association) to conduct a more thorough assessment of the adequacy of existing principles and practices that protect the privacy and confidentiality of the information contained in administrative databases.

Assessing and improving the quality and across-state comparability of administrative data for public assistance programs

Great strides have been made in the “science” of developing administrative databases, especially those that contain longitudinal information on program participants and those that consist of data linked across various databases. Nonetheless, it is the Advisory Panel’s assessment that many unanswered questions persist regarding the quality and usability of administrative data for many types of research, and this is especially true for evaluating the impacts of emerging state- and county-based welfare programs under PRWORA. The panel strongly believes that more research on the comparability of administrative and survey data needs to be done if administrative data are to become a trusted and appropriately used source of data in high-quality research.

- **The Advisory Panel urges that funding be provided by agencies like the National Science Foundation, private foundations, and government agencies themselves to further research and analysis on such questions as: (1) quality of administrative data; (2) comparability with other data sources, such as survey data; (3) methodological strategies for dealing with such analytic issues as the denominator problem that affect the range of use of data; and, (4) the interactions of research and management objectives and how this affects the structure and quality of such data.**
- **The panel also urges research organizations, such as the Joint Center for Poverty Research, and academic publishers and journals to encourage and help legitimize such research by creating “outlets” for it, including convening conferences and supporting volumes or special issues of journals on these topics.**
- **Further, the panel would urge those working on the “management” side of the equation, including professional organizations for the public sector, to collaborate and help support efforts to develop higher quality administrative data.**

The Advisory Panel’s final recommendation concerns data comparability across states. If across-state comparisons are to prove informative as alternative programs and policies developed by states over the next few years are monitored and evaluated, data that contain comparable measures and populations at the state level are essential. Clearly, administrative databases can play a crucial role in across-state comparisons. But to play that role, attention must be paid to achieving greater comparability of information and populations in these databases. Accordingly, the panel offers the following recommendation to highlight this important issue.

- **Develop guidelines and standards to ensure that comparable and high-quality data are gathered across states and across agencies within states.**

We suggest that the National Research Council’s Committee on National Statistics be commissioned to establish an expert panel to assess and make recommendations on ways to foster such data comparability. This expert panel should include and work with representatives from state and local governments. It should also seek input from professional organizations such as the APWA and NAWRS. Finally, this panel should recommend structures and institutional arrangements that will encourage an ongoing partnership between the states and federal government for gathering objective, high-quality, and comparable data on populations receiving (or at risk to receive) public assistance provided under PRWORA and related programs. The National Center for Health Statistics

and National Center for Education Statistics might be considered possible institutional models for public assistance data-gathering.

CHAPTER 6
**Developing the Research Potential
of Administrative Data: Summary of Findings
and Recommendations**

The Advisory Panel on the Research Uses of Administrative Data foresees substantial need for data able to support the rich array of research and monitoring needs of states and the nation in the changing climate of welfare reform and the “New Federalism.” In the concluding chapter, a summary is offered of what has been learned about administrative data sources and what appear to be the critical obstacles to increasing their usefulness in research and in monitoring policy effectiveness.

As this report makes evident, the utility of administrative data for policy-relevant research is considerable at present and promising for the future. Several states have developed or are developing structures and institutional arrangements to sustain ongoing state administrative databases to support research on state-level public assistance programs. These databases contain longitudinal information on individuals and/or households who have participated in one or more public assistance or training programs. The data have been (or can be) linked across a number of different administrative databases to produce a range of measures needed to assess the performance of these programs. Nevertheless, such databases have been developed only in a limited number of states. Further, their sustainability, where they have been developed, remains less than certain.

A number of recommendations for developing the research potential of administrative data emerge from the Advisory Panel findings. Because the experiences with building and using administrative databases for policy research, though quite promising, are still limited, the recommendations tend to be rather modest. The panel offers them to stimulate ways to think about information needs in a new era of social programs and policies that are more local in their orientation and more varied in their objectives. The recommendations are offered in the hope that they will help forge a mutual investment on the part of policymakers and analysts, agency officials and program managers, and researchers inside and outside the academy, to pursue the more sophisticated program monitoring of a larger number of antipoverty strategies, imperative for both broad policy interests and for administration of newly emerging programs and policies.

