

**Policy Research:
The Field Dimension**

Lawrence M. Mead
Department of Politics
New York University
726 Broadway, #765
New York, NY 10003-9580
Phone: 212-998-8540
Fax: 212-995-4184
E-mail: LMM1@nyu.edu

Prepared for delivery at the conference of the
Association for Public Policy Analysis and Management
Atlanta, GA, Oct. 30, 2004.

ABSTRACT

Field research, meaning unstructured contact with public problems and programs, is essential to realistic policy research. Research tying governmental action to good outcomes is rare, because those who study government and those who analyze public problems are seldom the same. Field inquiry can help give more governmental content to policy research. A lack of field contact was one reason why much of the research surrounding welfare reform proved to be incorrect. Ideally, the connections between policy and outcomes that respondents claim during field research should be verified by statistical analyses using program data. Unfortunately, field research is discouraged by academic incentives favoring rigor at the expense of realism.

Introduction

Field research is an essential component of realistic policy research.¹ By “policy research” I mean inquiry into the nature and origins of problems that public policy aims to solve. Policy research is distinct from purely academic research that seeks only theoretical knowledge. It is also distinct from policy analysis done to inform a specific decision, but it supports such analysis by identifying the causes of problems that policy might change.²

By “field research” I mean inquiry into programs or policies through direct contact, such as by reading government documents, observing operations, inspecting program data, or interviewing clients or staff—what I call laying hands on the institutions. Field research emphasizes unstructured learning about a program, and also serendipity—discovering the unexpected. It is guided only to a limited extent by prior hypotheses. Above all, it is unstructured by prior research or established data bases. If it includes quantitative analyses, these will address hypotheses derived from the field, and will be based on program data.

A paucity of field inquiry has weakened recent policy research on the problems of poverty and welfare. Inquiry into these questions consists mostly of statistical models estimated from academic or government data bases, without a hands-on sense of what is going on.³ Such modeling shortchanges the political and administrative forces that often drive change. These forces dominated welfare reform, meaning the institution of work tests in AFDC/TANF in the late 1980s through the 1990s.⁴ As a result, some social research omitting these factors has been erroneous.

¹ This paper builds on these earlier publications: Lawrence M. Mead, "Welfare Policy: The Administrative Frontier," *Journal of Policy Analysis and Management* 15, no. 4 (Fall 1996): 587-600; idem, "Comment on Mark R. Rank and Thomas A. Hirschl, 'Rags or Riches'," *Social Science Quarterly* 82, no. 4 (December 2001): 670-5; and idem, "Performance Analysis," in *Policy into Action: Implementation Research and Welfare Reform*, ed. Mary Clare Lennon and Thomas Corbett (Washington, DC: Urban Institute Press, 2003), chap. 6.

² Erwin C. Hargrove, *The Missing Link: The Study of The Implementation of Social Policy* (Washington, DC: Urban Institute, July 1975), pp. 9-11.

³ So it seems based on my own experience of poverty research as it appears at major conferences and in the leading publications. Data bases certainly dominate social research at such centers as the Institute for Research on Poverty at the University of Wisconsin--Madison, other major universities, or the leading think tanks in Washington. But I have done no formal survey.

⁴ The changes were driven by the Family Support Act of 1988 and the Personal Responsibility and Work Opportunity Act of 1996 (PRWORA), although in some states work demands went back earlier. Aid to Families with Dependent Children (AFDC) was recast as Temporary Assistance for Needy Families (TANF) by PRWORA.

Below, as background, I first situate public policy research within political science. The policy field has failed to meet expectations, I argue, because of a lack of governmental content, among other problems. More field inquiry is one means of changing that. I then give several instances where field research in social policy brought into question findings based on other methods. I also discuss ways of making field inquiry more rigorous. Finally, I describe the academic incentives that currently discourage field research.

Public Policy Research

I understand public policy as a subfield within political science. It is often located within American studies, but it need not be. Its defining feature, in my view, should be that it seeks to improve actual public policy.⁵ The ancients treated politics as the master science, meaning the highest means by which communities could achieve the good life. Politics had value in itself, but principally it was a means for realizing the good society.

To fulfill that vision, political science must show how government can promote good outcomes for the society. In the policy field, political science becomes a form of policy analysis—a reasoning that uses government to achieve an improved society. Public policy seeks to understand government and politics, but ultimately as a means to this further end. Much in this spirit, the earliest thinking about public policy, by Harold Lasswell and others, imagined it to be a rather catholic form of reasoning about how to reach public goals, which themselves had various meaning.⁶

The general model of public policy is thus:

(1) Government —————> The good society

With regard to a specific problem, policy research must show which public policies will achieve good effects, however “good” is understood:

(2) Public policy —————> Good effects

⁵ In his prospectus for this panel, Hank Jenkins-Smith has implicitly taken this position by asking what various approach to policy research imply for “influencing or shaping public policy.”

⁶ Daniel Lerner and Harold D. Lasswell, eds., The Policy Sciences (Stanford, CA: Stanford University Press, 1951); Harold D.Lasswell, "The Emerging Conception of the Policy Sciences," Policy Sciences 1, no. 1 (Spring 1970): 3-14.

For this connection to hold, however, government must be able to generate the desired policies.

That means both making policy and implementing it:

(3) Government \longrightarrow Effective public policy

For government to do this assumes that it has effective policymaking processes. These will include some form of policy analysis, or deliberate reasoning about how beset to solve public problems, and also administrative mechanisms. Policy research seeks general knowledge of problems so as to support analyses done for specific public decisions. Only if model (2) is combined with (3) can research realize the vision of model (1), the master science.

The policy field has not approached this ideal, chiefly due to developments within social science. The most fundamental problem is that most political scientists do not embrace equations (1) or (3). They do not appraise government in terms of policy performance. They are interested in politics or government for themselves, not as means to policy ends. Policy is of interest to them only as a scorecard for the political process; it has no importance in itself.

Most political scientists who are policy specialists study the policymaking *process* rather than how to improve policy. Their modeling is increasingly sophisticated, but it is not different in kind from what other scholars do outside the policy field.⁷ Policy research may focus on policymaking, whereas other specialists in American studies focus on the presidency or Congress. But neither group is interested in improving policy as such. This version of public policy research is not policy analytic. Only a handful of political scientists are known as policy experts.⁸ For this reason, most policy research done by political scientists is of little interest to actual policymakers.⁹

Conversely, most of those who do seek to improve policy in the sense of equation (2) take no interest in government. Most experts in specific policy areas such as poverty, national security, or

⁷ Kim Quaile Hill, "In Search of Policy Theory," *Policy Currents* 7, no. 1 (April 1997): 1-9; Paul A. Sabatier, ed., *Theories of the Policy Process* (Boulder, CO: Westview, 1999).

⁸ In social policy, those few would include myself, John DiIulio, Charles Murray, and James Q. Wilson. Wilson, of course, also is prominent as a political scientist outside the policy field, a measure of his unusual achievement.

⁹ James M. Rogers, "Social Science Disciplines and Policy Research: The Case of Political Science," *Policy Studies Review* 9, no. 1 (Autumn 1989):13-28; Dennis J. Palumbo, "Bucking the Tide: Policy Studies in Political Science, 1978-1988," in *Policy Studies Review Annual, Volume 10: Advances in Policy Studies Since 1950*, ed. William N. Dunn and Rita Mae Kelly (New Brunswick, NJ: Transaction Publishers, 1992), chap. 2.

the environment have backgrounds in economics or some other technical subject. They study policy problems and make recommendations for their solution. They are usually more comfortable doing this than political scientists are. But government itself is not part of their analysis. That is, features of government or politics are not usually given as reasons to prefer one course of action to another, nor are changes in government proposed to make possible better policies.

