Forthcoming in David L. Weimer and Aidan R. Vining, Editors, *Investing in the Disadvantaged: What We Know, and What We Need to Know, about the Benefits of Social Policies*. Washington: Georgetown University Press, 2008.

Chapter 11

Next Steps in Welfare to Work

Michael Wiseman

Welfare-to-work policies seek to build human capital by encouraging and facilitating greater or more beneficial participation in labor markets. Effective policies not only increase income but also generally raise the return to additional human capital investment. What are possibly effective policies? How can we know if they would be effective? How do we know if they are desirable?

In this chapter I answer the first two questions by proposing several policy demonstrations. Each of the demonstrations is motivated to some extent by existing research. Its execution would generate information that would enable researchers to determine its effectiveness. I answer the third question by reviewing the application of cost-benefit analysis (CBA) to the Minnesota Family Investment Program, one of the most important state initiatives in the welfare policy area in terms of breadth of assessment and contribution to policy development.

I propose three demonstrations: first, an experiment with subsidized third-party provision of strategic financial advice and support for working low-income families; second, coordinated state experimentation with a transitional incentive package for recipients of Temporary Assistance for Needy Families (TANF); and third, an employment intervention for TANF applicants and recipients with substantial disabilities. There is a common theme to these proposals. Each involves changes in factors that influence both what is gained from human capital currently possessed and from additional investment in skills and reputation for targeted

families. For working families, the third-party advisor demonstration targets access to benefits, information, and sense of personal control. The transitional incentive package demonstration intends to rekindle state interest in active efforts to improve TANF policies that influence incentives to work. The employment intervention creates an incentive for refocusing the attention of TANF applicants and recipients with substantial disabilities on habilitation and skill development rather than benefits acquisition. All three involve uncertain benefits or costs about which a well-designed experiment should produce considerable information.

Financial Intermediation and Strategic Support (FISS) Demonstration

My first proposal is to subsidize a financial planning and services firm to act as an agent for working low-income families. My proposal includes using such firms to administer certain publicly financed work and income supports, most notably food stamps. The motivating hypothesis of this demonstration is that a third party, behaving as an agent for working families, would help increase their savings and improve their labor market outcomes. These good things would come about because such firms would provide better integration of support services, better explication of strategy, and more consistent and income-focused family support through time. The same firms, acting as agents of government, would be able to improve accuracy in both eligibility determination and benefits provision. The value added in this function plus benefits to target families potentially justify the subsidy. This proposal grows out of discussions of the behavioral effects on households of the Earned Income Tax Credit (EITC) and my observation of certain developments in the private tax-preparation market the EITC has generated, specifically the activities of H&R Block, Inc., with which I consulted briefly in 2005. For shorthand, I refer to this Financial Intermediation and Strategic Support demonstration as the FISS proposal.

There is widespread agreement that the Earned Income Tax Credit played a central role in the increase in employment rates among single parents in the 1990s (Eissa and Hoynes 2006). The contribution of the EITC to the dramatic declines in TANF receipt that occurred over the same period is more controversial. Scholars such as Jeffrey Grogger (2004) attribute much of the caseload decline to the combination of EITC benefits and a very strong economy. Others, most notably Lawrence Mead (2006), argue that while the EITC played an important role in "making work pay," the principal factor behind decline in welfare receipt was terminations generated by aggressively applied work requirements. This controversy may in part be resolved by appreciating that caseloads can decline both because exit rates rise and because accession rates fall. It is possible that EITC has had little direct effect on job-taking by TANF recipients (as Mead claims), but has substantial consequences for the likelihood that families at risk of welfare dependence are in fact forced to apply for TANF. Studies of working poor life and interaction with welfare—notably Jason DeParle's (2004) *American Dream*—often provide anecdotal information on the importance of EITC-based tax refunds in preserving the financial viability of self-support for families at risk of TANF accession.

The average beneficiary of an EITC refund receives a substantial amount of money—$1,815 per tax return in 2004. This creates what some analysts call a "moment of opportunity" for savings and has spurred an effort to offer recipients the opportunity to divert part of the refund to some sort of saving instrument. Several changes in tax policy have been introduced to encourage saving part or all of tax refunds, including an option allowing tax payers to designate that part of their refund be deposited in up to three separate accounts (including IRAs). Those contributing to an IRA or other employer-sponsored savings plan can also benefit from the Saver's Credit. One suspects that these initiatives would be more effective if delivered in

context of more general and trusted advice on financial strategy, but giving such advice is not an Internal Revenue Service responsibility.

A substantial portion of EITC returns are received by filers who use tax preparation services, most notably H&R Block. The EITC is arguably the single most important factor contributing to recent growth of Block and its aggressive expansion in recent years into areas with substantial numbers of low-wage workers. Block and its competitors have suffered considerable bad publicity for provision of short-term “refund anticipation loans” (RALs) that provide its clients with an instant EITC for a charge that is probably in line with costs but, when converted into an annual interest equivalent appears usurious (Berube et al. 2002). In addition, the firm has been sued by Elliott Spitzer, in his former capacity as the New York State Attorney General, for promoting investment of EITC refunds in Individual Retirement Accounts (IRAs) managed by the firm. Like many others, in the period 2001-2003 these investments yielded very low and sometimes negative returns, and Spitzer has attempted to hold the firm liable. Block has responded aggressively, both directly and indirectly. The direct response has been to argue that over the longer run the firm’s mutual funds are good investments and that such accounts offer low-income workers a useful opportunity. While asserting that the RALs provide a useful service, sought by consumers, Block has reduced its marketing of RALs and is instead promoting savings.