6.1 Findings: Where We Are and Where We Need to Go

This section provides a summary of conclusions the panel drew from its investigation of the present utility of administrative data. Many of these conclusions confirm the increasing understanding of what is entailed in the use of administrative data for research, while others present possibilities for improving their usefulness in the future.

A number of observations shared by the panel about administrative data and its role in future research efforts were important in motivating the panel’s work. They were:

I. Administrative data sources will need to play a greater role in the monitoring and evaluating of the impacts of social assistance programs in the coming century.

As discussed throughout the report, the devolution of social assistance policy and programs to the state and local level will make the degree of reliance on standard surveys, such as the CPS and SIPP, less tenable in the future. The national cross-sectional and longitudinal data sets derived from

standard surveys currently provide inadequate state-level sample sizes. In order to support separate analyses of poverty-related issues for the majority of states, these surveys will have to develop measures of such phenomena as program participation in an increasingly heterogeneous policy climate. Forming such measures will take time and, even so, may ultimately have limited usefulness as a source of information for comparing programs and outcomes across states, let alone within states. The shortcomings of standard national survey data in the wake of devolution motivate a more serious effort to develop administrative databases that can compensate, in part.

II. To meet the research needs of future evaluations of policy and programs, there will be a growing emphasis on building administrative databases for linking information across time and across programs and agencies.

Several aspects of the changing nature of social assistance policy places great value on the ability to link administrative data across time (for a given case or individual) and across administrative record systems. The value of being able to reliably link case or individual data arises for at least two reasons.

- First, the new state- and local-based social programs generally have multiple objectives. Under PRWORA, for example, states are encouraged to develop TANF plans and programs with a multifaceted set of objectives, including encouraging self-sufficiency, reducing out-of-wedlock births, and improving the well-being of children. To assess each of these objectives, and the extent to which states can meet them as a group, clearly requires monitoring information from more than one administrative domain.
- Second, time-limited aid under PRWORA necessitates a capacity to develop longitudinal databases on individuals who participate in one or more of these programs over their lifetimes.

These fundamental needs for linking information over time and across program databases are clear. However, the ability of states to develop “linkable,” research-ready administrative databases places great strains on existing data systems for the variety of reasons discussed in Chapters 2 and 3.

The panel spent considerable time identifying and studying existing efforts to develop administrative databases. Those that appeared to provide an ongoing capacity for supporting multiple types of research were most intriguing. The Advisory Panel became particularly interested in how these centers formed, how they were structured, and what keeps them going, and came to several conclusions.

III. Several lessons may be learned from the experiences of existing efforts to develop ongoing administrative data for use in policy-relevant research:

- Ongoing databases are more typically the result of “bottom-up” rather than “top-down” development efforts. That is, they tend to be the result of localized and more idiosyncratic efforts as opposed to the mandates of any one agency.
- A key element in the development of successful administrative databases for research, especially those of an ongoing nature, usually have involved a

collaboration between one or more state agencies and outside academic and independent research groups or institutions.

- A key feature of the entrepreneurial effort that initiated and sustained the existing databases was the presence of someone or some group that held a longer-run perspective, so that the database was not viewed as useful for just a single project or contract.

The panel was struck by the similarity of these efforts to small start-ups. In terms of the energy and ingenuity that comes with such start-ups, they possess the same attraction as developing a new product in one's garage. But, as the statistics on the turnover of small firms indicate, such enterprises also can be unstable and short-lived. Below, the panel offers some suggestions for ways to help sustain the development of these ongoing databases, while maintaining the benefits of the bottom-up approach.