Thus, while political scientists typically study government without telling it what to do, policy experts tell it what to do while ignoring it. Neither group makes the connection shown in equation (3). But without this, policy reasoning consists only of equation (2), and the master science suggested by equation (1) cannot be realized.

These tendencies lead to the bifurcation of the policy literature as a whole. There are works about how government makes policy, written almost entirely by political scientists, and there are works about how to handle various policy problems, written almost entirely by economists and other policy experts. There is very little research that tries to do both, first making arguments for policies on-the-merits, as suggested by equation (2), and then adding governmental analyses about how to generate such policies, as suggested in equation (3). Only this small literature even attempts to realize equation (1).¹⁰

One reason for the division is academic specialization. To connect political with policy analysis takes a wider range of knowledge than most academics claim today. Perhaps only senior scholars with unusually broad backgrounds can do this. To practice the master science, it also helps to have government experience, but most policy researchers never work outside the university or think tanks. Another cause is the longstanding reluctance of political scientists to take stands on policy issues, lest they seem to trench upon the democratic process.¹¹

¹⁰ Examples include my books on poverty and welfare reform (e.g., Beyond Entitlement: The Social Obligations of Citizenship, New York: Free Press, 1986) and Allen Schick's books on the federal budget problem (e.g., The Capacity to Budget, Washington, D.C.: Urban Institute Press, 1990). But the best example is Derek Bok's ambitious study of the federal government. In The State of the Nation: Government and the Quest for a Better Society (Cambridge, MA: Harvard University Press, 1996), he appraises the government's performance in policy terms, and in The Trouble with Government (Cambridge, MA: Harvard University Press, 2001) he traces its problems back to features of the regime and its politics.

¹¹ Richard R. Nelson, The Moon and the Ghetto (New York: Norton, 1977).

A more recent cause is the tightening grip of the journals on political science. Peer-reviewed research places the value of rigor and precision above all others. Policy research addresses important issues and, potentially, has a large audience, but these values are secondary to reviewers. Any inference connecting government to policy outcomes is always more judgmental than inferences about political behavior alone. So most political research today focuses on the precise modeling of processes within government, neglecting effects on the society. The recent vogue for game theory and other quantitative methodologies has intensified these trends. As a result, political science is retreating from relevance to policymaking. Law and sociology are expanding to fill the void.

Governmental Content

Another way to put the problem is that most policy research lacks governmental content. That is, the characteristics and influence of government itself, as against its policies, is not part of the analysis. This shows up in two ways.

Institutional constraints

First, as mentioned, most research that recommends best policies is confined to the logic of equation (2): It argues what sort of policy would best solve some problem, but there is no reasoning in the form of equation (3). That is, there is no serious discussion of how politics or administration might constrain or expand the options available to decision makers.

Policy reasoning that is comprehensive should discuss policies that are optimal, but then *go on* to discuss what options are actually politic or administrable.¹² The effect is to include political and bureaucratic factors as part of policy analysis. Often, there is tension between what one would like to do “on the merits” and institutional constraints. In order to improve policy outcomes, one might

¹² The following is based largely on Lawrence M. Mead, "Policy Studies and Political Science," Policy Studies Review 5, no. 2 (November 1985): 319-35; and idem, "Public Policy: Vision, Potential, Limits," Policy Currents 5, February 1995: 1-4.

propose changes in the institutions. But equally, in light of political reality, one might alter one's goal for what government should do.¹³ There is an interaction of policy and political reasoning:

(4) Policy \longleftrightarrow Politics

Policy reasoning becomes a form of systems analysis that seeks to reconcile goals with means, what Dror called “metapolicymaking.”¹⁴ At its best, such work approaches equation (1) and the master science.

A lack of this governmental dimension has seriously weakened social policy analysis. Repeatedly, antipoverty experts have recommended expanded benefits for the poor, only to have the proposals dismissed out of hand in Washington as impolitic. This was true of 1960s proposals for a negative income tax, later embodied in the Nixon and Carter welfare reform plans, which were rejected by Congress. The same fate befell proposals for public jobs programs in the 1970s; a limited expansion occurred under the Comprehensive Employment and Training Act (CETA), but it was ended in the Reagan budget cuts of 1981. In the 1980s, proposals for an assured form of child support, which might replace welfare, failed to advance, and in 1991 the National Commission on Children's plan for a universal and refundable child tax credit received no serious consideration.¹⁵

Policy researchers tend to blame conservative opposition.¹⁶ However, the defeats occurred when liberals and Democrats controlled more of Washington than they do today. Rather, advocates focused on the benefits of their proposals to the intended recipients. They failed to take seriously the preferences and priorities of the rest of the community. Politicians resisted partly on grounds of cost, but also because the proposals slighted public concern that the working-aged needy do more to help themselves. Some of the proposals created new opportunities and incentives for employment; none required it. Even the public jobs proposals treated work as a benefit to be given to the poor, not

¹³ The need to limit goals in social policy to those government could achieve was a theme of Aaron Wildavsky, Speaking Truth to Power: The Art and Craft of Policy Analysis (Boston: Little, Brown, 1979).

¹⁴ Yehezkel Dror, Public Policymaking Reexamined (Scranton, PA: Chandler Publishing Co., 1968), chap. 14.

¹⁵ Irwin Garfinkel, Assuring Child Support: An Extension of Social Security (New York: Russell Sage Foundation, 1992); National Commission on Children, Beyond Rhetoric: A New American Agenda for Children and Families (Washington, DC: U.S. Government Printing Office, 1991).

¹⁶ David L. Featherman and Maris A. Vinovskis, eds., Social Science and Policy-Making: A Search for Relevance in the Twentieth Century (Ann Arbor: University of Michigan Press, 2001).

as an obligation that they owed in return for aid. CETA also failed due to serious administrative problems and a lack of positive evaluation findings.¹⁷ By the end of the 1970s, further progressive programming to help the poor seemed unattainable.¹⁸

But if well used, governmental forces can expand options as well as limit them. In the 1990s, policymakers began to use welfare to promote employment rather than avoid it. This policy was so popular that it permitted spending more on welfare and related services than ever before. Substantial sums were invested in welfare work programs, state-level work incentives, the Earned Income Tax Credit (EITC), and child and health care for families going to work. Mandatory work programs tied to welfare became far larger and costlier than voluntary training programs. Virtually a new welfare state grew up around promoting and enforcing employment for the working aged. Had experts and their political allies appreciated the potential, these changes might have occurred ten or twenty years earlier than they did.

Why are institutional constraints and potentials commonly ignored in policy research? Part of the reason is political. Most policy experts are left of center, particularly those in the universities. Most find distasteful the recent vogue to enforce work and other good behaviors on the poor. If endorsing that policy is the price of expanding aid for the disadvantaged, some will refuse to pay it. But more important, probably, are disciplinary constraints. As mentioned already, political scientists understand government best, but they usually eschew policy arguments. Conversely, the experts who do make policy arguments tend to come from disciplinary backgrounds that neglect institutions.¹⁹

Especially if they are economists, policy experts tend to reify government into a unitary actor who already has the authority to make policy. This “Model 1” assumption, as Allison called it, construes the policymaker’s problem as choice rather than power. The policy analyst then advises

¹⁷ Mead, Beyond Entitlement, chaps. 3, 5.

