

# “Whatever we have been doing”: Policy control over TANF

Thomas Kaplan

---

Thomas Kaplan is Senior Scientist and Associate Director at the Institute for Research on Poverty, University of Wisconsin–Madison.

---

An influential early critic of the welfare reforms of 1996, Wendell Primus, recently acknowledged a change in his attitude toward current welfare reform, stating that “whatever we have been doing the last five years we ought to keep doing.”<sup>1</sup> This article focuses on a part of “whatever we have been doing.” Even as researchers assess the well-being of actual or potential participants in Temporary Assistance for Needy Families (TANF), the cash welfare program arising from the 1996 reforms, the detailed program experiences of TANF participants have remained unclear. This lack of clarity raises questions about how federal and state policymakers might know more or, if they wish, exert greater program control.

## Current murkiness

To some extent, the lack of clarity concerning what actually happens in TANF programs is an intentional outgrowth of policy design. Federal TANF legislation established mandatory performance standards covering the percentage of the caseload that must participate in work or work-related activities and imposed financial penalties on states failing to meet the percentages. The legislation also specified that separate TANF financial bonuses would go to “high-performance states” (defined administratively as those that most successfully move clients into stable, better-paying jobs) and states that most reduce their rates of nonmarital pregnancy. The legislation did not detail the programmatic or administrative structure that states must develop as they meet, or do not meet, the performance goals. Nevertheless, having shifted from an AFDC program offering direct aid to individuals toward a block grant that instead delivers money to state or local (usually public but also private) organizations, policymakers and other observers naturally want to know what these organizations are doing to and for program participants.

It is possible that policymakers will later decide that these questions no longer merit attention, the same fate that earlier questions about what AFDC recipients were doing with their grants—sheltering a “man in the house,” starting a garden that could offset their food needs, etc.—met in the 1960s. But for the foreseeable future, if only to determine

whether certain organizational practices lead with reasonable reliability to better or worse outcomes for program participants, interested observers are likely to want to know in some detail what experiences the organizations receiving public funds provide for TANF participants.

Yet vagueness concerning the experiences of program participants and the practices of the organizations that serve them may be inevitable. Good public management generally requires information systems that allow for varied and decentralized managerial procedures, “while at the same time assuring accountability to central authorities by reporting accurate and interpretable information on the achievement of program objectives, the provision of services, and program expenditures,” as Richard Nathan and Thomas Gais noted in their “first look” at the implementation of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA).<sup>2</sup> Yet good welfare-to-work practice—at least as it has so far been defined in many of the front-line agencies serving TANF participants (and partially corroborated by one independent study<sup>3</sup>)—makes that management ideal hard to achieve. Staff in these agencies often speak of their “whatever it takes” approach to moving participants into the labor market. Even making allowances for the inevitable “spin,” and for the favorable financial conditions under which most TANF programs have operated until recently, it is hard at least in Wisconsin to be unimpressed by the willingness of TANF programs to put together complex and individualized service packages—ranging from residential drug treatment in a program that allows children to live with their mothers during the treatment, to vocational training, to the imposition of benefit sanctions—for at least some participants.

In such an individualized system, in which case managers have much flexibility to exercise their judgment, it is difficult to summarize the operation of that flexibility in the rectangular categories (reading down or across a table) that management information systems best capture. We may know at any one time how many clients are in English as a Second Language programs, how many are in introductory programs that describe what TANF clients can expect from the agency and attempt to motivate program participation, how many are in postsecondary education, and how many work, if these programs are distinct. But at least in the Wisconsin TANF operation, a growing number of programs are hybrids, and they may be entered into the management information system in the way that is most advantageous to program operators. A program that combines classroom instruction with work experience, for example, may be entered into the state data system either as classroom instruction or as community service work. Because federal law

**Table 1**  
**Typology of Work First Programs in Michigan**

<b>General Approach</b>	<b>Dominant Activities</b>	<b>Primary Purpose</b>
Job-seeking support	Workshops on résumé preparation, interviewing techniques, etc. Staff develop job leads, bring employers to the program for interviews.	Place all or nearly all clients into a job.
Job-search preparation	Workshops on résumé preparation, interviewing techniques, etc. Staff do not develop job leads or bring employers in for interviews. Instead, clients do their own self-directed job search.	Place all or nearly all clients into a job.
Labor market sorter	Workshops on résumé preparation, interviewing techniques, etc. Staff develop job leads, bring employers to the program for interviews. Clients who do not find a job receive extra assistance.	Determine, through success or failure in the labor market, which clients need extra assistance to obtain a job, and provide that assistance only to those who need it.
Client responsibility	Clients receive individual assistance (but not organized workshops) in preparing résumés and finding job openings.	Place all or nearly all clients into a first job, but assure that the program signals that finding a job is the client's responsibility, not the program's.

**Source:** N. Anderson and K. Seefeldt, "Inside Michigan Work First Programs," Michigan Program on Poverty and Social Welfare Policy, April 2000.

allows no more than 20 percent of those in "work experience" to be engaged in vocational education, and because community service work generates more administrative credit for clients with a high school diploma or equivalency degree under Wisconsin state policy, the hybrid is more likely to be entered as community service, and policymakers then lose information about the actual nature of the program.

W. Edwards Deming, the founder of the quality movement in the United States, often expressed an aversion to numeric performance standards; indeed, the 11th of his 14 quality principles called for their elimination.<sup>4</sup> Deming was concerned that, in order to meet the quotas, organizations would either take undesirable shortcuts that hurt performance in other areas, or they would hide their actual activities to assure surface conformance. Under PRWORA, although most states have generated sufficient caseload reduction credits to be unaffected by nominal work participation requirements, the fact that work is promoted over training probably leads states to define activities that could be considered both as a form of community service work.<sup>5</sup>

One could imagine a data system with the ability to control for this tendency by capturing nuances of hybrid programs: there could be one entry for a hybrid program in which participants spend the majority of their time in the classroom and another if participants spend the majority of their time in work experience. But even assuming that such distinctions are meaningful—how would one treat extra tutoring that some but not all participants receive?—there is a limit to the number of choices a case manager would reasonably consider in entering data.

More fundamentally, important elements of the interaction between case managers and clients cannot be captured in any data system. If case managers have flexibility, they will exercise it differently in ways that are important. Some will

be more or less willing than others to make exceptions to general program rules and either sanction or not sanction program participants for particular infractions. Some case managers, believing that a participant will gain no benefit from the program if she and her family lack a stable residential setting, will be more willing than others to define as "community service job" activity the time a participant spends finding housing. The data system can only capture whether or not the participant received a sanction or was recorded as a community service job case, not an informal agreement concerning how the client should spend some of her community service job time.

### **Can we know anything useful about practice?**

Despite the murkiness, we know much that is important about the program experiences of TANF participants, particularly the experiences reflected in state budgets and state financial assistance policies. States appear to have substantially reduced their budgets for direct financial assistance to low-income households and increased their expenditures on child care, child welfare, and employment and training.<sup>6</sup> In addition, annual federal reports allow us to categorize states by whether they offer earnings supplements, the size of the supplements, and the length and nature of time limits on financial assistance. We could also categorize programs by whether participants interact with two case managers—one for employment and training and one for eligibility and benefits for cash assistance—or one case manager, and this would tell us something at least moderately useful about the experiences of program participants.

More subtly, some researchers have pointed us in the direction of categorizing programs by central tendencies in the relationship of programs to their participants. Two investigators familiar with county Work First programs

in Michigan, Nathaniel Anderson and Kristin S. Seefeldt, for example, organized the agencies into four “typologies to describe the variation of service delivery.” Table 1 offers a simplified description of the typologies.

Because all the programs had a general work-first emphasis, this typology rests on often subtle differences that the researchers observed in dominant activities and purposes. Some programs offered organized workshops in job-seeking skills and some did not; the staff in some programs developed job leads and brought potential employers to the program, whereas staff in other programs provided less assistance; some programs sought only job placement, whereas others were equally concerned that clients understand that finding a job was their own responsibility.