The panel also was struck by the fact that the cases studied for this report involved partnerships of state agencies and outside groups, including a group or institution that is part of a university or an independent research institute. Several features of these partnerships have proved beneficial to these developmental efforts, including: the independence of the research organization; the role that these partnerships have played in establishing sustaining institutional trust; and, the fact that these partnerships provide already strapped and overburdened state agencies.

- The panel's investigation also uncovered a number of important barriers that impede the development of administrative records into analytic data sets capable of supporting high-quality research and accessible to responsible policy and academic researchers. These barriers can be described as follows:

IV. Key institutional issues representing "make-or-break" factors in the development of administrative databases are:

- The ability to negotiate interagency agreements; and,
- The ability to obtain protocols that protect the privacy of clients and the agency and the confidentiality of data.

V. Key issues in the ability to make such databases accessible to more than in-house researchers are:

- The ability to find adequate safeguards for protecting the privacy of clients and confidentiality of data; and,
- The ability to establish contractual (or less formal) relationships giving researchers freedom to conduct research while protecting the integrity of the agency, the privacy of clients, and the confidentiality of the data.

Two common themes underpin these obstacles. The first is the issue of confidentiality and privacy and the profound implications of this issue for the entire enterprise of using administrative data for research. The panel tried, in the entire chapter devoted to this subject (Chapter 3), to compile and discuss existing principles and standards that provide guidance for ways to safeguard confidentiality while still enabling responsible researchers to gain access to sensitive data. But, as also has been discussed, using administrative databases for research raises new

issues for which the existing standards, developed primarily in the context of large-scale national surveys, may not be adequate. The combination of the technological revolution in information systems and electronic sources of information and some unfortunate abuses of confidentiality have resulted in a level of anxiety and skepticism on the part of the American public that makes this topic all the more sensitive and in need of further assessment. This issue is addressed below.

The other common thread running through the obstacles outlined above is the element of trust, or lack of it, between various communities that produce, analyze, and ultimately own administrative data. On this score, the Advisory Panel is encouraged by the extent to which the efforts described in Chapter 5 seem to demonstrate that cooperation and understanding emerge and are strengthened through partnerships. Agencies have learned that properly informed researchers can use their data responsibly. Academic researchers involved in these efforts have come to better understand the political and operational responsibilities these agencies bear precisely because the information in their databases exists, first and foremost, to fulfill management and service-delivery needs. What remains is finding ways to replicate and sustain the trust-building experiences that have so benefited the research value of administrative data.

Finally, as discussed in Chapter 5, there are limitations to administrative data that affect their usefulness and appropriateness for many types of research. In particular, the panel has noted that:

VI. Administrative data have a number of limitations that diminish their value in certain types of research. These include:

- The choice-, event-, or participation-based nature of administrative data limits inferences and gives rise to the “denominator problem.”
- Administrative data typically do not contain adequate control variables, e.g., demographics of clients.
- Administrative data do not measure all outcomes, e.g., some types of indicators of well-being.
- Data are only available when the client is “in the program”; little is known when a person leaves the program.

It is important to note that the limitations noted above need not be permanent. As discussed in Chapter 4, the capacity to link information across many administrative data sources can go a long way toward reducing both the problems of limited control variables and outcomes that would exist in any one administrative data system. Clearly, the technology exists today to perform those links. But the problems of confidentiality and obtaining the interagency agreements to perform these links represent significant costs (and potential barriers) to efforts to develop more comprehensive data systems.

Similarly, concerns about potential biases in administrative data, due to their use in managing programs and for measuring performance in accountability systems, may turn out to be unwarranted. But, to date, we simply lack sufficient experience in the research community from comparisons across survey and administrative data sources to know the nature or extent of these biases. In the next section, the panel recommends that such data comparisons be encouraged, rewarded, and funded so that limitations can be reduced or removed. But, at the same time, it is important that

all parties be aware of these current limitations and exercise caution, given the relatively early stage of using administrative data for program evaluation and analytic research in the context of social service programs.