The Block indirect response has been both political and strategic. The firm has formed alliances with significant left-of-center interest groups in Washington, including the Center for Budget and Policy Priorities and the New America Foundation. At the same time, it has begun actively participating in experimental work on savings incentives, presumably in part because a national policy of subsidizing saving by its clients would redound to the firm’s benefit. A signal

example of this effort was the funding and conduct in 2004 of an evaluation of the effect of financial incentives on the willingness of its clients in the St. Louis area to invest EITC refunds. The evaluation was done by random assignment, with oversight and results analysis by a group of researchers affiliated with the Brookings Institution (Duflo et al. 2005). The mixed results of this investigation prompted an expanded research effort with a greater variety of treatments and at more sites in 2005.

Congressional testimony by its officers and other firm activities seem to indicate that Block sees the public support system for working low income families to be an important source of opportunities for the future. The firm has begun to offer state and local governments support in program outreach for the State Children's Health Insurance Program (SCHIP) and for the Food Stamp Program. One of Block's subcontractors, Onesta, Inc. (now cleverly renamed "Nets to Ladders"), has developed an artful software system for support of integrated social assistance intake that has attracted widespread interest in both government and the philanthropic community.

All this serves to reassure us that H&R Block has turned from the Dark Side, even if we trust that the apparent reorientation is in part propelled by profit. Moving beyond cynicism, it is important to consider the general issue. What *in abstract* does H&R Block appear to be doing? The firm is developing a strategic financial management service for families that potentially provides something quite similar to the advice financial counselors provide to middle-income families. The difference, however, is that financial assistance services for middle-income families are typically almost entirely focused on portfolio management and tax avoidance, while the emphasis of the Block effort is on gaining benefits as an initial step in developing savings—getting that portfolio in the first place. From the perspective of the clients, Block potentially

offers both agency and integration. Unlike social assistance agencies, Block is seen by clients as their employee: Block *wants* the client's business, and Block is willing to open thousands of store fronts and to stay open at odd hours to get it. Block is adopting a more comprehensive approach to assisting its clients in marshaling all available resources for increasing income and, in consequence, increasing both work and tax refunds as well as savings deposited in Block-related financial instruments. What the firm's critics seem to miss is that Block's clients appreciate the services they receive, and they are loyal. This is the foundation for productive agency.

Block has made overtures to a number of states to become involved in Food Stamp program (FSP) outreach. At the simplest level, the firm would like to do what many nonprofit organizations do now: Assist clients in filling out FSP applications that are then forwarded to the local FSP office. But more substantial connections are possible. One model is for Block to establish direct data exchange relationships with local FSP offices and to transmit the required information directly to the FSP agency's system. Food Stamp regulations require a direct interview with a government agency employee; this could be accomplished from the Block office. In addition to the benefits of such a system for the firm's clients, conceivably states could pay Block for administering the program for its clients and guaranteeing payments accuracy.

In recent years and especially since passage of the Personal Responsibility and Work Opportunity Reconciliation Act, much attention has been paid to finding ways to support working low-income households. For a variety of reasons, benefits and support programs have multiplied, creating a difficult maze of rules and regulations for households that by definition are short on time. Time lost in consequence of financial mismanagement is time not available for human capital development, and obtaining and enjoying the gains from human capital

development may require financial expertise that households lack. I suggest the agenda for new social experimentation include an experiment in which some population of working low-income households is offered an opportunity for comprehensive assistance with financial management covering not only tax filing and saving but access to some subset of basic social services, including food stamps, as well. I suggest the service be offered at minimal cost subject only to compliance with reporting requirements, but I think that to retain the sense of agency it is important that the service not be free. The service provider would be challenged to both develop individualized strategy for clients and to find ways to maximize participation.

Development of the FISS experiment would be complex. It would require that public agencies agree to employment of the financial service provider in administration of the public benefits included in the basic services package. (To my knowledge all of the statutory/regulatory restrictions can be waived in experimental context.) If Block were the agent, then the political issue raised both by privatization in this way but also the residual RAL stigma must be addressed. One way to do this would be to emphasize that should the project prove successful, licensing of organizations for provision of these services could be made competitive and open to a variety of nonprofit and for-profit organizations, as well as partnerships cutting across groups.

To repeat, the motivating hypothesis of the demonstration is that a third party, behaving as an agent of working families, would by better integrating support services, explication of strategy, and consistent support through time, increase saving and improve labor market outcomes. The consequence for public costs is uncertain, because part of the strategy is to increase, rather than reduce, access to services and benefits supported from public funds. This uncertainty, combined with a plausible model that identifies a role for intermediation, the corporate enthusiasm exhibited by H&R Block, and the experimental evidence to date justifies

planning a more widespread demonstration supported by an evaluation partner experienced in labor force and assistance issues, the operation of relevant government agencies, and philanthropic underwriting.

Transitional Incentives Package (TIP) Demonstration

The second experiment I propose has objectives on two levels. The immediate focus is a structural choice states must make in designing their TANF programs. However, I want to embed this experiment within an effort to improve methods for collective learning by states and the federal government about TANF policy choices. This is the “Transitional Incentives Package” (TIP) demonstration.

As with the other proposals, some background is useful. The first major social policy experiment, the New Jersey Negative Income Tax experiment, focused on studying the effect on labor supply of variations in basic assistance benefits and the rate at which benefits were “taxed” as income increased. This “tax rate” is commonly called the benefit reduction rate. Forty years have now passed since the New Jersey experiment was fielded, and the financial work incentives created by means-tested benefits programs continue to be a matter of major concern. One design objective for benefit systems is commonly cast as “making work pay,” taken to mean ensuring a benefit reduction rate less than one. As is well understood, providing an adequate minimum income guarantee while preserving financial incentives to work, and restraining costs, is difficult given the distribution of earnings and income in the American economy. The problem is compounded when public assistance is delivered through multiple programs, each with its own benefit reduction rate.