In a broader effort to categorize the central tendencies of state TANF programs, some coauthors recently carried me along in the development of a taxonomy that was intended to capture the full range of state programs. Table 2 displays a simplified version of the taxonomy.

The taxonomy relies heavily on an assessment of the motivation behind each state program, not necessarily the motivation expressed by state governors and other policymakers, since such expressions are often varied and conflicting, but rather motivations as judged by expert observers inside the state.<sup>7</sup> We argued that state TANF programs attempt primarily to promote one of three activities—motivating clients, building their skills, or reducing barriers—so that clients will either work or not be dependent. Thus, we said, Minnesota motivates clients to work by providing large earnings disregards that allow the welfare grant to subsidize earnings, whereas Wisconsin motivates clients to not be dependent by providing assistance only for work. In contrast, states like Kansas and Oregon do extensive and individualized early assessments of clients and their families to discover and address barriers to work, and Michigan’s Work First agencies have moved increasingly to build work skills among those who have obtained entry-level jobs.

### Federal and state control

I stress that the state-level taxonomy presented in Table 2 captures only differences in emphases, that many state TANF programs falling into some categories have much in common with programs in other categories, and that state emphases change over time. Even if this taxonomy helps analysts think about differences among state programs at a particular moment, case managers in Wisconsin say in interviews and surveys that they select from all these strategies and goals, based on the apparent needs, backgrounds, and aspirations of their “customers.” Although Wisconsin policy may overall be aimed at changing client motivation, many individual sessions between case managers and their clients seem primarily directed at reducing barriers to employment. Case managers in

**Table 2**  
**Taxonomy of State TANF Programs**

Programs Seek Primarily to:	So That Program Clients Will:
Motivate clients	Work or Not be dependent
Build client skills	Work or Not be dependent
Reduce barriers faced by clients	Work or Not be dependent

Source: T. Gais, R. Nathan, I. Lurie, and T. Kaplan, “Implementation of the Personal Responsibility Act of 1996,” in *The New World of Welfare*, ed. R. Blank and R. Haskins (Washington, DC: Brookings Institution Press, 2001), pp. 35–64.

other states are likely to be similarly eclectic in their actual interactions with program participants.

Is there a place for oversight and control in such an eclectic and flexible policy environment? A useful starting point is to assure that large evaluations of TANF programs include a significant study of program implementation, particularly if the outcome or impact parts of the evaluation rely heavily on administrative data. Evaluators should spend enough time observing and assessing daily activity in the programs they study to allow them to understand, and perhaps even compensate for, any bias in administrative data.

A second useful step may be to expand current federal reporting requirements imposed on states to include all clients receiving TANF-funded services (and state-funded services under maintenance-of-effort requirements), even if the clients receive no cash assistance. Under present law, states must report the household relationships, age, race, and educational background of household members, the type of service provided, and other data for households receiving cash assistance, but states do not report to the federal government on households receiving TANF-funded services without cash assistance. One reason for the distinction between families receiving TANF cash assistance and services and those receiving just TANF services has been to assure that only those receiving cash assistance “tick” their time limit clocks. It should be possible, however, to distinguish, for time-limit purposes, between households receiving and not receiving cash assistance, even if the reporting pool includes both kinds of households. If that were done, policymakers would have at least some understanding, consistent across states, of the range of TANF-funded services provided to households not currently receiving cash assistance.

Even with this change, however, an effective oversight role for state and federal policymakers would be hard to develop. If particular activities that clients are supposed to be

in are specified in legislation—as they are under the current iteration of TANF, with its emphasis on allowable activities that look like jobs—then states are likely to use their inevitable flexibility to report in a way that is to their advantage. If federal funds go primarily to programs that achieve particular outputs, such as a specified percentage of program graduates who are employed for 30 or 180 days at a specified wage, then policymakers will need to be alert to the kind of client “creaming” that occurred in the 1980s under the Job Training Partnership Act.

An alternative or additional form of control that Congress might consider in its reauthorization debate would govern program inputs: the credentials of the staff who exercise the discretion at the heart of local programs. When program operators have flexibility to choose among competing approaches to particular clients, and when the outcomes of those choices cannot be known with certainty (the client may obtain a job with particular wages and benefits, but the contribution of the case manager or the program to that outcome will be uncertain), a common way to ensure that the choices comply with some standard is through insisting that the staff making the choices have a particular educational background or professional certification.

Case managers in Wisconsin’s TANF program are not professionals in the ordinary sense of that term: in a 1999 survey, 14 percent of the case managers said that their highest level of education was a high school diploma, 42 percent had some college but lacked a bachelor’s degree, 31 percent had a bachelor’s degree, and about 11 percent had done some graduate work.<sup>8</sup> Wisconsin does, however, require that new case managers take a common state training program that lasts (in Milwaukee, where case managers do not handle Food Stamp and Medicaid eligibility determination) about two months, and case managers cannot have their own caseload until that training has been completed.

As one form of control over the eclectic and flexible decisions of case managers, TANF reauthorization legislation might specify either what educational background TANF case managers must have or the kind of preservice training that state or local organizations must provide these staff.

Developing such specifications may be difficult, since the ideal educational background for case managers in TANF programs is not obvious: case managers often address a variety of educational, family, cognitive, health, and other barriers as they help place clients in jobs and training, and no single profession has yet become dominant in welfare-to-work programs. Social work is a potential candidate for professional dominance, but its educational programs have not yet emphasized work and training. Rehabilitation counseling—another possibility for professional dominance—now deals only with clients who have disabilities. Because of these complexities, and because states could reasonably differ in their judgment of what constitutes good training, TANF

reauthorization might instead require states to establish and report on their own educational background and training requirements.

I acknowledge that the establishment of professional or quasiprofessional background requirements for case managers is unlikely to stir enthusiasm in the public or among politicians, and that any requirements will inevitably create, as well as address, problems. Public policymakers—who must necessarily cede discretion to school teachers operating in their individual classrooms, to police officers responding to emergencies, and to other “street-level bureaucrats”<sup>9</sup> in addition to TANF case managers—often try to exert some of their limited control through the establishment of common certification standards or initial training requirements. As the many criticisms of current certification requirements for teachers suggest, such efforts generate their own problems: certain requirements are inevitably too rigid for some and too flexible for others. Yet the difficulty of achieving consensus does not suggest that we should have no standards at all. Even critics of current teacher certification standards seek generally to substitute other requirements, such as college graduation or particular work experience, for the courses in education now specified, not to eliminate background requirements entirely. State or federal policymakers with a clear vision of what they want TANF case managers to do should be expected to back that vision by specifying a required training or experiential background. ■

---

<sup>1</sup>Quoted in E. Harden, “Two-Parent Families Rise after Changes in Welfare,” *New York Times*, August 12, 2001.

<sup>2</sup>R. Nathan and T. Gais, *Implementing the Personal Responsibility Act of 1996: A First Look* (Albany, NY: Rockefeller Institute Press, 1999), p. 53.

<sup>3</sup>D. Bloom and C. Michalopoulos, *How Welfare and Work Policies Affect Employment and Income: A Synthesis of Research*, Manpower Demonstration Research Corporation, New York, 2001, find that programs that carefully combine job search and short-term training yield better results than either job search or training alone.

<sup>4</sup>See, for example, M. Walton, *The Deming Management Method* (New York: Perigee Books, 1996).

<sup>5</sup>The nominal requirement in 2001 holds that 50 percent of single parents must be “engaged in work.”

<sup>6</sup>D. Ellwood and D. Boyd, *Changes in State Spending on Social Services since the Implementation of Welfare Reform: A Preliminary Report*, Rockefeller Institute of Government, Albany, NY, 2000.

<sup>7</sup>A recent essay by Irene Lurie provides a description of the methodology behind the studies that arrive at such conclusions. I. Lurie, “Field Network Studies,” in *Implementation Analysis: An Evaluation Approach Whose Time Has Come*, ed. M. Lennon and T. Corbett (Washington, DC: Urban Institute Press, forthcoming).