Key Findings

- I. Administrative data sources will need to play a greater role in the monitoring and evaluating of the impacts of social assistance programs in the coming century.
- II. To meet the research needs of future evaluations of policy and programs, there will be a growing emphasis on building administrative databases for linking information across time and across programs and agencies.
- III. Several lessons may be learned from the experiences of existing efforts to develop ongoing administrative data for use in policy-relevant research:
 - On-going databases are more typically the result of “bottom-up” rather than “top-down” development efforts. That is, they tend to be the result of localized and more idiosyncratic efforts as opposed to the mandates of any one agency.
 - A key element in the development of successful administrative databases for research, especially those of an ongoing nature, usually have involved a collaboration between one or more state agencies and outside academic and independent research groups or institutions.
 - A key feature of the entrepreneurial effort that initiated and sustained the existing databases was the presence of someone or some group that held a longer-run perspective, so that the database was not viewed as useful for just a single project or contract.
- IV. Key institutional issues representing “make-or-break” factors in the development of administrative databases are: the ability to negotiate interagency agreements, and the ability to obtain protocols that protect the privacy of clients and the agency and the confidentiality of data.
- V. Key issues in the ability to make such databases accessible to more than in-house researchers are:
 - The ability to find adequate safeguards for protecting the privacy of clients and confidentiality of data; and,
 - The ability to establish contractual (or less formal) relationships giving researchers freedom to conduct research while still providing agencies with adequate controls over what researchers disclose.
- VI. Administrative data have a number of limitations that diminish their value in certain types of research. These include:
 - The choice-, event-, or participation-based nature of administrative data limits inferences and gives rise to the “denominator problem.”
 - Administrative data typically do not contain adequate control variables, e.g., demographics of clients.
 - Administrative data do not measure all outcomes, e.g., some types of indicators of well-being.
 - Data are only available when client is “in the program;” little is known when a person leaves the program.

6.2 Recommendations for Developing the Research Value of Administrative Data

In this last section several recommendations are outlined to help foster the construction and research value of administrative databases from public assistance programs over the next few years. It is always tempting for panels such as this one to offer bold and expansive recommendations. But, as noted in the Introduction and as this report indicates, despite a number of efforts around the country to develop and use administrative data for research on an ongoing basis, the “newness” of these efforts does not support sweeping recommendations. These efforts will, however, benefit from practical, though modest, recommendations, which may also support other incremental efforts to develop ongoing administrative databases. The Advisory Panel makes recommendations in three areas:

- Fostering Institution Building
- Further Assessment of Confidentiality and Privacy Concerns
- Assessing and Improving the Quality and Across-State Comparability of Administrative Data for Public Assistance Programs

6.2.1 Fostering institution building

Across the country, a number of opportunities are emerging for the development of ongoing administrative databases for research of social programs and policies. It is the panel’s view that states and the nation need to build on these promising efforts and develop permanent, ongoing administrative data capacities to monitor policy changes and their impact on the disadvantaged segments of the population. To help realize that goal, the panel offers three sets of recommendations to foster the construction of permanent administrative data institutions.

The first recommendation concerns the need to improve interactions between those designing, developing, and using state and local administrative data and their access to information of common interest. One of the messages heard from many state research staff and data managers was the difficulty they encountered in gathering information from other states and learning from their peers. Thus, the panel recommends:

I. Establish (and fund) a centralized and ongoing repository of information on administrative data.

A repository could be as basic as a web site maintained by an existing professional organization or research center. Crucial to its usefulness is that it collect and disseminate information on the following issues and topics:

- Legislative and administrative strategies at the state level for dealing with confidentiality and privacy concerns;
- Up-to-date documentation on federal regulations related to administrative data under TANF and other federal programs;
- Legislative and administrative strategies at the state level for dealing with the establishment of interagency agreements;
- Prototypes of agreements for providing nongovernmental researchers access to data that meet concerns of confidentiality and safeguard the political integrity of agencies and state and local governments;
- Reports on the ways in which different state agencies use administrative data to improve the management and accountability of their programs.