¹⁸ Henry J. Aaron, Politics and the Professors: The Great Society in Perspective (Washington, DC: Brookings, 1978); Peter H. Rossi, "The Iron Law of Evaluation and Other Metallic Rules," Research in Social Problems and Public Policy, ed. Joann H. Miller and Michael Lewis (Greenwich, Conn.: JAI Press) 4 (1987): 3-20.

¹⁹ I realize that there is a “positive” economics that seeks to explain politics and bureaucracy, but it is seldom practiced by the economists who study policy issues. One may also question whether economists’ views of how government works are realistic.

the decision maker about how best to choose. In fact, in a Madisonian political system with dispersed initiative, policymakers must first assemble the power to act. At best, that means defining public problems so that they cut with the grain of politics and the bureaucracy rather than against. At worse, it means compromising with political opponents or bureaucratic incapacity so that at least something can be achieved.²⁰ How to make these judgments is seldom analyzed explicitly.

Institutional influences

So policy research in the style of equations (3) or (4) is rare. Government also tends to be missing from policy research in the style of equation (2). This is a more serious omission, because it affects our very understanding of the problems that government aims to solve. In social policy research, the dependent term is usually some condition affecting people who are poor or disadvantaged, and the explainers are various factors thought to determine that condition.

Social policy research typically specifies a model like this:

$$(5) \quad P = \alpha + \beta_1 \mathbf{E} + \beta_2 \mathbf{S} + \beta_3 \mathbf{B} + \varepsilon$$

Here P is the problem of interest. Its explainers include \mathbf{E} , a vector of economic conditions; \mathbf{S} , a vector of social or demographic conditions; and \mathbf{B} , a vector of benefits provided by government. The unit of analysis could be geographic or governmental units such as Census tracts, in which case all the variables would describe these units. If the unit were the individual, \mathbf{E} might be demographic features determining employability, \mathbf{S} other attributes of interest, and \mathbf{B} benefits received.

Benefits may include welfare, Food Stamps, the EITC, and so on. But governmental influence in any other sense seldom appears. Especially, we seldom see measures for whether the locality enforced work or child support, or whether an individual was subject to these or other suasions of a noneconomic kind. But since these factors have a lot to do with the behavior of the poor and dependent today, the models are underspecified. I give examples below.

Again, the causes include the reluctance of most academics to discuss behavioral enforcement as well as the neglect of institutions in economics and other technical disciplines. But the largest cause may be the simple fact that the surveys used by social policy researchers seldom measure

²⁰ Graham T. Allison and Philip Zelikow, *Essence of Decision: Explaining the Cuban Missile Crisis*, 2nd ed. (New York: Longman, 1999).

institutional influences. Data bases such as the Census or the Panel Study of Income Dynamics (PSID) include data on whether an individual or family received income from welfare or other programs, but not on whether they faced pressure to work, or pay child support or do anything else in return. Such data bases promote a view of social problems in which disembodied individuals respond to impersonal social and economic forces. That vision largely omits the political and programmatic context that today is driving change, especially in connection with welfare reform.²¹

Even the surveys designed to track the effects of reform omit governmental factors other than benefits. This is true of the government's Current Population Survey (CPS), the Survey of Income and Program Participation (SIPP), and the Survey of Program Dynamics (SPD), the latter an expansion of SIPP focused on the effects of welfare reform. It is also true of the National Survey of America's Families (NSAF), a private survey run by the Urban Institute (UI) that focuses on family status. UI's project on Assessing the New Federalism, of which NSAF is a part, also includes case studies of how thirteen states implemented PRWORA and associated child care and health policies. These cases have a good deal of governmental content, but they are detached from the survey used to analyze the effects of reform.²²

Political research on policymaking that lacks a policy argument draws little interest outside political science. Orthodox policy research in the mode of equation (2) generally draws more interest, but it too is often ignored by policymakers. One reason is that it makes a policy argument, but it says little directly *about government*. The models may suggest that certain changes in a public policy or benefit would improve the condition of some social group. But the analysis usually does not show how to enact or implement that change, and these are public officials' chief concerns.

²¹ For a rare attempt to add indicators of state "requirements" to individual-level data from the National Longitudinal Survey of Youth, see Robert D. Plotnick, "Welfare and Out-of-Wedlock Childbearing: Evidence from the 1980s," Journal of Marriage and the Family 52 (no. 3, August 1990): 735-46.

²² There are also 21 case studies of TANF implementation done by the State Capacity Study at the Rockefeller Institute of Government at the State University of New York. These cases are even richer in governmental content, but no clear connections to policy outcomes are made. This in part reflects the fact that the project is run by political scientists. However, I used both the UI and Rockefeller cases to try to make a connection between governmental features and at least the implementation of TANF. See Lawrence M. Mead, "State Political Culture and Welfare Reform," Policy Studies Journal 32, no. 2 (May 2004): 271-96.

Field Research

The virtue of field inquiry is that it restores governmental content to policy research. It gives a more realistic picture of the forces surrounding social problems, and this can open the way to fresh solutions. Field contact can break the grip that survey-based inquiry has on policy research today, at least in social policy.

Survey research

Survey-based research generates accurate depictions of the clientele served by social policy, especially the poverty or welfare populations. If samples are large, surveys describe social problems and their correlates reliably, and breakdowns by subgroups are possible.

But surveys by themselves simply measure the economic and social attributes of people. Research based on surveys tends to be atheoretical, embodying little understanding of why social problems arise.²³ It is in fact difficult to account for much of the behavior of the seriously poor and dependent. Why do many appear to pursue lifestyles that defeat their own ends; especially, why do few poor adults work regularly? Confronted with these mysteries, much of poverty research responds with description. It portrays the poor or dependent population in withering detail. But it does not explain their situation.

For want of any more specific theory, survey-based research is often economicistic. It projects onto its subjects the psychology assumed by economics. The people surveyed are said to act so as to maximize their utilities, defined usually in terms of material interests such as money. They respond to “incentives” in the way we assume about economic actors. They will work or not according to the substitution and income effects. Their capacities or incapacities to work represent “human capital.” The programs that give them benefits are assumed to be economic entities, powered by budgets and generating income along with incentives, and so on.

Researchers who are economists make these assumptions naturally. Other researchers make them insensibly, simply because economizing assumptions have become so influential in academic culture. Yet on its face, economism in social policy is implausible. Economics is a science of

²³ I refer here to the main social surveys mentioned above. Some other surveys, such as political science’s National Election Studies, embody a more specific psychology.

market behavior, but, by virtue of being jobless and low-income, most poor adults live largely outside the market. To understand them requires a more complex psychology.²⁴

Survey-based research tends also to be conservative in the sense of incremental. Each run of the surveys uses essentially the same variables and categories as before. Research analyzing new data inevitably takes its departure from what has gone before. Hypotheses derive from “the literature,” meaning past research, a tendency also promoted by peer review. New inquiry either confirms or disconfirms those suppositions. But it is difficult to ask a new question.

Questioning economism

Field inquiry frees the researcher from these constraints. It brings scholars into direct contact with the people involved in a social problem. One no longer views them through the filter of data bases. One must explain behavior initially on the basis of observation, setting aside the academic edifice. The result is initially a more journalistic understanding, based on what one sees and hears from people, rather than on statistical inferences. As I note below, field work may also lead to more valid statistical inferences based on program data.