<sup>8</sup>T. Kaplan and T. Corbett, “The Implementation of W-2,” in *W-2 Child Support Demonstration Evaluation, Phase 1: Final Report*, Vol. II, ed. D. Meyer and M. Cancian, Institute for Research on Poverty, University of Wisconsin–Madison, 2001.

<sup>9</sup>The phrase comes from M. Lipsky, *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services* (New York: Russell Sage, 1980). The book offers several reasons for the discretion that is characteristic of many public sector jobs.

# Welfare reform: The institutional dimension

Lawrence M. Mead

---

Lawrence M. Mead is Professor of Politics at New York University and an IRP affiliate.

---

What does government have to do at the state and local level to make a success of welfare reform? The answers begin with policy but finally focus on the capacity of the regime. Government must want to “do the right thing,” but it must also be able to do it.

## Getting policy right

What is best policy for welfare reform, of course, depends upon one’s goals. Temporary Assistance for Needy Families (TANF) awards localities great discretion in redesigning family welfare. At the same time, the work participation standards in the program force states to give priority to moving adult recipients into jobs. All states also say they are emphasizing employment.

If so, then it is reasonably clear from past evaluations what successful welfare work programs look like. First, they must be mandatory, requiring participation for all adults deemed employable, rather than leaving work as a choice. Many states have enacted more generous work disregards, but these, experience has shown, are not enough to attain high work levels. Rather, work effort must be enforced as a condition of aid. Second, programs must emphasize actual work in available jobs, rather than education or training as was favored by federal policy prior to TANF.<sup>1</sup>

Although most states claim to be pursuing “work first” policies, there are differences of degree. In Wisconsin, to be working in some job, either private or public, is virtually a precondition for aid, whereas in Portland, Oregon, recipients are encouraged to seek out full-time positions with good pay and benefits even if this delays employment somewhat—an approach that has evaluated well.<sup>2</sup> It is enough that a program be work-focused, making clear to recipients that they must soon confront the labor market as it is.

## Implementation

Implementing reform programs is just as important as designing them. To execute work policies, localities must be willing to spend money on the necessary bureaucracy and support services, particularly child care. Fortu-

nately, the recent fall in the caseload virtually everywhere has given states a windfall. The decline is driven by good economic conditions and enhanced benefits such as the Earned Income Tax Credit as well as by welfare reform. Even states that have not yet done much to change welfare have gained. This has allowed them to afford reform, at least in the short run.

Money, however, is only the beginning. The administrators who run programs locally must be fully “on board.” They must initiate the new policies or at least embrace changes decided above them. If they do not “own” reform, they will implement it perfunctorily, or not at all.<sup>3</sup> In many localities, welfare administrations are still so oriented to accurate grant payment—the “quality control culture”—that employment remains an afterthought. In some places, staffs have resisted the shift toward “work first,” as this curbs their ability to send clients to school for several years in hopes of getting them better-paying jobs.<sup>4</sup> And with caseloads falling for other reasons, welfare agencies may feel little actual pressure to build work into their routines.

How to engage the bureaucracy fully in reform is a deep problem. Part of the answer is instituting incentives—for example, tying funding to performance measures such as job entries—that give local staffs reasons to implement. A longer-term answer is recruiting more talented and motivated people to work in welfare administration. The search for more dynamic structures explains much of the reorganization that has gone on around welfare reform, including the frequent devolution of some functions to nongovernmental organizations.

My studies of Wisconsin and New York convince me, however, that the institutional challenges posed by reform stretch beyond mere implementation to the general capacities of government. If one compares these two important states, one sees radical differences in their ability to handle reform. The contrast is not mainly in the generosity of policy, since both have high benefits and ambitious antipoverty programs. Rather, it is in their general ability to decide what to do about difficult issues, and then to carry it out.

## Wisconsin

Wisconsin, by common consent, has achieved the most radical welfare reform in the nation. Here, to an extent not seen elsewhere, welfare was totally redesigned. Rather than just add work programs to Aid to Families with Dependent Children (AFDC), as most states have done, Wisconsin recentered welfare on the individual rather than the family and modeled the new system’s

benefits and obligations on what low-paid families off welfare could expect.

Serious efforts to enforce work in the Badger State began in the mid-1980s, unusually early. After a decade of change that included multiple federal waiver programs, Wisconsin implemented Wisconsin Works (W-2) in 1997. The new system combines the most severe work tests in the nation with unusually generous support services for the entire working poor population. The combination has driven the cash welfare rolls down by about 90 percent and work levels up from already high levels.<sup>5</sup> In both 1997 and 1999, Wisconsin recorded the highest work levels among low-income parents of any of the 13 states in the Urban Institute's National Survey of America's Families.<sup>6</sup> We cannot attribute that strictly to Wisconsin's welfare policies, but they probably contributed.

The political basis of the Wisconsin reform was an unusual concordat in the state legislature. In 1985–86, the welfare rolls crested at over 300,000 people—in a population of less than 5 million. Both parties agreed that dependency was a crisis demanding action. The two differed in approach, but they also focused on the same problem, and this drove them toward agreement on a work-based reform. Each gave up a central political stake in order to accomplish that. Democrats abandoned entitlement, agreeing to condition aid on serious work tests rather than distributing it according to need alone. But Republicans abandoned the ambition to downsize government, agreeing to massive expansions of the bureaucracy and support services. The reform emerged as conservative but progovernment, transcending normal partisan divisions.

The essence of this deal was struck as early as 1986, when the state began to build up welfare work programs—prior to Tommy Thompson's election as governor. Thompson and his aides were masterful in enacting the numerous reforms that followed, but they drew on broad support in both parties. In addition, community groups and urban politicians did not try to halt fundamental change, as they have done in some other states with large caseloads. They criticized details of the reform, but they accepted the essential need to refound aid for the working-aged on employment.<sup>7</sup>

The bureaucratic statecraft behind reform was, perhaps, even more impressive. Because of its early development of work programs, Wisconsin implemented the JOBS program early and well. Then, in the early 1990s, it refocused that program less on remediation and more on work, using a variety of funding incentives. It was already apparent that the high-performing counties in JOBS were those that enforced participation and actual work, so Madison acted to promote this style statewide.<sup>8</sup> Local officials also gave a lead. Administrators in Grant,

Kenosha, and other counties crafted reform programs on their own. Some of these became models for state-level reforms. Other counties signed up to pilot various Thompson experiments.<sup>9</sup> The implementation of W-2 involved further innovations, including bureaucratic competition, extensive use of performance measures, and the privatization of administration in Milwaukee.

One might criticize the Wisconsin reform for a lack of policy analysis. Leaders proceeded sensibly on the basis of experience and past experiments, but only W-2 was planned in any depth. The state made little effort to evaluate its programs. One might also criticize W-2 for excessive severity. The new system took risks by driving too many people away from welfare too fast, and its strictures on education and training were probably excessive. The first implementation of the new system was also troubled, especially in Milwaukee. Fortunately, the system has since allowed somewhat more training, and the administrative problems have mostly been solved.

Overall, the Wisconsin reform was a triumph *for government*. The state vindicated a paternalist approach to poverty, showing what a combination of demands and new benefits could achieve. More important, it showed that an inspired regime can actually decide and execute such a policy. It thus opened a new vista for antipoverty policy. But unfortunately, few states are so masterful.

## New York

Greatly in contrast to the Badger State, the Empire State has had considerable difficulty coping with welfare reform.<sup>10</sup> Five years after the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), the welfare rolls have fallen substantially, but there is still no fundamental change. Although stability might reflect an active preference by the state, it has, rather, resulted from political paralysis and serious administrative problems.

New York's central problem is a deep partisan division about how to handle welfare and poverty. Most Democrats and liberals in the state adhere to the ideal of entitlement, the idea that the poor deserve public support simply on the basis of need and without behavioral expectations. Republicans and conservatives aim mainly to cut social spending in order to deal with the state's serious budgetary problems. Neither side shows the magnanimity or creativity displayed in Wisconsin. The division matters because each party has long controlled part of the state government. A Democratic majority is entrenched in the Assembly, a Republican one in the Senate, while the governorship oscillates between parties. In Wisconsin, because of broad agreement on change, shifts in partisan control of the legislature made much less difference.