TANF provides states great leeway in designing the interaction between a recipient’s earnings and benefits. This latitude results in great diversity and great confusion. In a recent

paper, Linda Giannarelli and I counted 43 different procedures for varying benefits with TANF recipient earnings across the 50 states and the District of Columbia (Giannarelli and Wiseman 2006). Moreover, despite frequent references to the states as "laboratories of democracy," in the ten years since TANF was established, there has been no appreciable convergence of work incentive policy. Arguably, such convergence has failed to occur because no institutions exist to organize learning from this diversity or to share lessons gained across states.

The design of TANF benefits policies for working families differs from the classical negative income tax because most TANF programs are time-limited. Moreover, negative income tax systems are basically static; they take income as given and augment it as appropriate given the benefits schedule. In contrast, TANF programs are typically intended as transitional—as *temporary* assistance for those needy families. This dynamic character is underscored in 12 states in which TANF policy is to "disregard" a substantial amount of earnings in benefit calculation for the first few months following employment initiation and then to reduce TANF benefits, presumably as the family stabilizes in its new situation and in coordination with food stamp adjustments. Some states disregard essentially all earnings for a time. Other states treat earnings in standard negative tax fashion. Maryland, for example, has a flat benefit reduction rate of 60 percent; California uses 50 percent.

If TANF were the only income support system available, then a system that limited the duration of benefits for working families would hardly be appropriate. But in the United States every dollar reduction in TANF benefits is offset by an increase of $0.30 in food stamps, and working households eventually receive the EITC. Thus, one choice states face might be cast as between a regime in which a constant benefit reduction rate is applied to earnings for as long as permitted by the state's time limit and a regime in which the benefit reduction rate starts very

low—perhaps at zero—and then rises with time on the job. Obviously some condition would be required to prevent people from cycling on and off TANF to obtain the zero disregard repeatedly. With such a restriction, it is not clear at all which of the two regimes would have the more favorable effect on employment and welfare utilization at, say, a year past employment initiation. Available experimental evidence (see Michalopoulos 2005) provides little information pertinent to such a prediction or for choosing the best length for the high disregard interval.

A demonstration of this sort is hardly glamorous. However, what is important about it is that it addresses a policy choice that lies wholly in the sphere of state discretion, and one could imagine hammering out details as a collaborative exercise among states. As it is difficult to see in advance which system provides the greater advantage for unemployed TANF recipients, setting up a randomized experiment would probably meet less line worker resistance than many other possible innovations. Thus regardless of glamour, the choice is real and possibly significant in view of TANF objectives. The fiscal risks are minimal, as are the potential costs for recipients.

This brings us back to the larger objective. This overview provides few details; actual implementation would require much more design work. The issues that have to be addressed—how much is disregarded in the first months of job taking, how long “first months” should be, under what circumstances families should get a second chance, how to conduct the randomized trials, the penalties for states that fail to deliver—are generally easy to understand and usefully discussed directly with and among policy makers. But there are no procedures for bringing states together with federal officials to discuss options, secure commitment, and plan distribution of evaluation resources. I see picking up on an issue like this and using it as a means for developing better institutions for making the 51 state (including the District of Columbia) social

policy laboratories work as a common discipline. The experience of collaborative execution of a simple experiment with transitional work incentives might lay the foundation for other, more ambitious interstate collaboration. The architects of such a project would benefit from careful study of the evolution of the European Union's "Open Method of Coordination" in social policy (Walker and Wiseman 2006).

Employment Intervention for the Dependent with Disabilities (EIDD)

All available evidence indicates that disabilities are common among both adults and children in families applying for, and receiving, TANF benefits (Nadel et al. 2003/2004). Those with substantial disabilities are candidates for Supplemental Security Income (SSI). It is possible that the SSI eligibility determination process retards rehabilitation or employability development and movement to self-support for some persons. The third demonstration proposal involves early work-oriented intervention for TANF applicants and recipients who are potential candidates for SSI. SSI is the responsibility of the Social Security Administration, but TANF is the responsibility of the Department of Health and Human Services and state social assistance agencies. As a result, any such intervention in the TANF/SSI interface is complicated by problems of interagency collaboration. External financial support might encourage such collaboration, improve design, and ensure rigorous evaluation of effects. As with the other proposals, I begin with background and then outline the demonstration, which I call the Employment Intervention for the Dependent with Disabilities or EIDD demonstration.

Disabilities vary both in character and severity. A disability severe enough to preclude "substantial gainful activity" (SGA) combined with little income and few assets in principle qualifies working-age adults for Supplemental Security Income (SSI). Moving people from TANF to SSI is attractive both to recipients and to states. For recipients, the result is typically a

substantial income gain. In 2005 moving the adult in a three-person TANF recipient family would have on average more than doubled family income and removed the time limit on TANF assistance for children in the family. For states, such a transfer at least initially offers a financial windfall. While states pay all TANF benefits on the margin, the federal government pays all SSI. Twenty-three states supplement the federal SSI benefit out of state funds. For these states the transfer from TANF to SSI offers less net financial reward but is generally still desirable.