Interpretation of social problems and programs gained from the field tend to conflict with economism. The influences on behavior will seem various, some of them stemming from the economy, but others from private life. The motivations of people will also seem diverse, hardly confined to optimizing incomes. Economists will object that they assume optimizing behavior, not because it describes actual psychology, but because it predicts behavior.²⁵ But in social policy, at least, economism is often not even predictive. It misleads policymakers about what measures can best improve lives for the seriously poor.

Field inquiry tends to force researchers away from economic understandings of poverty. For example, the view that work is commonly unavailable to the poor is strongly held by many antipoverty researchers. Elijah Anderson is one who believes that the disappearance of jobs from the inner city is the underlying reason why black youth from these areas tend to get in trouble with the

²⁴ Lawrence M. Mead, *The New Politics of Poverty: The Nonworking Poor in America* (New York: Basic Books, 1992), chap. 7.

²⁵ Milton Friedman, “The Methodology of Positive Economics,” in Milton Friedman, *Essays in Positive Economics* (Chicago: University of Chicago Press, 1953).

authorities—they lack legitimate opportunities. But he observes youth directly in these neighborhoods, and his accounts of their behavior make little reference to economics. Rather, he attributes the patterns to a breakdown of social authority—an inability by the police and the community to enforce order.²⁶

Katherine Newman began her study of low-wage workers in Harlem out of a conviction that blacks shared the same work ethic as the rest of the population. If many poor blacks did not work regularly, the reason must be the deficiencies of the labor market, not, as some conservatives said, an indulgent welfare system. She did find that there were fourteen applicants for every job at fast-food restaurants in Harlem, and that employers often preferred hiring people from outside the area. But she also found that most of the workers she interviewed had long employment histories. Their difficulty was not so much to find work as to keep it and get ahead. And one of their chief obstacles was that other youth in the area often opposed their taking low-wage jobs. This, again, traced the problem more to social factors than to opportunities.²⁷

Kathryn Edin and Laura Lein began their study of poor single mothers believing that these mothers could not make enough to live on through either welfare or low-wage work. Based on their welfare benefits alone, the single mothers they interviewed would have had incomes of just 53 to 84 percent of the poverty line, depending on their state of residence.²⁸ This comports with the idea most poverty researchers have that welfare mothers live well below the poverty line. But based on interviewing, Edin and Lein found a different reality. The welfare mothers averaged incomes of \$10,596 a year, which was 97 percent of the federal poverty line in 1991 (interviewing was done in 1989-92). The working mothers averaged incomes of \$14,712, or 134 percent of the poverty line. The mothers also had multiple income sources—child support and charity of various kinds as well as

²⁶ Elijah Anderson, Streetwise: Race, Class, and Change in an Urban Community (Chicago: University of Chicago Press, 1990); idem, Code of the Street: Decency, Violence, and the Moral Life of the Inner City (New York: Norton, 1999).

²⁷ Katherine S. Newman, No Shame in My Game: The Working Poor in the Inner City (New York: Alfred A. Knopf and Russell Sage Foundation, 1999).

²⁸ U.S. Congress, House, Committee on Ways and Means, Overview of Entitlement Programs: 1992 Green Book: Background Material and Data on Programs Within the Jurisdiction of the Committee on Ways and Means (Washington, DC: U.S. Government Printing Office, May 15, 1992), pp. 636-7.

welfare and work. They thus had more options than they seemed to have at the outset. Especially, whether to work was a choice, not something they were driven to by necessity.²⁹

Whereas data bases convey social and economic conditions, field research highlights informal and institutional suasions that are not so easily measured. In all the studies mentioned, the world of the poor is structured by the rules governing the work place and child support as well as welfare. Those forces would be even more important today, because it was chiefly changes in work rules that drove the recent welfare revolution. Another change is tougher child support enforcement, due to major changes in that system.³⁰ This is just the institutional content that policy research needs in order to realize the vision of a master science.

Serendipity

Above all, field research allows for the asking of new questions. Because of the complexity of what one learns about social problems and programs, ideas about causes and cures may emerge that fall outside the existing literature. That allows the researcher to pose hypotheses that are genuinely new. These in turn can be the basis for more realistic quantitative inquiry.

For example, before the 1990s, the usual view of employment programs was that they are economic in nature, driven by resources and incentives. Their effects depended on how many clients they were funded to serve. Each had a limited number of “slots.” Recipients were chosen for these slots and then were “invested” in by the program. There was thought to be a tradeoff between serving many clients minimally and serving fewer more richly. This perspective treated the clients as passive receptacles of social resources.³¹ But in the late 1970s, studies were done of the Employment Service and the Work Incentive (WIN) program that related their performance, in part, to how well they were run. WIN was the first of several mandatory work programs instituted in

²⁹ Kathryn Edin and Laura Lein, Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work (New York: Russell Sage Foundation, 1997).

³⁰ Ronald B. Mincy and Hillard Pouncy, “Paternalism, Child Support Enforcement, and Fragile Families,” in The New Paternalism: Supervisory Approaches to Poverty, ed. Lawrence M. Mead (Washington, DC: Brookings, 1997), chap. 4.

³¹ Judith M. Gueron and Edward Pauly, with Cameran M. Lougy, From Welfare to Work (New York: Russell Sage Foundation, 1991); Jan L. Hagen and Irene Lurie, Implementing JOBS: Progress and Promise (Albany: State University of New York, Rockefeller Institute of Government, August 1994).

AFDC to put employable recipients to work. This research, done at the Urban Institute, measured the performance of state programs and then traced it, in part, to features of organization and management, the latter assessed through fieldwork.³²

During the WIN interviewing, local staff gave me quite a different image of their task. It was not to invest in passive clients. Rather, it was to get the clients to help themselves by going out and getting a job. Whether the program could do that depended only partly on its funding. The clients would go to work if they were expected to, but the program commonly lacked the authority to expect work. I followed up this idea in my own studies. They showed that whether a program obligated its clients to participate, and whether it put them in job search, strongly predicted the job placements it made, even controlling for the employability of the clients and the local labor market. This finding jibed with evaluations of work programs in the 1980s showing that programs that enforced participation had the largest impacts.³³

In the fieldwork for these studies, further possibilities came to me: The share of clients that participated actively in a program was a measure of its authority, and to maximize participation required a paternalistic administration where staff oversaw clients to be sure they fulfilled their obligations. By these means, effective work programs could drive employment up and caseloads down. That potential was dramatically realized in Wisconsin.³⁴ In all these studies, field impressions were verified by statistical analyses using program data. But the original ideas came from the field. They could not have come from prior research, for until the 1990s, virtually nothing

³² Mark Lincoln Chadwin, John J. Mitchell, Erwin C. Hargrove, and Lawrence M. Mead, The Employment Service: An Institutional Analysis, U.S. Department of Labor, R&D Monograph 51 (Washington, DC: U.S. Government Printing Office, 1977); and John J. Mitchell, Mark L. Chadwin, and Demetra S. Nightingale, Implementing Welfare-Employment Programs: An Institutional Analysis of the Work Incentive (WIN) Program, R&D Monograph 78, U.S. Department of Labor (Washington, DC: U.S. Government Printing Office, 1980). A summary of the WIN study appeared as Mark Lincoln Chadwin, John J. Mitchell, and Demetra Smith Nightingale, "Reforming Welfare: Lessons from the WIN Experience," Public Administration Review 41, no. 3 (May/June 1981): 372-80.