The impasse means that New York cannot set its own direction in welfare. In response to both the Family Support Act (FSA) of 1988 and PRWORA in 1996, New York was able to do little more than comply with new federal requirements. Some contentious issues, particularly how tough to be about work requirements, were delegated to the counties because they could not be resolved in Albany. Although some change has occurred at the local level, there is no consensus about direction.

Rudolph Giuliani, while mayor of New York City, aggressively sought to reduce the welfare rolls, and brought in a welfare director from Wisconsin, Jason Turner, to accomplish this. Turner's attempts to divert people from TANF and put many recipients in government jobs, however, earned nothing like the acquiescence seen in Milwaukee. Opposition from community groups, poverty lawyers, and local Democratic politicians has been intense, and press coverage was far more hostile than in Wisconsin. Outnumbered conservatives support the Mayor with equal stridency.<sup>11</sup> The city has rebuffed all requests to research its program, so fearful is it of academic criticism.

New York also faces serious administrative problems, especially in New York City. The bureaucracy is large and well funded by national standards, but it is also passive. Administrators do not take the same initiatives seen in Wisconsin because they know political support may be lacking. They await direction from elected leaders, but since the parties are so divided, this is seldom forthcoming. The capability of personnel in the bureaucracy, although substantial, falls short of the impressive level seen in Wisconsin. Agencies also are more resistant to management improvements and accountability measures. Whereas Madison rebuilt welfare administration with only minor resistance from organized labor, New York's unions deter managers from comparable changes.

One long-standing problem has been an inability to build a job search requirement into welfare. Traditionally, New York counted upon the Employment Service to provide this aspect of welfare work programs. In the 1980s, after disappointment with the performance of the service, the welfare department took total control of work programs, a shift confirmed under JOBS. But lacking the Employment Service's performance measures for job entries and other outcomes, welfare could never deliver more than the activity levels demanded by JOBS. In New York City, Turner has tried to build up performance management in welfare. He also reorganized local welfare offices as "job centers" and brought in private contractors to conduct job clubs. But these changes have been slowed by lawsuits. New York has not approached the thorough integration of the assistance and work missions seen in Wisconsin's W-2.

## National patterns

The cases of Wisconsin and New York, it turns out, can be generalized to much of the nation. It is difficult to speak definitively about welfare reform nationwide, because the task of studying what all 50 states are doing would be overwhelming. However, the Urban Institute project, *Assessing the New Federalism*, includes case studies of 13 states, and the Rockefeller Institute at the State University of New York is following reform implementation in 21 states.<sup>12</sup> From other sources, I have information on Oregon as well.<sup>13</sup> The case studies of these states—24 in all—permit one to say something about how localities have handled the welfare issue. And it is remarkable how many of these states fall into patterns similar to those of Wisconsin or New York.

One group has handled reform more or less in the Wisconsin manner. These states were able to agree about what to do by focusing on welfare as a policy problem and adopting work-centered reforms. Change in these states often occurred in advance of legislation in Washington. Legislators acted with restrained partisanship and also on the basis of experience, drawing often on their own waiver programs or other homegrown experiments. Policy decisions had some coherence and were not merely compromises among parties or interest groups. Administrators then implemented the reforms with, at most, limited problems. The final result was a new welfare policy that truly represented the will of the community. Among the UI or Rockefeller states, this group includes Kansas, Michigan, Minnesota, Oregon, Tennessee, Utah, and Washington, as well as Wisconsin.

In a second group of states, politicians were more divided about reform. Change began later and was driven more by the outside pressure of federal reforms than by internal initiatives. Debate about what to do was long, partisan, and contentious, and decisions left bitterness in their wake. In some of these states, black leaders or community groups disowned reform as racist or unacceptable. Implementation was also troubled by entrenched agency or local interests. Typically, the welfare and labor departments battled for the control of welfare work programs, or counties resisted direction from the center. Although change occurred, it was usually incremental and, more important, conflicted. Consensus was lacking. One could not say that government had truly achieved a direction. Among the UI or Rockefeller study states, this group includes California, Colorado, Massachusetts, New Jersey, and Ohio, as well as New York.

A third group comprises states that had to frame a serious welfare policy for the first time. These localities never previously had much policy other than to set AFDC benefit levels and, with them, the potential size of the caseload. Other details they abandoned to federal

policymakers. Dependency was limited mostly by setting benefit levels low or allowing them to decline with inflation, not by attempting to promote self-reliance within welfare. These states were less likely than others to pursue waiver programs.

When these states acted, policy often arose from the personal ideas of governors or other politicians, without the institutional basis provided elsewhere by prior experiment or party positions. In some cases, mistakes were made, requiring hasty corrections. Administratively, too, prior development was rudimentary. Welfare work structures had been set up mainly to comply with federal requirements, not to manage the caseload. Only slowly, under pressure from FSA and TANF, with their work participation standards, did these states begin to try to enforce work effort within welfare. Among the UI and Rockefeller samples, these states include, for example, Alabama, Florida, Georgia, Mississippi, North Carolina, and Texas.

To summarize, one might say that in the first group of states, the capacity to form and execute a public will about welfare was fully realized, in the second it was partial and conflicted, and in the third nascent.

## Key issues

As one reads these case studies, three problems emerge as uppermost in the realization of reform. The first is *legislative decisions* concerning how tough to be about work expectations. The big issue is no longer benefit levels but rather how soon to require that welfare adults go to work. Legislatures tend to divide over time limits, over how much remediation to allow in place of work, and especially over sanctions. When a case violates work standards, TANF allows states to choose whether to exclude only the uncooperative adult from the grant or to end it entirely. States that are whole-hearted about work typically choose the full-family sanction. The more ambivalent states choose a partial sanction. Because the latter states often pay high benefits, many of the cases sanctioned simply accept the lower grant rather than meeting the work test or leaving welfare. This has kept California and New York from getting their large caseloads to take the work demand seriously. But it reflects these states' deep divisions about the welfare problem.

A second challenge is *program integration*. Most states are trying to simplify their administrative structures, at least as clients encounter them. They are trying, not only to raise child care funding, but to merge the various child care funding streams—some centered on welfare work and some not. This way, families do not have to shift programs as they leave welfare for work. The change seems vital to making subsidized care available to the working poor off welfare. Some states have managed

this, whereas others have been defeated by entrenched existing programs. A similar problem is combining various health programs for the welfare and low-income populations. A third instance is simplifying the child support enforcement structure, which in many states is divided between the welfare agency and the courts.

A third crux is *control of welfare work programs*. Historically, control has oscillated between state welfare and labor departments. The former have run cash assistance programs, including AFDC and now TANF, whereas the latter have run the Employment Service and federal training programs funded by the Workforce Investment Act (WIA). The Work Incentive (WIN) program, which was the earliest work program attached to AFDC, was run largely by labor departments, on the view that this agency was the more committed to the work mission. But labor department routines were built around serving voluntary, mostly male job seekers. These departments were commonly uneasy about enforcing work, as the welfare work mission required, and also about giving welfare families the special services and attention they often need to work. In short, they were unwilling to be paternalist. So starting with the Reagan reforms of 1981, federal policymakers shifted control of work programs to welfare departments, a process described above for New York. JOBS, the successor to WIN, was a welfare operation, and the employment agencies had no assured role, except as contractors to welfare departments.

Currently, however, many states are turning control of welfare work programs back over to labor departments. This is partly because WIA obligates labor departments to merge their own programs in local “job centers” that will serve all clientele in common. Policymakers reason that it is logical to include welfare work under the same aegis as well. And with half of recipients departed the rolls, mostly for employment, it is easy to think that now the voluntary structure of the employment agencies is sufficient to serve them.