The interaction between TANF and SSI is substantial. By FY2003 11 percent of all SSI awards to adults age 18-64 and 35 percent of awards to children were to TANF/SSP recipients. (SSP refers to Separate State programs.) About 16 percent of families receiving TANF/SSP in 2003 included an adult or child SSI recipient (Wamhoff and Wiseman 2005/2006). State efforts to qualify persons with disabilities for SSI assistance are increasing as a result of changes in TANF brought about by the Deficit Reduction Act of 2005 (DEFRA). Prior to DEFRA, states could circumvent TANF work activity requirements for adults with disabilities and\or who were caring for children with disabilities by placing them in a "Separate State Program." DEFRA altered the treatment of SSPs in assessing states' TANF-related fiscal effort so that the SSP option has become much more costly. This has increased the incentive to move recipients from TANF to SSI for all states.

Moving eligible persons from TANF to SSI is a good thing, because by definition Congress intended SSI for those entitled and their income increases, at least in the short run, as a result. However, over the long run the process appears to have perverse consequences for the employability of some adults. Establishing absence of capability for substantial gainful activity is a time consuming process, often requiring as much as two years. During the application period SSI applicants are disinclined even to explore opportunities for employment or skill

acquisition because it is commonly believed that such efforts may be taken as evidence of capacity for SGA. Oddly, once an applicant is "awarded" SSI, the Social Security Administration offers him or her a "Ticket to Work"—access to a voucher system for obtaining work-related rehabilitation and training. Those who fail to gain SSI awards have used up TANF eligibility time and possibly are even less employable after having spent time outside the workforce striving to demonstrate absence of employment capacity. That such a system may be counter productive for rehabilitation and movement to self-support is widely recognized (Government Accountability Office 2004).

The SSI application process is similar to that used for making awards in the much larger Social Security Disability Insurance (SSDI) program. SSDI differs from SSI in that it is not means tested. All that counts is entitlement (established by work history) and, again, loss of capacity for SGA. Far more policy attention has been directed at SSDI than at SSI because the numbers are greater, the caseload trend is more dramatic, and SSDI claims are charged against the beleaguered Social Security Trust fund, while SSI payments come from general revenues (Autor and Duggan 2006). Like SSI, SSDI is seen as being too focused on establishing eligibility and too little concerned with restoring employability. In its September 2006 report the Social Security Advisory Board states:

> We [the Social Security Advisory Board] are convinced that the primary issue with our national disability system is not to be found in how the current Social Security disability programs are designed and operated but rather in the absence of an adequate culture and methodology to guide and support impaired individuals into a more work-oriented path that could avoid or delay their need to

depend on those Social Security Programs (Social Security Advisory Board 2006, 22).

The Social Security Administration did considerable development work in 2006 on an Early Intervention Initiative (EII) for SSDI applicants. In broad outline, the EII process would have been as follows. Applicants for SSDI in demonstration locations would be pre-screened for potential eligibility using an assessment tool developed by IBM. A triage would take place that was intended to separate applicants into three groups: (1) Those unlikely to receive SSDI awards; (2) those likely to receive awards but were judged capable of rehabilitation and job entry within some horizon—probably two-three years; and (3) those not considered candidates for near-term rehabilitation. A randomly selected subset of group (2) were to be offered a cash stipend, immediate access to Medicare, and opportunity to participate in a rehabilitation program intended to achieve employment if they suspended their SSDI applications. This intervention would apparently have offered the “work-oriented path” to which the Social Security Advisory Board refers. It would have been attractive to SSDI applicants both because of the Medicare component and the immediate stipend. Otherwise SSDI awardees must wait two years to become Medicare eligible, and be subject to the uncertainty of the disability adjudication process. One of the planned objectives of the EII evaluation was to do a benefit-cost analysis of the results. Clearly there would be more up-front costs to this system than are incurred under current procedures, but if the “work-oriented path” reduced the likelihood of long-term dependence on SSDI, the savings might justify the investment on even narrow “benefits to government” grounds. Whatever the project’s merits, EII project was terminated without field trial in early 2007, apparently because of projected cost.

Despite this setback, I think it important to consider something akin to the EII for adult SSI applicants, and it may be useful to focus on those connected to TANF. However, there are significant administrative barriers. There is little communication between local TANF agencies and local offices of the Social Security Administration on matters other than the incidence of SSI awards. There are also incentive problems. States gain from moving people from TANF to SSI. TANF recipients are categorically eligible for Medicaid, so the health insurance incentive is missing for these individuals. Assessment triage would likely be much more difficult, since features of work history are important factors in assessing probability of rehabilitation, and TANF recipients typically have much weaker labor force connection than do applicants for SSDI. On the other hand, for the states that supplement the federal SSI payments SSI take-up is not costless, and in any event the children in such families typically stay on TANF, without time limit. Involvement of SSI applicants who are TANF recipients in work training and/or habilitation programs will serve to help states reach TANF participation goals.

I propose an early intervention initiative for adult TANF recipients who are potential or actual applicants for Supplemental Security Income. In this Early Intervention for the Dependent with Disabilities (EIDD) demonstration, TANF applicants and recipients would be reviewed for disabilities and some subset will be identified as candidates for SSI. This "at risk" group would be triaged into the three groups identified for the SSDI Early Intervention Initiative. Again, the target is the group that appears likely to be successful in applying for SSI but also capable of employment and/or rehabilitation and, eventually, some contribution to self-support. In return for deferring application for SSI, the selected group would be offered a stipend that approximates the SSI benefit for a horizon to be determined coupled with ongoing TANF eligibility for their children.

There are basically four hypotheses to be investigated in the context of design and implementation of the EIDD. One is that persons in group 2, i.e. promising candidates for such an intervention, can actually be reliably identified. The second is that there are enough of them within any administrative catchment to justify focus of a serious innovation effort. The third is that such persons will respond to such an offer. And last but hardly least comes the motivating hypothesis that such an innovation would lead over time to increased employment and reduced assistance dependence at cost commensurate with the benefits.