³³ Lawrence M. Mead, "Expectations and Welfare Work: WIN in New York City," Policy Studies Review 2, no. 4 (May 1983): 648-62; idem, "Expectations and Welfare Work: WIN in New York State," Polity 18, no. 2 (Winter 1985): 224-52; idem, "The Potential for Work Enforcement: A Study of WIN," Journal of Policy Analysis and Management 7, no. 2 (Winter 1988): 264-88.

³⁴ Lawrence M. Mead, "Optimizing JOBS: Evaluation Versus Administration," Public Administration Review 57, no. 2 (March/April 1997): 113-23; idem, "The Decline of Welfare in Wisconsin," Journal of Public Administration Research and Theory 9, no. 4 (October 1999): 597-622

in the poverty literature treated work as a problem of enforcement.³⁵ The view dominated that low work levels among the poor was an economic problem, rooted in the labor market.³⁶ It took direct contact with the problem, unmediated by data bases, to reveal the key role of the institutions.

Where Research Erred

In several instances related to welfare reform, impressions of social problems given by conventional social research were contradicted by field research as well as by policy experience.

Can welfare recipients work?

Was it possible for welfare mothers to work? Prior to reform, many experts doubted this. Most studied the employability of the poor statistically, by examining the features of needy adults who did and did not work in various data bases. They typically found that the nonworkers had more children, less education and work experience, and more personal and family problems than others who were employed. They treated these correlations as “barriers” to employment. They concluded that government should not demand greater effort of single mothers, unless it first alleviated the “barriers.” That meant providing recipients with better education and training and more child and health care. Even if they could work, some experts suggested, there was little point in their doing so because the jobs they could get would still leave them poor.³⁷

A typical conclusion was that “. . . the wages and job options available to less-skilled women are extremely limited; increased hours of work do not produce large economic returns. Female heads of households . . . face complex child care issues, and the welfare of their children could be

³⁵ One exception: Valerie Englander and Fred Englander, "Workfare in New Jersey: A Five Year Assessment," Policy Studies Review 5 (no. 1, August 1985): 33-41.

³⁶ Until at least the late 1980s, most experts rejected the idea that welfare work programs had to be mandatory to succeed. See Laurence E. Lynn, Jr., ed., "Symposium: The Craft of Public Management," Journal of Policy Analysis and Management 8, no. 2 (Spring 1989): 284-306. In the wake of welfare reform, few would dispute that idea today.

³⁷ E.g., Charles Michalopoulos and Irwin Garfinkel, "Reducing the Welfare Dependence and Poverty of Single Mothers By Means of Earnings and Child Support: Wishful Thinking and Realistic Possibility" (Madison: University of Wisconsin, Institute for Research on Poverty, August 1989); Maria Cancian and Dan R. Meyer, "A Profile of the AFDC Caseload in Wisconsin: Implications for a Work-Based Welfare Reform Strategy" (Madison: University of Wisconsin—Madison, Institute for Research on Poverty, September 1995); Demetra Smith Nightingale and Robert H. Haveman, eds., The Work Alternative: Welfare Reform and the Realities of the Job Market (Washington, DC: Urban Institute Press, 1995); and S. Jody Heymann and Alison Earle, "The Work-Family Balance: What Hurdles Are Parents Leaving Welfare Likely to Confront?" Journal of Policy Analysis and Management 17, no. 2 (Spring 1998): 313-21.

seriously harmed by further declines in AFDC benefits and any arbitrary termination of welfare support.” Indeed, a work demand “may be the proverbial ‘last straw’—one more pressure on an already strained situation in families already stretched to the limits of their coping ability.”³⁸

Some other research, however, suggested that work by welfare mothers was quite common, even if much of it went unreported to the welfare department to avoid grant reductions.³⁹ Edin and Lein’s research revealed that whether to work was a choice that low-income single mothers made depending on circumstances, not something they were barred from by rigid impediments. Since even the “welfare-reliant” mothers in this study sometimes worked, it was not obvious why more could not do so regularly. In the 1990s, policymakers decided that the majority could in fact work, and events confirmed their judgment. I discuss the rise in work among poor mothers further below.

Some recent academic studies appraise the employment prospects of welfare adults more positively. It now seems that positions are widely available to them. Wages coupled with remaining government benefits such as EITC can raise their incomes substantially. Provided they take jobs and work steadily, they can escape poverty and move up to higher incomes over time.⁴⁰ But these studies came in the wake of welfare reform when it was easier to see the advantages of employment.

³⁸ Rebecca M. Blank, “The Employment Strategy: Public Policies to Increase Work and Earnings,” in Confronting Poverty: Prescriptions for Change, ed. Sheldon H. Danziger, Gary D. Sandefur, and Daniel H. Weinberg (New York: Russell Sage Foundation, and Cambridge, MA: Harvard University Press, 1994), p. 203; Patricia Spakes, “Mandatory Work Registration for Welfare Parents: A Family Impact Analysis,” Journal of Marriage and the Family 44, no. 3 (August 1982): 688.

³⁹ Mildred Rein, Dilemmas of Welfare Policy: Why Work Strategies Haven't Worked (New York: Praeger, 1982), chap. 6; Kathleen Mullan Harris, “Work and Welfare Among Single Mothers in Poverty,” American Journal of Sociology 99, no. 2 (September 1993): 317-52.

⁴⁰ Gregory Acs, Norma Coe, Keith Watson, and Robert I. Lerman, Does Work Pay? An Analysis of the Work Incentives under TANF (Washington, DC: Urban Institute, July 1998); Harry J. Holzer, “Will Employers Hire Welfare Recipients? Recent Survey Evidence from Michigan,” Journal of Policy Analysis and Management 18, no. 3 (Summer 1999): 449-72; Harry J. Holzer and Michael A. Stoll, “What Happens When Welfare Recipients Are Hired?” (Washington, DC: Georgetown University, May 2000); Susanna Loeb and Mary Corcoran, “Welfare, Work Experience, and Economic Self-Sufficiency,” Journal of Policy Analysis and Management 20, no. 1 (Winter 2001): 1-20; LaDonna Pavetti and Gregory Acs, “Moving Up, Moving Out, or Going Nowhere? A Study of the Employment Patterns of Young Women and the Implications for Welfare Reform,” Journal of Policy Analysis and Management 20, no. 4 (Fall 2001): 721-36; Sheldon Danziger, Colleen M. Heflin, Mary E. Corcoran, Elizabeth Oltmans, and Hui-Chen Wang, “Does It Pay to Move from Welfare to Work?” Journal of Policy Analysis and Management 21, no. 4 (Fall 2002): 671-92.

Are jobs available?

As mentioned above, many experts doubt that jobs are available to the poor of the inner city, either women or men. Many support the “mismatch theory,” or the idea that, for economic reasons, jobs have tended to leave the inner city, where most urban poor reside, and shift to the suburbs, the South, or overseas. Or if jobs still exist in the center city, they demand more skills than the poor usually have. This idea is the centerpiece of a complex theory of urban poverty developed by William Julius Wilson and his followers.⁴¹ While the statistical evidence for it is generally weak,⁴² a few recent studies are in favor, and this has kept hope for the theory alive.⁴³ This is a case where field impressions—low work levels in poor areas—may seem to support the theory.