But this bit of streamlining seems to be a mistake. Not only do labor departments have to develop new routines to serve welfare, but control of welfare work often shifts from the state to the local level, where workforce investment boards oversee the WIA structure. The changeover has sown great confusion in a number of states, among them Florida, Michigan, New Jersey, and Texas. I also believe that poor families continue to need special oversight and assistance if they are to work, even off welfare. That is still a mission that most regular employment agencies perform poorly. It seems better to let welfare agencies—or their contractors—continue to run welfare reform and postwelfare services, at least locally, even if the operation is put under a work-labeled umbrella agency at the state level in order to honor the work mission. This was the solution in both Wisconsin and New York.



## Political culture

The above patterns are deep-seated, not confined to welfare or even social policy. Thirty-five years ago, the political scientist Daniel Elazar published an analysis of state political cultures that remains highly useful today.<sup>14</sup> He argued that there were three strands in national political culture, which he called moralistic, individualistic, and traditionalistic. All were found everywhere in the nation to some extent, but states differed in which was strongest.

In the states where moralism dominated, politics had a strong public interest orientation. Politicians were expected to focus on problem solving rather than partisanship. Although they might disagree, all sides were expected to pitch their proposals to the community as a whole, rather than defending narrower interests. A second attribute of moralism was strong public administration, because this culture affirmed strong government. The civil service was merit-based and drew capable individuals into public service. The presiding spirit of moralism was high-mindedness. This culture prevailed in northern New England, the upper Midwest, and parts of the West and Northwest.

In individualistic states, in contrast, government was seen more as advancing the private interests of groups and citizens. Politics, like the economy, was an arena in which people strove to advance themselves. Strong parties represented coalitions that hoped to benefit from their election. Policymaking involved practical compromises among parties and interests and was less often rationalized in terms of a public interest. Government was large, but administrative quality was lower. Public service was seen as a secure career rather than a calling, and appointments often hinged on patronage. The presiding spirit of individualism was tolerance. This culture dominated the Mid-Atlantic states and the lower Midwest.

In traditionalistic states, politics was more elitist and government played a more limited role, chiefly to defend society against fundamental changes. Party divisions were less important than personalized factions that arose within a single governing class. Public administration was limited and distrusted. The presiding spirit was the defense of traditional values. Traditionalism was found mainly in the South and Southwest. Indeed, its chief original purpose had been to maintain the Southern racial caste system.

Elazar traced the roots of these subcultures chiefly to the ethnic and religious origins of the states. The moralistic states were first settled by Protestants from Britain, Scandinavia, and Germany who believed in self-government and held demanding views of personal and public morality. The individualistic region was settled by diverse groups from Britain, Germany, and elsewhere in

Europe who accepted democracy and capitalism but had more self-centered goals. The South was settled by more traditional elites who came to preside over a slave-based agricultural economy. The differences were enduring because later migrants to all these regions tended to adopt the same attitudes. The cultures are also distinct from other divisions in opinion such as political partisanship or liberalism versus conservatism.

The remarkable thing is how well Elazar's scheme predicts which states have dealt well with welfare reform and which poorly. First, let us define performance in process terms, to mean how well a state handled reform, whatever its goals were. Politically, high performance means a state acted rationally and cohesively according to its own goals, avoiding obvious missteps or debilitating divisions. Administratively, high performance means that officials "owned" the reform programs, had the resources to carry them out, and minimized internal divisions.

By these two standards, most of the states that have performed best in welfare reform fall within Elazar's moralistic group. The states in the middle of both distributions tend to be individualistic, whereas the low performers tend to be traditionalistic. There are degrees and exceptions, of course. The moralistic states did not do everything well. I have mentioned Wisconsin's shortcomings, and Michigan was one of the states that unwisely gave control of welfare reform programs to the employment agencies. California, although rated by Elazar as moralistic, today has such serious divisions about poverty that it falls well down in my rankings. Conversely, traditionalistic Tennessee performed well.

A more policy-oriented assessment is also possible.<sup>15</sup> I have modeled differences in state welfare caseload changes from 1989 to 1994, when caseloads generally rose, and from 1994 to 1998, when they mostly fell. These analyses show that states generally had higher growth or less fall in welfare caseloads if they had higher welfare benefits than the average, alongside several social and economic determinants. They also had less growth or more fall if they enforced work and child support more vigorously than the average and if they adopted full-family sanctions. These policy features, in turn, correlate with various background features of government and politics, including the Elazar subcultures.

Typically, the individualistic states have high benefits and have adopted only moderate sanctions, a combination that reflects their ambivalence about welfare reform. The traditionalistic states have both low benefits and strong sanctions, reflecting their time-honored approach to limiting dependency—simply by keeping people off the rolls. On one side you have the maintenance of entitlement, on the other the denial of any broad responsibility for the poor. The moralistic states, however, typically combine high benefits with strong work and child

support enforcement. That is, they affirm a strong government role while denying entitlement. They seek to help the poor while also demanding that the adults help themselves.

Each tradition contributes something to the process of welfare reform. In part, that process is about compromising opposed interests, as individualism emphasizes. In part, it is also about maintaining established values, such as the work ethic, as traditionalism emphasizes. But above all, it is about combining support for families with work expectations. That is the goal that all the states say they have adopted, to one degree or another. The objective is also fiercely popular. Large majorities of Americans reject traditional welfare because of its permissiveness, but still think that needy children and families deserve aid, and this view does not differ much by subgroup.<sup>16</sup>

To realize such policies takes governments able to run complicated social programs that define and enforce social values while also providing support. This is where the moralistic states excel. They appear most able to decide a public will about these difficult issues and then to carry it out with efficiency and conviction. That is the ultimate sense in which successful welfare reform depends on good government. Welfare reform is finally an exercise in statecraft. To succeed at it requires a difficult elaboration of public policies and structures. To do this calls upon deep reserves of belief in the public enterprise. That is just what the moralistic states have.

Time has overtaken the approaches to welfare that the other states favored. In the individualistic states, aid was seen as a payoff to an important constituency; hence the comfort of this tradition with entitlement. In the traditionalistic South, to maintain social discipline meant to minimize dependency and force the needy to fend for themselves. But these are just the attitudes that recent decades have taught the nation to question. After the burgeoning of dependency and related social problems in the nation's cities, aid given without questions regarding the behavior of the recipients has become indefensible. But after civil rights and welfare rights, and with today's levels of affluence, the idea of doing little for the poor is equally unpalatable. On both policy and political grounds, governments everywhere in the nation are driven to embrace the difficult goal of combining aid with requirements. The goal is to fuse help and hassle in a paternalistic regime. In that endeavor, the moralistic states, chief among them Wisconsin, have profound advantages.

## Improving government

But if political cultures have deep historical roots, as Elazar suggests, how can states move toward a more moralistic style? Elazar himself thought that the culture

as a whole was becoming more moralistic, albeit slowly. The civil rights era, in destroying Jim Crow, dealt a severe blow to the traditionalistic culture, which was now in retreat. Individualism, for its part, had had to accept the welfare state and thus become less distrustful of big government. In later editions of his book, through 1984, Elazar rated some states as more moralistic than he had before, although the dominant cultures he attributed to them did not change.<sup>17</sup>

In the 1960s, federal policymakers tried to inculcate a more high-minded ethos in state welfare agencies by requiring that they hire more college graduates as a condition of federal funding.<sup>18</sup> That effort was downplayed in the 1970s and 1980s, and the legal basis for it was repealed in PRWORA. Today, with devolution, Washington would never attempt to promote a moralistic style overtly. Yet it probably is still true, as Elazar said, that long-term trends favor greater high-mindedness. One force in that direction is rising education levels in the populace and among officials, another the rising role of public interest groups in national (and often state) policymaking, a third the influence of experts and researchers, particularly in social policy. All these trends promote respect for broad interests and on-the-merits argument as the basis for policy decisions, while deprecating appeals to narrower interests or values.