The Minnesota Family Investment Program

Over the thirty year period leading up to the reforms instituted by the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), the evolution of welfare and work policy in the United States was closely intertwined with social experimentation. The experimental side of this policy double helix began in 1967 with planning for the New Jersey Negative Income Tax experiment and continued, in the 1970s, with the Seattle-Denver Income Maintenance Experiment and other related initiatives (Munnell 1986). In response to the welfare reforms introduced by the Reagan administration and the entrepreneurship of the Manpower Demonstration Research Corporation (now MDRC) in the 1980s initiative for welfare experimentation partially devolved to states. The result was a series of work-related state initiatives that produced evidence on the effects of work requirements that observers credited with influencing the content of the major welfare reform initiative of the 1980s, the Family Support Act of 1988 (Wiseman 1991). The Family Support Act in turn produced a set of initiatives collectively termed the National Evaluation of Welfare-to-Work Strategies, experiments that focused on strategic choices for the Job Opportunities and Basic Skills program established by the Act (Hamilton 2002). At the same time state experiments multiplied, creating

a national perception of innovation and competence at the state level that contributed to passage of PRWORA, the law that famously replaced Aid to Families with Dependent Children (AFDC) with Temporary Assistance for Needy Families (TANF). Since 1996 no significant welfare demonstrations have been initiated by the federal government or states, although some experimentation continues with welfare-related policies such as marriage promotion and efforts to support job retention and advancement among low-skilled workers. The reasons for this retreat are complex, reflecting both the ideology of PRWORA and the second Bush administration and the change in financial incentives for research and evaluation brought about by PRWORA and subsequent budget developments.

A central feature of this age of experimentation was the widespread utilization of random assignment and growing acceptance, at least within the policy analysis community, of classical experimental design as the "gold standard" for assessment of innovation effects. Judged in terms of refinement of methodology, breadth of assessment, and contribution to policy development, the Minnesota Family Investment Program (MFIP) is among the most important of the last generation of the state initiatives. MFIP outcomes are both fascinating and troublesome; the MFIP evaluation illustrates problems with achieving both internal validity and external utility of welfare and work research. To get to the problems, it is necessary to briefly explain the experiment and the cost-benefit analysis applied. It is impossible in a short overview to do justice either to MFIP or the literature to which the experiment's results have contributed; interested readers are encouraged to read more generally, beginning with the source for the present summary, the two volumes of the MFIP *Final Report*. Volume 1 covers effects on adults (Miller et al 2000), Volume 2 effects on children (Gennetian and Miller 2000). This

commentary focuses exclusively on effects on adults. Follow-up studies continue to be released by the evaluation contractor and posted on the firm's website (www.MDRC.org).

The Experiment

MFIP began in 1994. The program was an experimental change from public assistance as delivered in Minnesota's AFDC program.

As in most other states, Minnesota's AFDC program differentiated between policy for single- and two-parent families. To obtain assistance, single parents and two-parent families with a disabled adult were required to meet a standard of need based on income and assets. "Able-bodied" two-parent families were required in addition to satisfy the Unemployed Parent rule that the family's "primary earner" was involuntarily working fewer than 100 hours per month and in general were expected to engage more intensively in employment-related activities. Single parents were required to attend an orientation for the JOBS program (called STRIDE in Minnesota). The primary earner in a two-parent family was required to participate in job search and, in some instances, a Community Work Experience Program (CWEP) job in return for benefits.

Minnesota was exceptional in that the Food Stamp benefit was "cashed out" and included in recipient families' cash payment. However, for families with non-assistance income the benefit reduction was calculated separately for the AFDC and Food Stamp components of the grant. The larger AFDC grant calculation followed national AFDC regulations in excluding $120 and one-third of any monthly earnings in excess of $90 during the first 4 months of work, $120 during the next 8 months, and $90 per month thereafter. Thus AFDC incorporated a "transitional inventive" of the type targeted for experimentation earlier in this discussion. For single parent families, eligibility for AFDC was lost once earnings reached the point that the

AFDC component of the benefit fell to zero; the cutoff point for Food Stamp receipt was somewhat higher. In addition two-parent families were subject to the 100 hours restriction: No matter how large the family or how low the primary earner's earnings, once he (or, in rare instances, she) began working more than 100 hours a month, in principle eligibility for AFDC was lost.

MFIP was a state initiative that addressed four common concerns about AFDC: (1) long-term welfare dependency, (2) lack of work incentives, (3) the effects of the two-parent 100-hour rule, (4) and administrative complexity. The initiative altered treatment of families in several important ways. Food Stamp and AFDC benefits were integrated to produce a single eligibility standard with for some families an increased maximum benefit. The AFDC variable financial work incentive was replaced with a single more generous schedule that did not vary with time on the job. Work history requirements and the 100 hour test were eliminated as eligibility tests for two-parent households. Participation in employment and training programs was made mandatory for certain recipients: These included single parents with no child under 1 who had received AFDC payments for more than two years and one parent in two-parent families that had received assistance for more than 6 months.

The Evaluation

Alteration of AFDC and Food Stamp rules in this way required waiver of federal regulations. To obtain a waiver, it was necessary for the state to assume fiscal responsibility for costs above what would have been incurred in the absence of the intervention—from the federal government perspective the intervention needed to be "cost neutral." Assessing fiscal effects is difficult, but by the time of MFIP initiation, it was common to construct estimates of cost impact by comparing costs for cases subject to the innovation to costs for a control group subject to AFDC

rules. In general the administering agency, the Office of Policy Research and Evaluation in the Department of Health and Human Services Administration for Children and Families, required that controls be established by random assignment and that project evaluation include assessment of differences between costs for cases subject to MFIP rules and cases subject to AFDC. Minnesota agreed to this condition and contracted with the Manpower Demonstration Research Corporation, a nonprofit organization with substantial support from the Ford Foundation and other philanthropic organizations, to conduct the evaluation. In this and many other projects MDRC's exceptional third-party resources were used to expand substantially the purview of evaluation activity.