The bulk of observational evidence, however, cuts the other way. Wilson mounted a large study of poverty in Chicago in the 1980s, but it failed to find clear evidence that jobs were lacking in the inner city. Rather, interviews with inner-city fathers showed that work levels varied strongly by ethnicity. Mexicans had the lowest human capital, yet 93 percent of them were employed, compared to 82 percent of whites. Blacks had human capital nearly equal to whites’, yet only 66 percent of them were working.⁴⁴ In addition, the Wilson project commissioned an interview of employers in and around Chicago who frequently hired low-skilled workers. They were hostile to hiring blacks from the inner city, claiming to have found them unreliable.⁴⁵ Heavy immigration by the unskilled

⁴¹ William Julius Wilson, *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy* (Chicago: University of Chicago Press, 1987); idem, *When Work Disappears: The World of the New Urban Poor* (New York: Knopf, 1996).

⁴² Christopher Jencks and Susan E. Mayer, "Residential Segregation, Job Proximity, and Black Job Opportunities," in *Inner-City Poverty in the United States*, ed. Laurence E. Lynn, Jr., and Michael G.H. McGeary (Washington, DC: National Academy Press, 1990), chap. 5; Harry J. Holzer, "The Spatial Mismatch Hypothesis: What Has the Evidence Shown?" *Urban Studies* 28, no. 1 (February 1991): 105-22.

⁴³ Keith R. Ihlanfeldt and David L. Sjoquist, "Job Accessibility and Racial Differences in Youth Employment Rates," *American Economic Review* 80, no. 1 (March 1990): 267-76. John Bound and Harry J. Holzer, "Industrial Shifts, Skills Levels, and the Labor Market for White and Black Males," *Review of Economics and Statistics*, vol. 75, no. 3 (August 1993): 387-96.

⁴⁴ Robert Aponte, "Ethnicity and Male Employment in the Inner City: A Test of Two Theories" (paper prepared for the Chicago Urban Poverty and Family Life Conference, Social Science Research Council, University of Chicago, Chicago, IL, October 1991).

⁴⁵ Joleen Kirschenman and Kathryn M. Neckerman, "'We'd Love to Hire Them, But . . .': The Meaning of Race for Employers," in *The Urban Underclass*, ed. Christopher Jencks and Paul E. Peterson (Washington, DC: Brookings, 1991), pp. 203-32.

from Latin America and Asia is the best evidence against the mismatch theory. For the native-born poor, not jobs but basic work discipline appeared to be what is lacking.

This evidence applies to the 1980s. The economy was still better later. When radical welfare reform hit the cities in the mid 1990s, the nation enjoyed the best labor market in thirty years—not only massive job creation and low unemployment but a recovery of earnings growth even for the low-skilled, whose wages had fallen for two decades. Nevertheless, while work levels for poor black women rose sharply in the 1990s, employment for black men continued to fall.⁴⁶ That trend cannot be explained by the opportunity structure.

Why did caseloads fall?

After national welfare caseload peaked in 1994, they fell by 60 percent through 2002.⁴⁷ The collapse accelerated with the passage of PRWORA in 1996, but began well before it. It was easily the steepest decline in the history of AFDC/TANF.

The most dramatic policy change in the period was the institution of serious work requirements in AFDC/TANF. Yet several economic studies attribute the bulk of the decline to the excellent economic conditions of the 1990s, demoting welfare reform to a secondary role.⁴⁸ Some other studies give reform a greater role, but give it primacy only in the late 1990s.⁴⁹ Overall, economic causes dominate these analyses.

That is surprising in light of the failure of the caseload to decline by much in previous tight labor markets. Indeed, in the 1960s the rolls grew relentlessly despite forceful economic growth. In

⁴⁶ Harry J. Holzer and Paul Offner, “The Puzzle of Black Male Unemployment,” *The Public Interest*, no. 154 (Winter 2004): 74-5.

⁴⁷ Caseload data from the U.S. Administration of Children and Families.

⁴⁸ James P. Ziliak, David N. Figlio, Elizabeth E. Davis, and Laura S. Connolly, “Accounting for the Decline in AFDC Caseloads: Welfare Reform or Economic Growth” (Madison: University of Wisconsin, Institute for Research on Poverty, November 1997); Council of Economic Advisors, “Explaining the Decline in Welfare Receipt, 1993-1996” (Washington, DC: Executive Office of the President, May 9, 1997); idem, “The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads, An Update” (Washington, DC: Executive Office of the President, August 3, 1999); Stephen H. Bell, “Why Are Welfare Caseloads Falling?” (Washington, DC: Urban Institute, March 2001).

⁴⁹ John H. Bishop, “Is Welfare Reform Succeeding?” (Ithaca: Cornell University, School of Industrial and Labor Relations, August 14, 1998); David T. Ellwood, “The Impact of the Earned Income Tax Credit and Social Policy Reforms On Work, Marriage, and Living Arrangements” (Cambridge, MA: Harvard University, Kennedy School of Government, November 1999); June E. O’Neill and M. Anne Hill, “Gaining Ground? Measuring the Impact of Welfare Reform on Welfare and Work” (New York: Manhattan Institute, Center for Civic Innovation, July 2001).

the boom of the late 1980s, they changed little, then grew by a third during 1989-94, an expansion that began before the recession of the early 1990s. Observational accounts of the recent fall, by journalists and field researchers, trace it principally to welfare reform. This was what forced welfare mothers to take employment seriously as an alternative to welfare. Jobs also had to be available, and they were, but the new work policy was what triggered change.⁵⁰

In Wisconsin, the state with the sharpest caseload fall, one economic study attributed 53 percent of the decline over 1987-96 to a favorable economy, only 11 percent to welfare reform.⁵¹ However, when I asked state and local welfare officials in 1995 what was driving the caseload down, 52 percent cited the economy, but 58 percent cited increased work requirements, 37 percent credited changed political expectations surrounding welfare, and lesser percentages cited other policy changes such as increased support services and improved child support collections. Respondents said that mere controversy over welfare reform had begun to drive the rolls down, as recipients heard a message about self-reliance—even before reform was legislated, let alone implemented. Overall, policy and political causes clearly outweighed economic ones.⁵²

The difference probably is due to the fact that the economic studies cannot measure welfare reform as well as they do economic conditions. They gauge reform very imperfectly, in terms of whether a state had waivers from Washington to run experimental programs prior to PRWORA. As a result, some of the credit for change that should go to welfare reform is attributed to economic variables. If one measures policy more accurately using program data, the importance of reform in explaining caseload change appears greater.⁵³ Field research is vital to appreciate the global ways that reform influenced the caseload. It was more than a change in certain indicators that show up in

⁵⁰Jason DeParle, “Success and Frustration, as Welfare Rules Change,” *New York Times*, December 30, 1997, p. A1, A16; Richard P. Nathan and Thomas L. Gais, “Is Devolution Working? Federal and State Roles in Welfare,” *Brookings Review* 19, no. 3 (Summer 2001): 25-9.

⁵¹Ziliak et al., “Accounting for the Decline in AFDC Caseloads,” pp. 22-6.

⁵²Lawrence M. Mead, *Government Matters: Welfare Reform in Wisconsin* (Princeton: Princeton University Press, 2004), pp. 178-81. Respondents might cite any of these factors, so the percentages add to more than 100.

⁵³Lawrence M. Mead, “Caseload Change: An Exploratory Study,” *Journal of Policy Analysis and Management* 19, no. 3 (Summer 2000): 465-72; idem, “State Political Culture.” However, these are cross-sectional studies and thus not strictly comparable to the economic studies, most of which use pooled models.

regression equations. Rather, an entire political system mobilized to transform welfare, and succeeded

Why did work levels rise?