The irony is that welfare reform itself may be a leading force in this direction. For the past decade, probably no domestic issue has received more public attention than how to recast AFDC. The press—itsself a strong moralizing force—gave wide coverage to early experiments in welfare reform in California, Wisconsin, and elsewhere. These in turn helped trigger the promises by Bill Clinton in 1992 and by congressional Republicans in 1994 to end traditional welfare. These were the key commitments that led to PRWORA.<sup>19</sup> And since the enactment of TANF, equal attention has flowed to the effects of reform. States have produced studies of those leaving welfare, and there is competition among them to generate the largest and best effects. All this has forced them to approach reform in a moralistic style—as a problem in policymaking and administration. Ongoing national attention to the poverty problem will have the same effect in future. Thus welfare reform may help to generate its own mandate. ■

---

<sup>1</sup>L. Mead, "Welfare Employment," in *The New Paternalism: Supervisory Approaches to Poverty*, ed. L. Mead (Washington, DC: Brookings Institution, 1997), chap. 2.

<sup>2</sup>S. Scrivener, G. Hamilton, M. Farrell, S. Freedman, D. Friedlander, and others, *National Evaluation of Welfare-to-Work Strategies: Implementation, Participation Patterns, Costs, and Two-Year Impacts of the Portland (Oregon) Welfare-to-Work Program*, U.S. Department of Health and Human Services and U.S. Department of Education, Washington, DC, May 1998.

<sup>3</sup>M. Meyers, B. Glaser, and K. MacDonald, "On the Front Lines of Welfare Delivery: Are Workers Implementing Policy Reforms?" *Journal of Policy Analysis and Management* 17, no. 1 (Winter 1998): 1–22.

<sup>4</sup>L. Mead, "Welfare Reform in Wisconsin: The Local Role," *Administration and Society* 33, no. 5 (November 2001): 523–54. Resistance to work first also occurred in several of the case study states referred to below, for instance Arizona, California, and New York.

<sup>5</sup>Among mothers who were on AFDC in 1990, 57 percent worked in that year, rising to 65 percent by 1998, according to S. White and L. Geddes, *Economic Lessons for Welfare Mothers*, Wisconsin Policy Research Institute, Thiensville, WI, February 2001, pp. 5–7.

<sup>6</sup>*Snapshots of America's Families II*, Urban Institute, Washington, DC, October 24, 2000.

<sup>7</sup>L. Mead, "The Politics of Welfare Reform in Wisconsin," *Polity* 32, no. 4 (Summer 2000): 533–59.

<sup>8</sup>L. Mead, "Optimizing JOBS: Evaluation versus Administration," *Public Administration Review* 57, no. 2 (March/April 1997): 113–23.

<sup>9</sup>Mead, "Welfare Reform in Wisconsin: The Local Role."

<sup>10</sup>The following is based largely on my experience consulting with welfare officials in Albany and New York City. See also L. Mead, "Kicking New York's Dependency Habit," *City Journal* 1, no. 4 (Summer 1991): 41–49.

<sup>11</sup>Heather MacDonald, "Welfare Reform Discoveries," *City Journal* 7, no. 1 (Winter 1997): 14–31. This journal is published by the Manhattan Institute, a prominent conservative think tank in the city.

<sup>12</sup>The Urban Institute states include Alabama, California, Colorado, Florida, Massachusetts, Michigan, Minnesota, Mississippi, New Jersey, New York, Texas, Washington, and Wisconsin. The Rockefeller Institute states include Arizona, California, Colorado, Florida, Georgia, Kansas, Michigan, Minnesota, Mississippi, Missouri, New Jersey, New York, North Carolina, Ohio, Rhode Island, Tennessee, Texas, Utah, Washington, West Virginia, and Wisconsin. For the UI studies I made use of the baseline reports, which usually cover up to around 1997. For the Rockefeller studies, I read the first reports, also covering

up to around 1997, or summaries of these materials appearing in C. Weissert, ed., *Learning from Leaders: Welfare Reform Politics and Policy in Five Midwestern States* (Albany, NY: Rockefeller Institute Press, 2000), and S. Liebschutz, ed., *Managing Welfare Reform in Five States: The Challenge of Devolution* (Albany, NY: Rockefeller Institute Press, 2000). The assessments I make about the high or low performance of states based on these sources are entirely my own. The UI and Rockefeller projects do not make such judgments.

<sup>13</sup>M. Buis, "Taking States Seriously: Welfare Reform Implementation in Massachusetts and Oregon", Ph.D. Dissertation, Brandeis University, Department of Politics, Waltham, MA, 2001; Scrivener and colleagues, *National Evaluation of Welfare-to-Work Strategies: Portland (Oregon)*.

<sup>14</sup>D. Elazar, *American Federalism: A View from the States* (New York: Crowell, 1966). Later editions of this book were published in 1972 and 1984.

<sup>15</sup>The following is based on two articles by L. Mead, "Caseload Change: An Exploratory Study," *Journal of Policy Analysis and Management* 19, no. 3 (Summer 2000): 465–72, and "Governmental Quality and Welfare Reform," paper presented at the conference of the Association for Public Policy Analysis and Management, Seattle, WA, November 2000.

<sup>16</sup>M. Gilens, *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy* (Chicago: University of Chicago Press, 1999), chaps. 2, 8.

<sup>17</sup>Elazar, *American Federalism*, 3rd ed. (New York: Harper & Row, 1984), pp. 132–37.

<sup>18</sup>M. Derthick, *The Influence of Federal Grants: Public Assistance in Massachusetts* (Cambridge, MA: Harvard University Press, 1970).

<sup>19</sup>R. Weaver, *Ending Welfare as We Know It* (Washington, DC: Brookings Institution, 2000).

# Reauthorization and the data infrastructure: Can we do better this time around?

Robert Moffitt and Michele Ver Ploeg

---

Robert Moffitt is Professor of Economics, Johns Hopkins University, and an IRP affiliate. Michele Ver Ploeg is a Study Director at the Committee on National Statistics of the National Research Council.

---

The reforms of the nation's major social assistance programs during the 1990s fundamentally changed the way support is provided to low-income families with children. As policymakers consider the reauthorization of the reform legislation, there is little doubt that the general direction of the reforms—the promotion of work and the end of entitlement to assistance—will continue into the future. Besides focusing on funding levels for the renewed block grants to states that are at the center of the reforms, policymakers may consider further changes in policies to promote employment retention and advancement, reduce barriers to work, build strong families, and improve the well-being of low-income children.

Whether policymakers choose to make large or small changes in the 1996 legislation, a data and information system must be fully in place to continue current evaluation efforts and to prepare to evaluate the effects of further changes that Congress might make or that states might make on their own. Whether such a data and information system, a “data infrastructure,” is currently in place and, if not, what improvements are needed, is the subject of this article. Among the questions we address are:

- How has the existing data infrastructure performed in supporting monitoring and evaluation research to date?
- As Congress begins to consider reauthorization, do policymakers have enough information about the effects on families of existing reforms to know which policies need to be changed and which kept the same?
- Are the available data adequate to monitor the well-being of low-income families over time and to determine how different policy components have contributed to that well-being? If there are any inadequacies, how should they be addressed? If there are no inadequacies, are the kinds of data currently available likely to be adequate for the future?

We conclude that the data infrastructure has been, at best, only modestly successful in supporting the types of re-

search needed so that policymakers can be thoroughly informed about the effects of the legislation. Although the data infrastructure for welfare research is rich in many respects and the 1996 welfare reform legislation itself set up new and innovative data collection systems in several important areas, the coverage, quality, and availability of the current datasets still fall far short of what the United States as a society should demand. In part because the data are inadequate, there is, indeed, almost no good research on the effects of some of the key reforms that Congress is likely to consider in its reauthorization discussion. Major improvements are needed if analysts are to be in a better position to provide good information to Congress and the public the next time reforms of welfare policy are considered.