The second recommendation concerns the importance of bringing together researchers and program administrators in the development and maintenance of administrative databases that can and will be used to conduct policy-relevant research. Existing efforts to build such databases (see Chapter 5) required involvement between researchers and state and local program administrators. Therefore, the panel believes that an important way to support the

use and improvement of administrative data is to encourage states to establish such “partnerships” as they develop their administrative databases for research.

II. Encourage states without administrative databases to establish partnerships with independent research organizations, such as those at universities, to help develop and use administrative databases on an ongoing basis.

The development of administrative data in the five cases examined for this report provides several different models for partnerships. Regardless of the particular structure, the panel considers partnerships between states and research institutions an important role if both the development and quality of administrative data for research is to be assured and improved over the coming years.

When considering such partnerships, the Advisory Panel takes seriously the inherent tension between establishing a trusting relationship between researchers and government agencies and the need for researchers to maintain their professional integrity so as not to undermine the credibility of research findings. “[T]he value of research. . . is its credibility in the policy debate, most of which has to do with its objective scientific respectability...but some of [which] has to do with the absence of any non-scientific motivation of the researcher” (Bouman 1997). Researchers need to maintain a fine line between recognizing the political and bureaucratic realities and pressures that impinge on agencies and their administrators when conducting and presenting their research while retaining the necessary independence in the work that they do. This tension was evident in several of the efforts to develop ongoing administrative data research efforts discussed in Chapter 5. For example, some of the reports produced by the Center for the Study of Human Resources (CSHR) at the University of Texas were critical of the procedures and practices of the state agencies whose cooperation they needed to get their data. But, as panel member Dr. Deanna Schexnayder noted, the center was able to establish a credible voice within the state, which is respected and listened to by various parties, precisely because researchers with the CSHR, as well as the center itself, have maintained their independence.

Third, it is important that mechanisms be created to more visibly recognize and encourage efforts to build and use administrative databases in research.

III. National organizations (such as American Public Welfare Association (APWA) or the Welfare Information Network (WIN) or National Governors Association (NGA)) as well as organizations and groups within the academic community, (such as Association for Public Policy and Management (APPAM) and National Association of Welfare Research and Statistics (NAWRS)) need to find ways to recognize and encourage the use of administrative data in research.

Rewarding the painstaking and often thankless work that must be undertaken to develop and sustain ongoing administrative databases will help to encourage their development. Organizations such as the American Public Welfare Association or the Welfare Information Network or National Governors Association should identify state or local examples of accomplishments in the development of administrative databases. They could even establish awards or grants to encourage these efforts. Organizations such as the Association for Public Policy and Management, the American Statistical Association, and the National Association of Welfare Research and Statistics, might encourage presentations and sessions at their annual conferences devoted to research on the methodologies supporting administrative data, and on research findings that used administrative data sources and program applications that have redesigned administrative data systems to serve both the ends of research and evaluation and program monitoring. These latter efforts would signal the importance

and legitimacy of using and analyzing administrative databases in scholarly research as well as make useful design changes more attractive to program managers.

6.2.2 Further assessment of confidentiality and privacy concerns

In its assessment of confidentiality and privacy concerns, the panel found that many of the existing principles and recommendations of previous initiatives apply to the research uses of administrative data. At the same time, new issues were identified that threaten confidentiality and privacy protections when using administrative data. The panel also found that while federal confidentiality guidelines are well-established, guidelines and legislative acts at the state level are quite diverse and, in some cases, the principles are potentially inadequate to protect privacy and assure the public (including program participants) that information on them will not be used improperly. Therefore, the panel thinks it wise to re-examine and further assess the adequacy of existing practices, especially in light of growing public skepticism about the privacy of information governmental units possess.