Linked to the caseload fall was a dramatic rise in work levels among the poor. In 1993, only 44 percent of poor mothers with children had worked at all, but by 1999 64 percent did; the level then declined to 57 percent by 2002, probably due to the recession.⁵⁴ Again, one might attribute this change to welfare reform, since raising work levels by poor parents was one of its leading goals. The principal economic study, however, attributes most of the increase to the EITC, which was expanded in 1990 and 1993, rather than to work enforcement.⁵⁵

Again, the finding conflicted with earlier policy experience. In the 1970s and 1980s, giving work incentives to welfare mothers (that is, allowing them to keep part of their earnings if they worked, rather than deducting all of them from grants) had shown little power to raise work levels.⁵⁶ This was one reason why policy shifted toward administrative work requirements— demands to work as a condition of aid. In the Minnesota Family Investment Program (MFIP), which used both incentives and requirements to promote work, the two together had much stronger impacts on employment and earnings by long-term single mothers than incentives did alone.⁵⁷ How then could the EITC have such large pro-work effects?

The result also is contradicted by field evidence. In the Wisconsin interviews mentioned above, not a single respondent—out of 60—mentioned the EITC as a cause of the caseload decline. This even though the state had added its own wage subsidy to the federal one.⁵⁸ If no one had heard

⁵⁴ March 1994 Current Population Survey, table 19; March 2000 Current Population Survey, table 17; March 2003 Annual Demographic Survey, Table POV15.

⁵⁵ Bruce D. Meyer and Dan T. Rosenbaum, “Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers” (Cambridge, MA: National Bureau of Economic Research, September 1999).

⁵⁶ Robert Moffitt, "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature* 30, no. 1 (March 1992): 1-61.

⁵⁷ Cynthia Miller, Virginia Knox, Lisa A. Gennetian, Martey Dodoo, Jo Anna Hunter, and Cindy Redcross, *Reforming Welfare and Rewarding Work: Final Report on the Minnesota Family Investment Program, Volume 1: Effects on Adults* (New York: Manpower Demonstration Research Corporation, September 2000), pp. 83-113. After the first year, the addition of the requirements to the incentives more than doubled the employment impact and accounted for all of the earnings impact. The incentives were more important for raising incomes and reducing poverty.

⁵⁸ One respondent mentioned increased welfare disregards for dependent care.

of these subsidies motivating respondents to leave the rolls, how can we believe the large effect claimed? Again, the likely cause is that quantitative studies cannot measure welfare reform accurately, so that some of its effects end up being attributed to the EITC or other economic factors.

Journalistic observation suggests that in Wisconsin welfare recipients went to work chiefly because they were pushed by policy changes, and also because jobs were available. They then heard about the EITC and received it as a windfall. Either they filed for it after they were working, or it was sent to them by tax authorities because they appeared eligible.⁵⁹ A recent analysis of work increases in California finds, similarly, that many single mothers who received the credit did not file for it themselves, suggesting that they may not have been aware of it.⁶⁰ It appears that higher work levels are driving up EITC claims, not the other way around. Again, field observation clarifies realities missed by the most elaborate statistical analysis.

Optimizing Field Research

The limitation of survey-based research is that it obtains precision and generalizability at the expense of over-simplifying reality.⁶¹ Only attributes that are easily measurable can be included in large data bases, and these may miss the institutional forces that increasingly structure the world of the poor. The whole governmental effort to promote welfare work and child support occurs largely outside these data.

Field research has opposite limitations. Hands-on contact does convey a fuller picture of reality, but drawing conclusions is more judgmental. Fieldwork can become a form of journalism, where impressions of a problem are knit together out of diverse comments from the people one happens to interview. The picture may not be wrong, but we cannot know that it is right without

⁵⁹ Jason DeParle, "On a Once Forlorn Avenue, Tax Preparers Now Flourish," New York Times, March 21, 1999, pp. 1, 20.

⁶⁰ V. Joseph Hotz, Charles H. Mullin, and John Karl Scholz, "The Earned Income Tax Credit and Labor Market Participation of Families on Welfare" (Los Angeles: University of California—Los Angeles, Department of Economics, January 9, 2003).

⁶¹ This discussion draws on Gary King, Robert D. Keohane, and Sidney Verba, Designing Social Inquiry: Scientific Inference in Qualitative Research (Princeton: Princeton University Press, 1994), pp. 151-5.

comparing it to other, more systematic research.⁶² The subjectivity comes partly from judgments about whom to talk to, partly from how one sums up what people say.

One solution is to make the respondents in field inquiry as representative as possible. Define the universe of officials or recipients one wishes to tap, and then choose a selection that captures the full range of variation. For staffs, it is seldom feasible literally to draw a sample; at each site, one has to interview the people available on the day; but one can at least choose people from a variety of positions inside the organization. In a program with multiple offices, interview at all sites if possible, otherwise at a selection that is representative. Such a sample usually has to be constructed rather than randomly chosen, since the number of sites visited is likely to be small.⁶³ The goal is to obtain the full range of variation in the program dimensions that one believes might influence performance.⁶⁴

Most important is to verify the picture of problems and solutions that comes from the field in some more rigorous way. The evil of research premised on “the literature” and established data bases is that it forecloses serendipity. But once the field has offered up new hypotheses, they should be verified statistically if possible. Do the connections that respondents assert between certain policies and results really hold; and if they hold at one site, are they valid across the entire program? This is possible if and only if the program dimensions of interest can be measured and thus related to performance measures, and if one can control for other influences.

To test hypotheses that come out of programs usually takes program data, as only this can capture the institutional or policy features highlighted in interviews. In my welfare work studies, respondents identified the program’s authority as a crucial determinant of performance. I measured it using the participation rate, while the proportions of clients assigned to varying activities captured the program’s emphasis on job search versus education and training. Staff suggested that high

⁶² Some field accounts of Wisconsin welfare reform were less systematic than my own but reach similar conclusions. See Jason DeParle’s series which ran in the New York Times between May 7, 1997, and December 30, 1999, and Amy L. Sherman, “The Lessons of W-2,” The Public Interest, no. 140 (Summer 2000): 36-48.

⁶³ In my welfare work studies I either interviewed all local programs (New York City), visited some and talked to others by phone (New York State), or visited a constructed sample of 10 counties meant to be representative of 72 counties total (Wisconsin).

⁶⁴ King et al., Designing Social Inquiry, chap. 4.

participation and an emphasis on job search tended to produce more job entries, and the statistics confirmed this. It was also vital to these analyses that the programs possessed, or produced for me, data on the demographics of their caseloads, to generate control variables. Labor market controls were available from U.S. Bureau of Labor Statistics data. The effect was to change equation (5) into:

$$(6) \quad P = \alpha + \beta_1\mathbf{E} + \beta_2\mathbf{S} + \beta_3\mathbf{B} + \beta_4\mathbf{A} + \varepsilon$$

Now alongside variables expressing economic and social conditions, or benefits, we have a vector expressing various administrative dimensions of the program. This model is much more able to capture all the influences on the clients and hence on outcomes.⁶⁵

Interview findings backed up by statistical analysis are a powerful combination for understanding how programs affect social problems. Just as the numbers serve as a check on the interviews, so the interviews were vital to explaining the numbers. Unlike survey-based research, this sort of modeling is tied into bureaucratic context. It dramatizes the role of the institutions in shaping outcomes, just as standard modeling tends to conceal it. The combination of the field and program data begins to demonstrate the sort of connection seen in equation (3). Governmental content is brought into policy research, and the master science of equation (1) becomes imaginable.