The field evaluation of MFIP was conducted in four rural and three urban Minnesota counties. Random assignment began in April 1994 and continued for 24 months. At the end of March 1996, 14,639 families had been added to the research (control and treatment) sample. Random assignment was applied both to new applicants for assistance and to on-going assistance at a point of periodic case review. Outcomes for both MFIP and the control groups were then followed, using a variety of instruments, for up to four years. Differences between the treatment (MFIP) and control groups are the foundation for a cost-benefit analysis. As is often the case in the conduct of social policy experiments, various external developments complicated execution of the research plan. In the Minnesota experiment such changes included first general increases in AFDC work requirements for single parents and, in 1997, replacement of AFDC with a statewide TANF program, called MFIP-S (for statewide), that included time limits, more aggressive work requirements, and a reduction in financial work incentives. The 100-hour rule was eliminated for on-going eligibility for all two-parent families, including those involved in the MFIP control group, but otherwise most rules were retained for control and experimental

families until mid-1998. The MFIP experiment was conducted in the context of a very tight job market, an environment likely to affect the productivity of the initiative's components.

Outcomes

The MFIP demonstration produced a wide range of outcomes. The final report differentiates between outcomes for single and two-parent families and, within each of these groups, between applicants and recipients. Applicants and recipients are defined differently for the two family types. Effects were measured over the first 28 months following random assignment. (Other parts of the evaluation incorporate longer windows—Lisa Gennetian, Cynthia Miller, and Jared Smith (2005) report six-year impacts.)

For single parents the MDRC analysis defines a "recent applicant" group that includes, in addition to actual applicants, persons who at the point of random assignment had received benefits for less than two years. "Long-term recipients" were those who had received benefits for 24 months at the point of random assignment. Families in MFIP stayed on welfare longer than their counterparts in the control group but tended to be more likely to combine work with welfare. In general most positive effects were found for long-term recipients. Some important benefits were not financial. For example, after three years women who were at random assignment to MFIP long-term recipients were more likely to be married than were their counterparts in the AFDC group. These outcomes came at a price: MFIP cost more per case per year than did comparable cases in the control.

For two-parent families MDRC differentiates between persons randomly assigned at application ("applicants") and persons randomly assigned while receiving assistance ("recipients"). By design applicants received the experiment's financial incentives immediately, but work requirements were deferred for the first six months. For both applicants and recipients

family labor supply fell in the experimental group relative to the control, generally as a result of lower employment rates and hours of work among women. Both applicants and recipients in the experimental group were more likely to receive some welfare benefits subsequent to assignment than were controls. For recipients (but not applicants) the greater receipt of benefits offset the loss of earnings so that experimental households had greater income. Survey evidence combined with study of administrative data indicates that couples in the two-parent recipient experimental group were more likely to stay married over the three years following assignment than were the controls; the survey sample size for applicants was insufficient to support analysis of marital stability for that group.

The Cost-Benefit Analysis

Like virtually every other MDRC program evaluation, the MFIP study includes a benefit-cost analysis. The summary results are reproduced in Table 11.1. In addition to the distinction already drawn between applicants and recipients (defined, it should be recalled, differently for single- and two-parent families), the final analysis differentiates between single-parent families living in urban and rural counties. Costs of service delivery were greater for rural families.

[Table 11.1 about here]

The template analysis distinguishes four different perspectives on the experimental outcomes and recognizes a variety of effects, not all commensurable. The first is that of families receiving the experimental treatment—the “welfare sample.” For them, the gains, compared to controls, are positive. Gains are measured over a five-year horizon, discounted to year one. Thus, at the end of the first year the sum of income gains in the first year of random assignment plus discounted (at 5 percent) income gains in years 2 to 4 is $10,222 in 1996 dollars for urban single parent long-term recipients, $521 for two-parent applicant families. The second

perspective is that of the combined federal and state government budget, basically the combination of incremental transfer and services costs attributable to the intervention minus net additional taxes paid by treatment group families. “Taxpayers” in the analysis is everyone except those in the experimental group; taxpayers differ financially from government only because government counts net change in business payroll tax contribution as a benefit, while this is not included for taxpayers. The social perspective is the sum of net gains of the welfare sample and taxpayers. It is only for the long-term urban recipients sample that benefits exceed costs; for families in the two-parent applicant sample, MFIP cost taxpayers a lot ($12,173 per treatment group member) for insignificant ($512) financial gain. Both financial benefits and costs are measured quite comprehensively.

“Total financial gains and losses” are just that: financial. The MFIP analysis recognizes that public support for transfer programs involves many qualitative considerations. The bottom section of the table identifies qualitative outcomes produced by the MFIP treatment. Thus if taxpayers define “dependency” as degree of reliance on welfare payments for family needs, MFIP may have increased the percentage of time families received welfare, but the program reduced the proportion of families reliant solely on welfare for support.

Why the MFIP Contribution is Important, and the Issues It Raises

The MFIP analysis reflects a template developed over two decades of welfare initiative analysis. That template has many desirable attributes, including random assignment design.

Nonetheless, it is important to recognize the limitations of the MFIP. First, the more comprehensive the reform effort, the greater is the difficulty in implementing both the policy change and the evaluation in a way that produces well-defined treatment and control groups. In

the MFIP case, general policy developments and a changing public perception of welfare trends served over time to alter the difference between treatment and "control."