Our findings and recommendations in this article are partly based upon those of a report issued in August 2001 by the National Academy of Sciences (NAS). The NAS Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs was established in 1998. Sponsored by the Assistant Secretary for Planning and Evaluation of the U.S. Department of Health and Human Services, the panel was composed of 13 experts in welfare evaluation research and data. In its report, *Evaluating Welfare Reform in an Era of Transition*, the panel concluded that there are serious limitations in the nation's data infrastructure for monitoring low-income populations and evaluating the welfare programs targeted to them, even though the research on welfare reform has been unprecedented in its scope, volume, and diversity. Limitations in both national and state-level surveys, in administrative data, and in what is known about the policies states have actually enacted and implemented hinder efforts to thoroughly evaluate reforms that have already been passed and to inform reauthorization debates over what other reforms should be considered.

The bulk of our article summarizes the panel's findings and its recommendations for improving the nation's data infrastructure. In some cases we go beyond what is in the report, which was a consensus report of the panel's members, and insert our own views. We are careful to note the places where we do so.

Although the panel is critical of the capabilities of the current data infrastructure, it recognizes that the data challenges are great and the nation's governmental data administrators are in most cases doing the best they can with limited resources. With a reform as broad and deep as the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), which not only changed the rules of the program but has also devolved control of

the program from the federal level to the state level, improving the infrastructure will not be an easy or a short task. We do think, however, that it is important to consider steps that can be taken to improve the data systems now, as Congress begins consideration of reauthorizing PRWORA.

## Research questions of interest

The NAS panel recognized that data needs are necessarily driven by the questions to which policymakers and researchers need answers. The panel broadly outlined the questions of interest for welfare reform research by considering all types of major questions in which significant segments of the public and government have an interest.

The first distinction the panel made was between monitoring questions and evaluation questions. Monitoring questions are those that seek to track the well-being of the low-income population, or the well-being of families who have left welfare, are still on welfare, or have been discouraged or prevented from entering welfare. Answering these questions requires data over time on the relevant groups of interest. But because monitoring studies do not attempt to rigorously assess the contributions of different policies or other causes to how those outcomes have evolved, data structures can be relatively simple. The panel concluded that monitoring questions are an essential part of welfare reform research that have been and will continue to be important.

Evaluation questions, in contrast, impose greater demands on data because they require an assessment of cause and effect. To answer evaluation questions, researchers must determine a counterfactual—the value of an outcome that would have occurred if a policy had not been implemented. The counterfactual can then be compared to the value of the outcomes that have occurred after the policy is actually implemented. One approach to answering evaluation questions is to use random assignment studies. The panel urged that such studies be continued and in fact increased in the future for the analysis of certain types of questions. But it also concluded that such studies have both inherent and practical limitations that have permitted them to answer only a restricted set of questions so far. Nonexperimental methods must necessarily be a major part of the evaluation toolkit, and these require the right type of high-quality data.<sup>1</sup>

The panel determined that there are three broad types of evaluation questions in which different audiences are interested: (1) What are the overall effects of welfare reform, taking the entire bundle of changes in policies, programs, and practices? (2) What are the effects of individual, but still broad, components of welfare reform (e.g., work requirements, sanctions, time limits, family caps)? (3) What are the impacts of individual, detailed

strategies within each of the broad components (e.g., type of work strategy, specific cash assistance level, nature of sanction policy)?

The panel concluded that the amount of research that has been devoted to these evaluation questions has been considerably less than that devoted to monitoring, and that this represents a misallocation of research attention. In addition, it concluded that most evaluation research had concerned itself with the first and third questions, that is, with the overall effects of welfare reform and with the effects of individual detailed strategies. There has been very little evaluation of the broad components of reform.

## Types of data

There are five types of data that have been and will continue to be useful for these monitoring and evaluation questions. These are (1) national survey data, (2) state survey data, (3) administrative datasets, primarily at the state level, (4) qualitative data, and (5) data on policies actually adopted. We will discuss each of these in turn.

### National survey data

The major national household surveys which are useful for welfare reform research include the Survey of Program Dynamics (SPD), Survey of Income and Program Participation (SIPP), March Current Population Survey (CPS), Decennial Census Long Form, American Community Survey (ACS), the Michigan Panel Study of Income Dynamics (PSID), the National Longitudinal Surveys (NLS), and the National Survey of America's Families (NSAF). These surveys cover a wide variety of time periods and contain different types of information; some are longitudinal and some are cross-sectional, and many have specific welfare reform content.

These surveys provide the basis for a rich analysis of welfare reform, at least for monitoring, if not for evaluation. But collectively they have limitations which have constrained their contributions to research on welfare reform. These include their survey content, response rates, statistical power for detecting impacts of reform, and timeliness in release.

Survey content is clearly central to obtaining information on welfare reform. All of these surveys contain, to at least some degree, basic information on income, family structure, welfare program participation, and other behaviors.<sup>2</sup> However, problems have emerged in most of these national surveys because the devolved program environment has made eliciting information on welfare program participation an increasingly complex task. Names of welfare programs vary from state to state and sometimes within states, noncash programs have grown to rival cash programs, and the very concept of welfare participation is more ambiguous than it once was. Each

of these factors makes it difficult for a national survey to ask the right questions about participation in what are now more localized programs. In addition, determining the actual services that those on welfare have received and the requirements, sanctions, and other regulations they have faced is even more difficult.<sup>3</sup> Limitations also exist in obtaining detailed monthly information on labor force participation and earnings of recipients in an era when turnover is quite high, and determining what barriers exist to employment—in health, transportation, and child care, to mention only three—is a particularly acute problem. These dimensions of current household surveys can potentially be greatly improved. The panel recommended regular and frequent review of national household survey questions regarding participation in welfare programs to ensure the surveys are up to date in capturing participation in state welfare programs. The development of mechanisms for regular communication between the Census Bureau, the Department of Health and Human Services, and the states to discuss how surveys can ask appropriate questions was also recommended.

Obtaining high response rates has been a major challenge in most of these surveys as well. Probably the most serious problem has occurred in the SPD, which is likely to suffer a 50 percent attrition rate over the life of its panel. But nonresponse problems are also present for the SIPP, the PSID, and the NSAF. Whether the nonresponse that has occurred biases estimates and when corrections for nonresponse can be made is a complex statistical question, but high rates of nonresponse create a *prima facie* layer of uncertainty about the quality of the dataset. There are ways to increase response rates by using better field procedures, incentive payments, and other mechanisms.

Statistical power for detecting the counterfactual effects of different policies is a serious problem for all datasets except the CPS and the ACS, which have the largest sample sizes of those listed.<sup>4</sup> A study prepared for the NAS panel by John Adams and V. Joseph Hotz concluded that the CPS has only barely adequate sample sizes for the detection of reasonable overall reform impacts using cross-state comparison methods, the most common method in the literature. Even the CPS has low power to detect the effects of individual broad welfare reform components—the second evaluation question of interest identified by the panel. Necessarily, the other datasets, which have much smaller samples than the CPS, have very little power for the detection of policy impacts.

Timely release has been a barrier to the use of many of the surveys for welfare reform research. The SPD, SIPP, and PSID have been released very slowly; only recently have post-1996 data which can be used to assess the impact of welfare reform become available. Often this is the result of insufficient resources and overworked staff given other priorities. Whatever the reason, the fact is that neither the SPD nor the PSID, and to a lesser extent

the SIPP, have played a significant role in the generation of research findings on welfare reform and they are unlikely to do so in time for the reauthorization discussion in Congress.

Probably the most difficult issue is that of statistical power, which requires larger sample sizes. The ACS holds promise for this purpose, as do increases in CPS sample size. The other datasets will have to be used to answer a more limited set of questions or researchers will have to use alternative statistical techniques to circumvent the problem of low statistical power. Nevertheless, it is possible to address these problems. Much can be done to improve the ability of the household datasets to deal with the other issues we have discussed.