IV. Independent organizations (such as the Committee on National Statistics), as well as professional organizations (such as the American Statistical Association) need to conduct a more thorough assessment of the adequacy of existing principles and practices that will protect the privacy of individuals and confidentiality of the information contained in administrative databases. Special attention should be paid to such questions as:

- How should informed consent of program participants to be handled with respect to the use of information on them for research?
- What mechanisms and procedures should be adopted that will provide access of these data to responsible researchers while still safeguarding the privacy of individuals?
- What guidance can be provided for crafting interagency agreements?
- What are the proper “disclosure” standards for these databases when reporting on results from research based on these data?

6.2.3 Assessing and improving the quality and across-state comparability of administrative data for public assistance programs

The Advisory Panel’s final two sets of recommendations concern data quality and comparability across units. Great strides have been made in the “science” of developing administrative databases, especially those that contain longitudinal information on program participants and those that consist of data linked across various databases. Nonetheless, it is the panel’s assessment that many unanswered questions persist regarding the quality and usability of administrative data for many types of research, including uses that will require across-state comparisons.

The situation today confronting the use of administrative data to monitor and analyze the impacts of the nation’s devolving public assistance programs bears a remarkable resemblance to that faced in the 1980s with respect to the monitoring and assessing the condition and performance of the nation’s primary and secondary educational systems. As with the new emphasis on state and local control in the recently enacted welfare reform legislation, the primary responsibility for financing and governing of schools and education programs has long rested with state and local governments. But over the last three decades, there has been a growing concern about the quality and success of the nation’s schools, sparked in part by performance comparison of American students in math and science with students in other countries. These concerns were first chronicled in *A Nation at Risk*, a report of the National Commission on Excellence in Education, released in 1983. Furthermore, the debate about appropriate curriculum and teaching methods in the U.S.

gave rise to an increasing need for high-quality and comparable data on student performance and what was happening in our nation's schools.¹

While the National Center for Education Statistics, within the U.S. Department of Education, had responsibility for providing such data, serious concerns existed about its adequacy. As described in a report issued in 1986 by the "Panel to Evaluate the National Center for Educational Statistics (NCES)," created by the National Research Council Committee on National Statistics,² the quality and comparability of data on educational performance for the nation's elementary and secondary schools was inadequate. As noted by one of the commentators on this report, "if the data continue to be as inaccurate in the future as they have been in the past, all other issues are moot." The panel summarized their findings about the quality and comparability of national education statistics and data as follows:

The poor quality of the data is generally attributed to the fact that data are collected, in large part, from administrative records maintained at the local level, which record "official" rather than "real" behavior; that the data are the product of diverse record-keeping systems that lack comparability in definitions and time periods; that the data provided to the center [NCES] are at such gross levels of aggregation, such as for a state as a whole, as to seriously limit anyone's ability to check them for accuracy, consistency, and reasonableness; and that the data as published are at some summary levels of geography, such as a region, as to seriously limit their analytical usefulness.

This criticism was accompanied by a series of recommendations for improving the quality and comparability of educational data collected and compiled by NCES. While still in progress, NCES has adopted many of these recommendations. For example, it has sponsored more research on the quality of its data and appropriateness of its various measures of performance. And, reflecting the political and historical realities of the local control of educational data, it has worked in partnership with states and professional organizations (such as the American Statistical Association and the American Educational Research Organization) to develop standards to improve the quality and comparability of common data across states.

In an effort to avoid repeating some of the problems confronted in developing education statistics, as well as to learn from the strategies for dealing with them, the panel offers several recommendations for ways to improve the quality of administrative data and to promote greater comparability of data elements derived from administrative data for the state- and local-level public assistance programs emerging under PRWORA.

6.2.3.a. ASSESSING THE QUALITY AND VALIDITY OF ADMINISTRATIVE DATA

The concerns outlined above about the quality and comparability of state-based educational statistics derived from administrative records in the 1980s offer parallels to the situation facing human service program administrators and researchers interested in poverty and policy today. As discussed in Chapter 4, when it comes to using administrative data in evaluations of the impacts of emerging state- and county-based welfare programs under PRWORA, a number of important questions remain unanswered. There have been a few studies of the comparability of variables such as income and program participation status across administrative and survey data (several of which are cited in this report). But the panel strongly believes more research on the comparability of administrative and survey data needs to be done if administrative data are to become a trusted and appropriately used source of data in high-quality research. Therefore, the panel recommends that funding agencies and foundations, as well as professional and research organizations, give more

¹ For a more complete discussion of these developments, see Elliott and Ralph (1997).