Second, the experiment cannot capture all the consequences of general implementation. Changes in the character of assistance systems are likely to alter the number and characteristics of people seeking help and given assistance. Eligibility for MFIP was determined by AFDC standards for both the treatment and control groups; any general implementation would presumably affect both initial and on-going eligibility. This problem is particularly important for the two-parent group, where the 100-hour rule was applied in determining initial eligibility for the treatment group but not thereafter. This made the treatment an exceptional prize. This attractiveness is possibly reflected in the substantial difference (72.1 percent for MFIP, 59.7 percent for AFDC-UP controls) in welfare take-up rates following random assignment (Miller et al.1997: 154). Those informed of the special opportunity proffered by the experiment may have been uncertain concerning whether should they return the offer would have still be available. In any general implementation, elimination of the 100-hour rule only post-entry would over time create a growing inequity between those families with employed principal earners fortunate enough to have gained access to the special work incentive and families in similar current circumstances but ineligible. This is the inequity that creates interest in the sorts of time-limited support policies that are proposed for testing in the Transitional Incentives Package Demonstration described earlier.

Third, while the cost-benefit analysis has internal validity provided by random assignment, the external relevance of analysis is problematic. In any full-scale implementation of a program like MFIP, the mix of applicant and on-going/long term cases would change. While it might be tempting, for example, to argue that because the efficiency of the program

seems greatest for long-term single parent cases, one has no guarantee that implementation of the program only for such cases would not alter the incentives for those in newly-opened cases to continue receipt until qualifying for the advantages of long-term status. This would raise costs.

Finally, in constructing MFIP, an important objective of Minnesota's policy makers was to simplify and better integrate support provided for working-low income households. Within the project this was accomplished by making the benefit disregard within the program more generous, eliminating the short-time horizon for work incentive in AFDC, and cashing out food stamps. While the period of MFIP implementation coincided with a substantial increase in the Earned Income Tax credit, the EITC was not part of the program and, like other taxes, EITC receipt was imputed for the MFIP cost-benefit analysis. An alternative strategy would have been to experiment with better integration of the EITC and food stamps and treating MFIP as a transitional program designed to move people out of welfare and into employed status. The result of the treatment as implemented was, not surprisingly, a tendency for those who received the MFIP prize to stay within the welfare system. The choice of strategy was largely dictated by what was under the state's control: Federal EITC policy definitely was not, and the cash-out of the Food Stamp benefit came about only by congressional authorization achieved by the state's adroit legislators. Grass-roots welfare reform efforts produced important innovations in the 1990s, but states faced then, and continue to face, important constraints in choosing alternatives.

In sum, the past quarter century has seen remarkable development in procedures for evaluation of welfare initiatives. The central achievement is more procedural than methodological, a common template for such evaluations that is founded on random assignment and incorporates cost-benefit analysis as an accounting framework. The evaluation of the Minnesota Family Investment is exemplary of this template and state welfare reform initiatives.

While the MFIP experience provides many insights pertinent to welfare strategy, significant questions remain concerning the external utility of the cost-benefit analysis.

Conclusion: Welfare AND Work

This essay has explored various possibilities for additional experimentation in matters related to the intersection of welfare and work. While the initial charge referred to "welfare *to* work" initiatives, I have argued that there are important aspects of the interplay between public assistance *and* work that call for thinking more generally.

Before moving on to new demonstrations, it might be useful to develop a template for reviewing the benefits and costs of new evaluations. Benefits are hard to predict: I doubt that few people expected to learn much about the consequences of expanded work support for marriage when the MFIP experiment was initiated. Well-run evaluations have important spillover effects. To improve the efficiency of the national system of evaluation it is important to find ways to identify these effects and to create incentives and procedures to ensure that the public interest generally is reflected in choice of next steps to be taken.

References

Autor, David & Mark Duggan (2006) The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding. NBER Working Paper No. 12436 (August). Cambridge, Mass.: NBER.

Berube, Alan, Anne Kim, Benjamin Forman & Megan Burns (2002) *The Price of Paying Taxes: How Tax Preparation and Refund Loan Fees Erode the Benefits of the EITC*. Washington, D.C.: Brookings Institution, 2002.

DeParle, Jason (2004) American Dream: Three Women, Ten Kids, and a Nation's Drive to End Welfare. New York: Viking.

Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orszag & Emmanuel Saez (2005) *Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H&R Block*. Retirement Security Project Working Paper 2005-5. Washington, D.C.: Brookings.

Eissa, Nada & Hilary Williamson Hoynes (2006) Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply. In James M. Poterba (Ed.) *Tax Policy and the Economy*, 20. Cambridge, Mass.: MIT Press, 73-110.

Gennetian, Lisa & Cynthia Miller (2000) *Reforming Welfare and Rewarding Work: Final Report on the Minnesota Family Investment Program*. Volume 2: Effects on Children. New York: Manpower Demonstration Research Corporation.

Gennetian, Lisa, Cynthia Miller & Jared Smith (2005) *Turning Welfare into a Work Support: Six-Year Impacts on Parents and Children from the Minnesota Family Investment Program*. New York: MDRC.

Giannarelli, Linda & Michael Wiseman (2006) The Working Poor and the Benefit Doors: Entrance, Exit, and Employment Incentive Policy in TANF. Paper presented at the 46th Annual Workshop,National Association for Welfare Research and Statistics, Jackson, Wyoming, 21-23 August.

Government Accountability Office (2004) *TANF and SSI: Opportunities Exist to Help People with Impairments Become More Self-Sufficient*. Report GAO-04-878. Washington, D.C.: GAO.