One might say that making connections between policy and outcomes is the job of evaluators. At their best, as in well-run experimental trials, evaluations do test the effects of some program on a clientele. The MDRC evaluations on welfare work programs contributed much of the institutional content that social policy research does have. But evaluations usually reveal little about how the institutions function. They can show the impact of a policy, but programs are complicated, and it seldom clear which of their elements produced the effect, or how.⁶⁶ By comparing the effects of different client assignments on performance measures, statistical models help to break into the

⁶⁵ Unfortunately, the reporting now available on welfare work programs today is less adequate than in the WIN era. See Mead, "Performance Analysis," pp. 123-6/

⁶⁶ An experimental evaluation can isolate the effect of specific program elements only if there are multiple program groups so as to estimate the effect of each aspect of the treatment. MDRC has accomplished this in some of its evaluations, notably in the comparisons of "labor market attachment" and "human capital development" approaches in the National Evaluation of Welfare to Work Strategies.

“black box.” This even though the results are not strictly speaking an evaluation; they compare subunits within a program, rather than the program against no program at all.

The “new public management” has recently proposed a more ambitious form of program modeling. This school, developed by Laurence E. Lynn, Jr., and his collaborators, aspires to link features of organization more rigorously to outcome measures than has traditionally been done in public administration. The group uses sophisticated methods, including hierarchical modeling, to relate both the services provided to clients and higher-level features of a program to its results.⁶⁷ In some cases, performance is stated in terms of impact, not just outcomes, because the data come from experimental evaluations.⁶⁸

In principle this is an advance over the simpler modeling suggested above, but all depends on whether researchers can measure the dimensions of interest. In many instances, modeling has been confined to peripheral aspects of organizational performance because these were the only ones where data was available. In some cases also, measures of service provision are calibrated using interviews rather than program data, there is no demographic data on caseloads, or there is no field dimension to the research at all.⁶⁹ These studies suffer from some of the same validity doubts as other academic research on social policy. There is no substitute for field inquiry backed up by program reporting that can measure the major determinants of interest. Without these resources, improved modeling techniques cannot achieve much.

Why Field Research is Rare

Field inquiry is vital to bringing governmental content back into policy research, and thus realizing the promise of the master science. But academic resistance to doing this is strong.

⁶⁷ Carolyn J. Heinrich and Laurence E. Lynn, Jr., eds., Governance and Performance: New Perspectives (Washington, DC: Georgetown University Press, 2000).

⁶⁸ Carolyn J. Heinrich, “Organizational Form and Performance: An Empirical Investigation of Nonprofit and For-Profit Job-Training Service Providers,” Journal of Policy Analysis and Management 19, no. 2 (Spring 2000): 233-61; Howard S. Bloom, Carolyn J. Hill, and James Riccio, “Linking Program Implementation and Effectiveness: Lessons from a Pooled Sample of Welfare-to-Work Experiments,” Journal of Policy Analysis and Management 22, no. 4 (Fall 2003): 551-75.

⁶⁹ Lynn, Laurence E., Jr., and Patricia W. Ingraham, eds., “Government and Public Management: A Symposium,” Journal of Policy Analysis and Management 23, no. 1 (Winter 2004): 3-96.

Notwithstanding the inability of survey-based studies to understand the recent revolution in welfare, I see no tendency for more researchers to get out into the field and find out what is going on.

When I began my research into welfare reform in Wisconsin, I expected that I would have many competitors. The state was too prominent for other scholars not to out there laying hands on the institutions. But in months of fieldwork in the Wisconsin counties, I never encountered another researcher doing similar work, nor heard of any. I was also virtually alone in studying the politics of reform in the state.⁷⁰ The people who came to Wisconsin to study the reform directly were all journalists or welfare officials from elsewhere in the country, or overseas.

Many scholars at the Institute for Research on Poverty (IRP) at the University of Wisconsin—Madison and elsewhere did want to study the reform. Mainly, they wanted to study its social and economic effects. Requests to the state for data or for permission to interview clients were so heavy the state had to create a special body to process them.⁷¹ But nearly all these scholars operated from behind their computers. I had the political and administration dimensions of reform—to me the most significant—all to myself.

One reason, again, was politics. IRP was largely excluded from the planning of welfare reform in Wisconsin because its researchers largely disagreed with the conservative tenor of that change. One reason I could do fieldwork in Wisconsin was that, as a recognized conservative, I could get the data and permission to interview. Even had other scholars sought entrée, they might have been denied it. Access is a precondition for field research, and it is far from automatic.⁷²

The greater fact, however, is that for most researchers improving public policy is not a high priority. Even many of those who do policy research in the universities or think tanks are not centrally interested in equation (2), let alone (3) or (1), in my experience. Other things are more important. One of these is disciplinary respectability. All the social sciences today have modeled

⁷⁰ Michael Heaney, a graduate student at the University of Chicago, also researched the politics. His methods were more systematic than mine, but he drew conclusions similar to mine, reported in Lawrence M. Mead, “The Politics of Welfare Reform in Wisconsin,” *Polity* 32, no. 4 (Summer 2000): 533-59.

⁷¹ Mead, *Government Matters*, pp. 153-5.

⁷² Mead, *Government Matters*, pp. 56, 60, 107-8, 256; Mead, “Performance Analysis,” pp. 121-3.

themselves on the hard sciences. They seek theoretical knowledge, as the various disciplines understand that. By these lights, most policy research appears “applied” and low-tech.

Even more important is rigor, or the derivation of results that are precise and unquestionable. There is a tradeoff between rigor and realism. To achieve bulletproof results usually requires simplifying information sources, so that one can give an exact account of how inferences were drawn. Realistic policy research, however, involves multiple information sources. Integrating them to draw conclusions is necessarily judgmental. That is particularly true with field research, which is more suggestive than conclusive. Such inquiry faces skepticism from peer reviewers. To them it can easily appear as little better than consulting or journalism. This makes the work less eligible for journal publication, which is the coin of the realm in today’s academe.⁷³

For these reasons, most researchers who claim to pursue policy subjects today are primarily mathematicians. Only such people can make it in academe or the leading think tanks. Their strength is in sophisticated forms of data analysis, and they display it doing survey-based research. However, the contribution this work makes is largely to method, not policy substance. This research community is alive with acute intelligence and intense methodological rivalries, but it is remote from reality, because it involves very thin information sources.

The issue between field research and academia goes to the nature of knowledge. Do you believe that there is a reality “out there” that can be simply observed, or do you believe that knowledge is created by method. Field research asserts the former. The method involves dangers of subjectivity that need to be countered in the ways I have suggested. But the reality sought is nevertheless seen as something waiting to be discovered. Academia, however, tends to believe that truth cannot be simply apprehended. It must be constructed by some deliberate methodology. It has no reality apart from statistical inference.

Field research places faith in stamina and shoe leather, not mathematical methods. One leaves one’s computer and goes out into the world with open hands, and without strong hypotheses. One is assaulted by information from many sources, in many forms. One struggles to assimilate and

⁷³ The research for Government Matters generated nine academic articles. But only one of them appeared in a major political science journal. In most cases, the obstacle was inability to meet canons of rigor set largely outside policy research. All the other papers appears in public policy outlets.

integrate it. It is like the moment in “The Wizard of Oz” when the tornado transports Dorothy into the magic kingdom, and the screen turns from black and white to color. Government as it really operates is alive with energies and complexities invisible to the data bases. It is out there in the real world that the battle for better public policy must be won. To support that quest, policy research must lay hands on the institutions. Or the master science cannot be realized.