² See Levine (1986).

attention to and expand their support for research that validates and assesses the relative quality and adequacy of administrative data for all types of research, especially evaluation research.

V. Funding needs to be provided by agencies (such as National Science Foundation), private foundations, and government agencies themselves to further research and analysis on such questions as:

- Quality of administrative data;
- Comparability with other data sources, such as survey data;
- Methodological strategies for dealing with analytic issues, such as the denominator problem, that affect the range of data use; and,
- The interactions of research and management objectives and how this affects the structure and quality of data.

VI. Research organizations (such as the Joint Center for Poverty Research), and academic publishers and journals must encourage and help legitimize research on these questions by creating outlets for it, including convening conferences and supporting volumes or special issues of journals on these topics.

VII. Those working on the management side of the equation, including professional organizations for the public sector, must collaborate and help support efforts to develop higher quality administrative data.

For example, the panel noted in Chapter 4 the importance of having high-quality data for implementing “results-based accountability” systems developed by government agencies. Such groups have a direct self-interest in the improvement of administrative databases for the types of “research” that must be done to assess where praise and blame should be lodged. The panel would urge organizations such as the National Academy for Public Administration and the Council for Excellence in Government to take a proactive role in promoting the assessment of data quality in information systems used for performance assessment.

6.2.3.b IMPROVING ACROSS-STATE COMPARABILITY OF ADMINISTRATIVE DATA

The Advisory Panel’s final recommendation concerns data comparability across states. As noted in the introduction, the current trend in social policy places unprecedented responsibility and control in the hands of state and local governments. Many have predicted that this change in the locus of control is likely to result in even less research and monitoring of program performance than in the past. But, as the initial assessment of several states has indicated, these dire forecasts seem premature. As the examples of welfare reform legislation from Illinois and California illustrate, states are not ignoring the research component in their implementation of welfare reform. Moreover, administrative data are likely to serve as a key source of data for whatever research and program evaluation states *do* perform. But as promising as these commitments to research are on the part of a few states, it is the panel’s assessment that an important national goal remains—namely, being able to monitor and evaluate the impacts of the alternative policies developed by states over the next few years. If across-state comparisons are to prove useful and informative, data that contain comparable measures and populations at the state level are needed.

One potential source of comparable data will be from surveys of nationally representative populations, including the CPS, SIPP, and the new Survey of Program Dynamics (SPD) for the population of social program participants. But, as also noted in the introduction, the sample sizes from these surveys will only be adequate for the largest states. Clearly, administrative databases can play a crucial role in across-state comparisons. But to play that

role, attention must be paid to achieving greater comparability of information and populations in these databases. Accordingly, the panel offers the following recommendation to highlight this important issue.

VIII. Guidelines and standards must be developed to ensure that comparable and high-quality data are gathered across states and across agencies within states.

Following the model used for educational statistics, the Advisory Panel suggests that the National Research Council Committee on National Statistics be commissioned to establish a panel to assess and make recommendations on ways to foster data comparability. Some crucial issues include:

- the availability of universal identifiers to facilitate linking administrative records across states, which should be addressed in the context of PRWORA;
- review of the comparability of state-provided measures of outcome and demographic variables mandated under PRWORA; and,
- assessment of what other data elements could be made available by most or all states that would be valuable in monitoring and evaluating the impacts of social assistance programs on a nationwide basis.

In developing their recommendations, this panel will need to be mindful of the current political climate in which calls for stringent and mandatory guidelines or standards from the federal government are not likely to be palatable to the states. Therefore, a panel on data comparability must include in its membership representatives from state and local governments, and seek input from professional organizations such as APWA and NAWRS.