Grogger, Jeffrey (2004) Welfare Transitions in the 1990s: The Economy, Welfare Policy, and the EITC. *Journal of Policy Analysis and Management* 23(4), 671-695.

Hamilton, Gayle (2002) *Moving People from Welfare to Work: Lessons from the National Evaluation of Welfare-to-Work Strategies.* New York: Manpower Demonstration Research Corporation.

Mead, Lawrence (2006) Did EITC Raise Work Levels? If So, How? Annual Research Conference of the Association for Public Policy Analysis and Management, November 3, Madison, Wisconsin.

Michalopoulos, Charles (2005) *Does Making Work Still Pay? An Update on the Effects of Four Earnings Supplement Programs on Employment, Earnings, and Income*. New York: Manpower Demonstration Research Corporation.

Miller, Cynthia, Virginia Knox, Patricia Auspos, Jo Anna Hunter-Manns & Alan Orenstein (1997) *Making Welfare Work and Work Pay: Implementation and 18-Month Impacts of the Minnesota Family Investment Program*. New York: Manpower Demonstration Research Corporation.

Miller, Cynthia, Virginia Knox, Lisa A. Gennetian, Martey Dodoo, Jo Anna Hunter & Cindy Redcross (2000) *Reforming Welfare and Rewarding Work: Final Report on the Minnesota Family Investment Program.* Volume 1: *Effects on Adults*. New York: Manpower Demonstration Research Corporation.

Munnell, Alice (Ed.) (1986) *Lessons from the Income Maintenance Experiments*. Boston: Federal Reserve Bank of Boston.

Nadel, Mark, Steve Wamhoff, and Michael Wiseman (2003/2004) Disability, Welfare Reform, and Supplemental Security Income. *Social Security Bulletin* 65(3), 14-30.

Social Security Advisory Board (2006) *A Disability System for the 21st Century*. Washington, D.C.: The Board.

Walker, Robert & Michael Wiseman (2006) Opening Up American Federalism: Improving Welfare the European Way. Unpublished manuscript.

Wamhoff, Stephen & Michael Wiseman (2005/2006) The TANF-SSI Connection. *Social Security Bulletin* 66(4), 21-36.

Wiseman, Michael (1991) Research and Policy: A Symposium on the Family Support Act of 1988. *Journal of Policy Analysis and Management* 10(4), 588-666.

Table 11.1

Financial and Nonfinancial Gains and Losses per MFIP Group Member, by Subgroup and Accounting Perspective (in 1996 Dollars)

Reproduced from Miller et al 2000, Table 7.1

	Single-Parent Long-Term Recipients			Single-Parent Recent Applicants			Two-Parent Families	
Perspective	Urban	Rural	Total(a)	Urban	Rural	Total(a)	Recipients	Applicants
Total financial gains and losses () over five years								
Welfare sample	$10,222	$9,301	$9,891	$5,967	$10,477	$7,762	$6,855	$521
Government budget	($8,465)	($12,068)	($9,762)	($8,122)	($11,912)	($9,630)	($19,147)	($12,762)
Taxpayers	($8,678)	($12,113)	($9,915)	($8,111)	($12,008)	($9,662)	($18,669)	($12,173)
Society	$1,545	($2,812)	($24)	($2,144)	($1,531)	($1,900)	($11,814)	($11,652)
Non-financial effects over observation period (from welfare sample perspective)b								
Work, welfare, and income per quarter[b]								
Percentage with income below poverty	↓	↓	↓	↓	↓	↓	↓	0
Percentage working	↑	↑	↑	↑	↑	↑	0	0
Welfare use								
Percentage receiving welfare	↑	↑	↑	↑	↑	↑	↑	↑
Percentage relying solely on welfare	↓	↓	↓	↓	0	↓	0	0
Other family outcomes								
Continuous health coverage[c] (%)	↑	0	↑	↑	↑	↑	0	n/a
Homeownership[d] (%)	0	0	0	0	↓	0	↑	n/a
Mother currently married and living with spouse[e] (%)	0	0	↑	0	0	0	↑	n/a
Time spent out of the home[f]	↑	0	↑	↑	0	↑	0	n/a
Child environment and child well-being (measured for families with children age 2-9)[g]	↑	n/a	n/a	0	n/a	n/a	n/a	n/a

SOURCES: MDRC calculations using data from Minnesota's Unemployment Insurance (UI) and public assistance benefit records, the 36-month client survey, state and federal tax codes, aggregate fiscal data, and county child care payment records. See Gennetian and Miller(2000).

NOTES: The arrows on this table reflect positive and negative statistically significant effects. Outcomes indicated as n/a are not measured. A more in-depth explanation of these impacts can be found in Miller et al (2000) and Gennetian and Miller (2000).

[a]Total gains and losses were estimated as a weighted average of urban and rural results, based on urban and rural proportions in total caseloads of the seven field trial counties.
[b]Average quarterly during the follow-up period.
[c]Percentage who had continuous health insurance coverage from random assignment through time of the 36-month survey.
[d]Percentage who owned their home at the time of the 36-month survey.
[e]Percentage married and living with spouse at the time of the 36-month survey.
[f]Measured on 36-month survey as average hours worked per week in current or most recent job. For two-parent families, measured only for the survey respondent (usually the mother).
[g]Summary of full MFIP impacts on domestic abuse and on behavior and school outcomes for children age 2-9 at random assignment. For urban long-term recipients, MFIP produced statistically significant impacts on domestic abuse and on children's behavior and school performance. For urban recent applicants, MFIP produced few statistically significant impacts on child wellbeing. For single-parent families in rural counties and for two-parent families, the results are not reported due to small sample