We believe that the dataset with the most potential for improvement at this stage is the SIPP, an issue that was not addressed in the NAS report. The major disadvantage of the SIPP is the limited length of its panels, but it does have many advantages. It has reasonably large sample sizes which can be increased by pooling across panels. It also has a monthly accounting time frame which more easily measures labor force and welfare dynamics, and which is begun anew periodically—allowing the incorporation of improved response rate procedures and survey content on a regular basis. With more attention and resources devoted in this direction, with a more timely release of data, and with increased sample sizes, the SIPP could become the overall best national-level dataset for welfare reform evaluation.

### **State surveys**

State surveys are in their infancy and have so far contributed in only limited ways to welfare reform research. Historically, state governments have not conducted much survey research. This responsibility has been thrust upon them by the devolution of program responsibility—and, to a great degree, responsibility for research—down to the state and local levels. Conducting surveys of the low-income and welfare recipient populations, which is a difficult task requiring significant resources and high levels of expertise (as demonstrated by the experience of national household surveys) is a challenge for states. Most state surveys have taken place as part of studies of welfare “leavers,” that is, families which have left the welfare rolls in the wake of welfare reform. Most of these have been telephone rather than in-person surveys and they have encountered significant problems of coverage, nonresponse, and survey content. Nevertheless, they represent a promising start.

Seeing a need for improvements in state-level surveys, the NAS panel organized a workshop on methods for improving state household surveys that brought together experts in survey methodology and state welfare program administrators and researchers. Numerous suggestions for improvements in data collection were made; all of

them are part of a forthcoming volume of papers from that workshop.<sup>5</sup> Given that devolution will surely continue past reauthorization, improved state ability to conduct surveys is a high priority. In its report, the panel recommended investments in building the capacity of states to conduct household surveys.

### **Administrative data**

Administrative data, or data collected as part of the administration of programs and services, are increasingly being used for monitoring and evaluation. For welfare program research, administrative data from TANF, Food Stamps, and Medicaid are used to measure characteristics of recipients and their benefits, and Unemployment Insurance administrative data are frequently used to measure earnings and employment. Administrative data can provide valuable descriptive information about program participants for monitoring purposes and they are usually integral parts of experimental program evaluations. Because administrative data are only available for those who use a program, they are less often used in nonexperimental evaluations, unless they are linked with other data sources.

A key advantage of administrative datasets is that they are often quite large and dwarf the sample sizes of most surveys. They are also relatively inexpensive, available on the state level for analysis of state programs, and potentially available relatively quickly. Further, state program administrators usually have a great deal of knowledge about these datasets and are comfortable using them. However, administrative data are collected to run a program and are not readily available for use in research. Because coverage is limited to those who receive the program, such data are problematic when research questions concern broader populations of interest. The content of administrative databases is also limited to the few variables collected during the course of administering the program. Because of this, databases are often linked with other administrative databases or with coordinated survey data collections. Linking data is not always straightforward, however, and issues of confidentiality and access can be difficult.

When the goal is cross-state analysis, the comparability of administrative data across states is often a problem. Each government tends to define variables in its own way and to collect data in a form suitable to its primary goals. Yet cross-state comparisons are one of the most important nonexperimental evaluation methods, and pooled administrative data across states could solve the difficulties of sample size and statistical power noted above. For administrative data to be used in this way—to date, they have generally only been used for within-state studies—would require some coordination and cooperative agreements among the states to generate data that are sufficiently comparable to make this type of analysis feasible.

State administrative data have great potential for studying welfare reform, but will need to be developed and supported for that potential to be realized. The panel recommended several capacity-building steps for state-level administrative datasets. In particular, states need technical assistance and funding to develop, clean, and improve the quality of administrative data systems. Such assistance is also needed to help states link separate administrative and survey datasets and to negotiate data-sharing agreements among agencies. The panel also recommended that the federal government should work with states to help build comparable data definitions and formats so that cross-state analyses can be conducted.

### **Program description data**

Devolution has led to a flowering of local innovations in social welfare programs, and there is evidence that states are using their discretion to design welfare programs in many different ways. Specific information about the rules of each state's programs has, however, been slow to emerge. Such information is not only an outcome of interest in itself, since policymakers want to know what kinds of variation PRWORA has spawned, but is also necessary to evaluate which programs work and which do not. The absence of descriptive data on program rules has been a major limitation, especially for cross-state evaluation of welfare reform. The task is difficult because rules in each state and jurisdiction are complex. Further, the way rules are implemented in local offices may be particularly hard to ascertain and rules for programs other than TANF are also relevant. If high-quality data concerning program rules are to be collected, substantial resources must be committed.

The best data of this type have thus far been assembled primarily by private organizations, with limited help from the federal government and major assistance from the private sector. But private-sector support is episodic, and hence the projects have not always had a solid commitment to permanence. The panel concluded that collecting data on rules is a federal responsibility because of its general importance to all states and all members of the public and the research community. It recommended more federal support for systematic and continual collection of program description data.

### **Qualitative data**

Although most of the recommended improvements are directed toward quantitative data collections, the panel also emphasized that qualitative data sources are a vital part of the data infrastructure. Qualitative data, whether ethnographic studies that seek to provide more detail on the lives of low-income families, or process studies that attempt to assess exactly how state programs are implemented, can provide added texture to what is known about the programs and the people upon whom they are targeted. The panel concluded that these data sources

have thus far been underutilized in research on low-income populations, but that the development of standardized documentation of data collection methods and more inclusive representation of groups and areas would improve this segment of the data infrastructure.

## Conclusions

The enactment of PRWORA was accompanied by much public fanfare and media attention. Yet regrettably, more than five years after the legislation was passed, very little is known about the effects of many of its important components. A mature society like the United States should be expected to evaluate the effects of legislation that it passes and the policies that it enacts.

What further steps need to be taken to improve the nation's data infrastructure for evaluating the effects of PRWORA and any future policy changes? In addition to its recommendations regarding data collection methods, the panel recommended administrative changes in how these data are collected. Currently, no single federal statistical agency has the authority and responsibility to collect and develop data on low-income populations and the programs that serve these populations. Human services does not have a statistical agency comparable to the Bureau of Labor Statistics, the National Center for Health Statistics, or the National Center for Education Statistics. This lack of federal authority and leadership has been especially detrimental in a time when more coordination of data collection activities and development of data sources is needed to evaluate the multiplicity of state and local programs.

To begin to remedy this gap in authority, the panel recommended that a new or existing agency take over responsibilities to collect data needed for program monitoring and evaluation. The panel did not make recommendations about the form of that agency, but it did outline its responsibilities, including:

- Sponsorship of national surveys of social welfare programs and low-income populations.

- Development of administrative data for research purposes, including standard definitions and data formats; leadership in linking administrative data to other sources of data.
- Provision of technical assistance to states in data collection for monitoring and evaluation.
- Production of periodic reports on the utilization of social welfare programs.
- Leadership in developing data archives on topics for use in social welfare program evaluation and research.

Data for monitoring and evaluating social welfare programs will continue to be needed and will change as new legislation is passed and as programs change. A permanent data infrastructure must be built up so that it is in place for the next round of reform. We believe that reauthorization is an opportune time to begin on this task. ■

---

<sup>1</sup>Monitoring questions are, of course, inherently a nonexperimental exercise and hence must also be built up from nonexperimental data.

<sup>2</sup>One problem, for example, is that only some of the surveys attempt to include information on child outcomes and even then not always with very extensive modules.

<sup>3</sup>These issues all pertain to eliciting a family's actual welfare experiences. A separate but related issue is determining the program rules that each family faces in its local area. None of the datasets has matched information on welfare program rules in the local area of residence to their surveys (except, that is, for crude state-level rules). Such matching would be a major improvement in the capacity of the data to address evaluation questions.

<sup>4</sup>We ignore the Decennial Census Long Form microdata, which we list only because of its potential. It has not been used to date because the 2000 data are not yet available.

<sup>5</sup>National Research Council, *Studies of Welfare Populations: Data Collection and Research Issues*, ed. M. Ver Ploeg, R. Moffitt, and C. Citro, Committee on National Statistics, Division of Behavioral and Social Sciences and Education (Washington, DC: National Academies Press, 2002).