Finally, this recommended panel on data comparability should be asked to assess what institutional and governmental structures might be put in place to improve the quality of administrative and other data sources used to monitor and evaluate public assistance programs in the U.S. Entities such as the National Center for Education Statistics (NCES) or the National Center for Health Statistics (NCHS) provide useful models that could be adapted to the public assistance context. Furthermore, such a panel should consider ways in which an NCES- or NCHS-like arrangement could foster an ongoing partnership between states and the federal government, and begin to separate national data gathering efforts in the area of public assistance from the enforcement and sanctioning that has been the focus of many past and existing federal reporting requirements in this area.

Recommendations in Three Key Areas:

A. Fostering Institution Building

- I. Establish (and fund) a centralized and ongoing repository of information on administrative data.
- II. Encourage states without administrative databases to establish partnerships with independent research organizations, such as those at universities, to help develop and use administrative databases on an ongoing basis.
- III. National organizations (such as APWA or the WIN) as well as organizations and groups within the academic community (such as APPAM and NAWRS) need to find ways to recognize and encourage the use of administrative data in research.

B. Further Assessment of Confidentiality and Privacy Concerns

- IV. Independent organizations, such as the Committee on National Statistics, as well as professional organizations (such as the American Statistical Association) need to conduct a more thorough assessment of the adequacy of existing principles and practices that will protect the privacy of individuals and confidentiality of the information contained in administrative databases. Special attention should be paid to such questions as:
 - How should informed consent of program participants with respect to the use of information on them for research be handled?
 - What mechanisms and procedures should be adopted that will provide access of these data to responsible researchers while still safeguarding the privacy of individuals?
 - What guidance can be provided for crafting interagency agreements?
 - What are the proper “disclosure” standards for these databases when reporting on results from research based on these data?

C. Assessing and Improving the Quality and Across-State Comparability of Administrative Data for Public Assistance Programs

- V. Funding needs to be provided by agencies (such as the National Science Foundation), private foundations, and government agencies themselves to further research and analysis on such questions as:
 - Quality of administrative data;
 - Comparability with other data sources, such as survey data;
 - Methodological strategies for dealing with analytic issues such as the denominator problem, that affect the range of use of data; and,
 - The interactions of research and management objectives and how this affects the structure and quality of data.
- VI. Research organizations (such as the Joint Center for Poverty Research) and academic publishers and journals must encourage and help legitimize research on these questions by creating outlets for it, including convening conferences and supporting volumes or special issues of journals on these topics.
- VII. Those working on the “management” side of the equation, including professional organizations for the public sector, must collaborate and help support efforts to develop higher quality administrative data.
- VIII. Guidelines and standards need to be developed to ensure that comparable and high-quality data are gathered across states and across agencies within states.

6.3 Concluding Observations

Social policy is undergoing dramatic changes today, with the responsibility for design and implementation of these policies devolving to state and local governments. Because of the scarcity of resources and the reluctance of the nation's citizens to support "big government," what is at stake today in social policy is higher for states, agencies, and clients than it was 20 years ago. To address the uncertainty about "what works" and "for whom," it is important that the focus and tools of research adapt if we are to accurately and fairly describe, monitor, and evaluate just what these changes imply for our nation's poor and disadvantaged.

The Advisory Panel on Research Uses of Administrative Data is convinced that administrative data can and ought to be one of the important tools in this research effort. Administrative data can provide a cost-effective yet extremely useful source of information with which to monitor and evaluate the impacts of changes in social policy at the state and local levels. At the same time, much work is needed to develop administrative data systems that can routinely provide information for this research on an ongoing basis. It is the Advisory Panel's hope that this report stimulates and encourages policymakers, program managers, researchers, and funding agencies and foundations to join in the effort to strengthen administrative sources of data and to ensure that administrative data play an expanding role in monitoring the well-being of the nation's disadvantaged.